September 19, 2024

Neighborhood Change Indicators

HUD PD&R Contract #: 86614623c00008

URBAN INSTITUTE · ELEVATE · THE · DEBATE

Alena Stern and Claudia Solari

Agenda

- Background & motivation
- Types of neighborhood change measured
- Input data for predictions
- Results and machine learning models
- Limitations & next steps
- Audience thoughts and Q&A
  - How might you use this in your work?

Background & Motivation

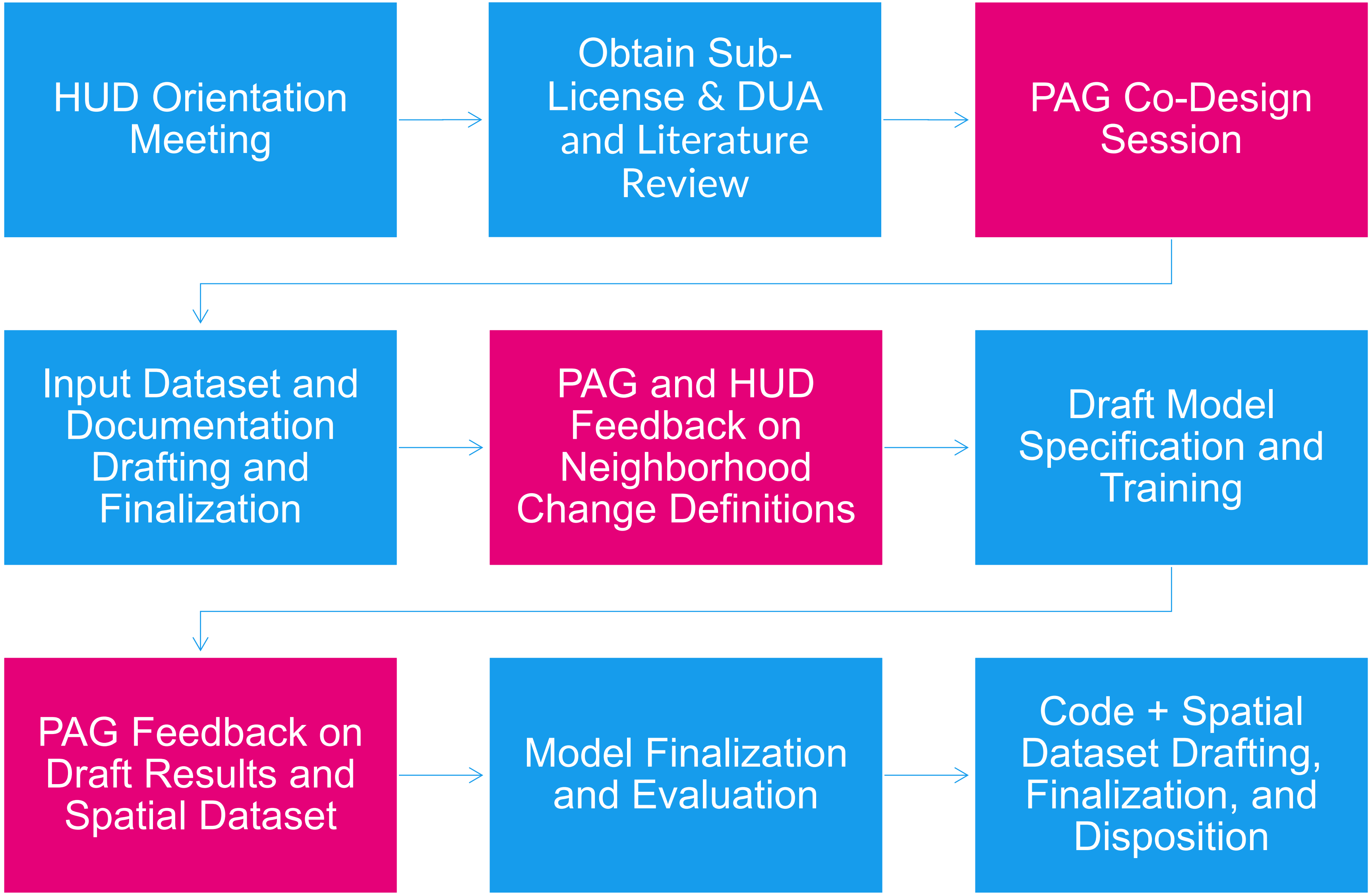
Background & Motivation

- Stern 2021
  - Machine learning models of neighborhood change on 4 metro areas in DC and OH
- Expand
  - Nationwide
  - Data sources beyond HUD Administrative data & ACS
  - Incorporate learnings from pilot and user engagement

User Engagement

- Project Advisory Group (N=11)
  - Geographic coverage: Denver, Phoenix, Charlotte, New Orleans, Philadelphia, Rochester, Detroit, DC, rural areas
  - Various roles: Lived experience of neighborhood change, local practitioners/advocacy groups, NNIP affiliates, local government, scholar
- Subject Matter Experts
  - Rural areas, tribal areas, climate impacts, historical preservation, equity, fairness

Implementation Process



Types of Neighborhood Change

Types of Neighborhood Change - Goals

- Provide actionable information
  - Prevent displacement of incumbent residents
  - Target resources to mitigate displacement
  - Address the “missing middle” from outmigration
- Not stereotyping/stigmatizing communities
  - Act of change (e.g., population loss) vs label (e.g., declining)
- Understandable definitions
- Split data into rural and urban subgroups for modeling

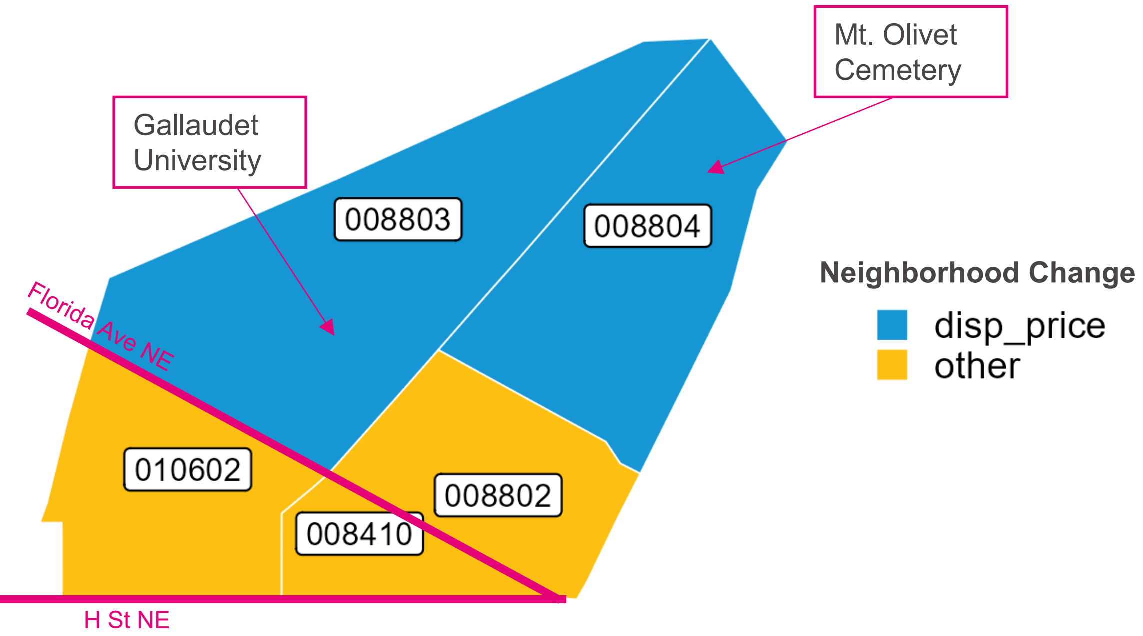
Selected Types of Neighborhood Change (% tracts 2022)

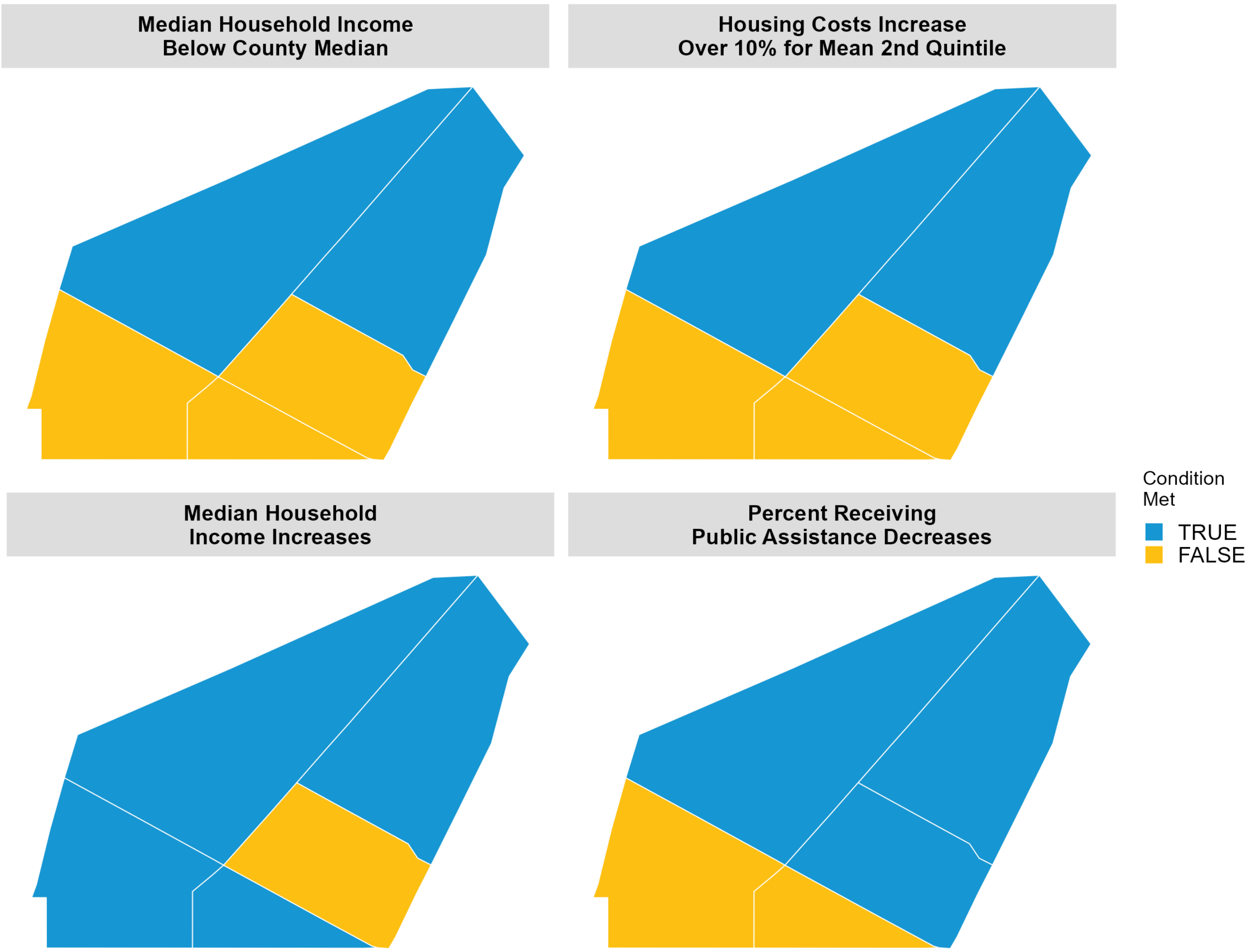
- Displacement due to price pressures (only urban areas) (2.2%)
  - Lower income at period start, increased housing cost, increased income, decreased public benefits
- Population loss due to economic disinvestment (4.2%)
  - Decreased households, decreased income, decreased college-educated
- Inclusive growth (2.2%)
  - Lower income at period start, increased income, stable housing cost, stable voucher tenants, increased population
- Mutually exclusive, with “Change not measured” (91.4%)

Time Periods of Change and Prediction

Period of Change Predicted	Date Prediction Made
2017-2022	December 2021
2016-2021	December 2020
2015-2020	December 2019
2014-2019	December 2018
2013-2018	December 2017

Neighborhood Change Classifications in Trinidad/NoMa





Input Data for Prediction

Neighborhood Change Dimensions and Data Sources

- Demographic composition (e.g., race, ethnicity, age)
- Income and education composition (e.g., education, household income)
- Land and dwelling use (e.g., vacancy, tenure, housing choice voucher tenants)
- Cultural and Institutional characteristics (e.g., libraries, historic sites)
- Quality of housing stock (e.g., housing costs, mortgage activity, disaster shocks)
- Economic investment (e.g., jobs, commute time)

Feature Generation

- Create new variables capturing changes prior to prediction date
  - Change in variable in the given tract
  - Change in variable in the neighboring tracts
  - Change in the rate of change of a variable in the given tract
  - Consistency of change in a variable in the given tract
- Drop variables that have no variation, high degrees of missingness after imputation, and are highly correlated with each other
- From 449 raw variables, results in 261 (rural) and 249 (urban) features for training

Time Periods of Prediction and Data Availability

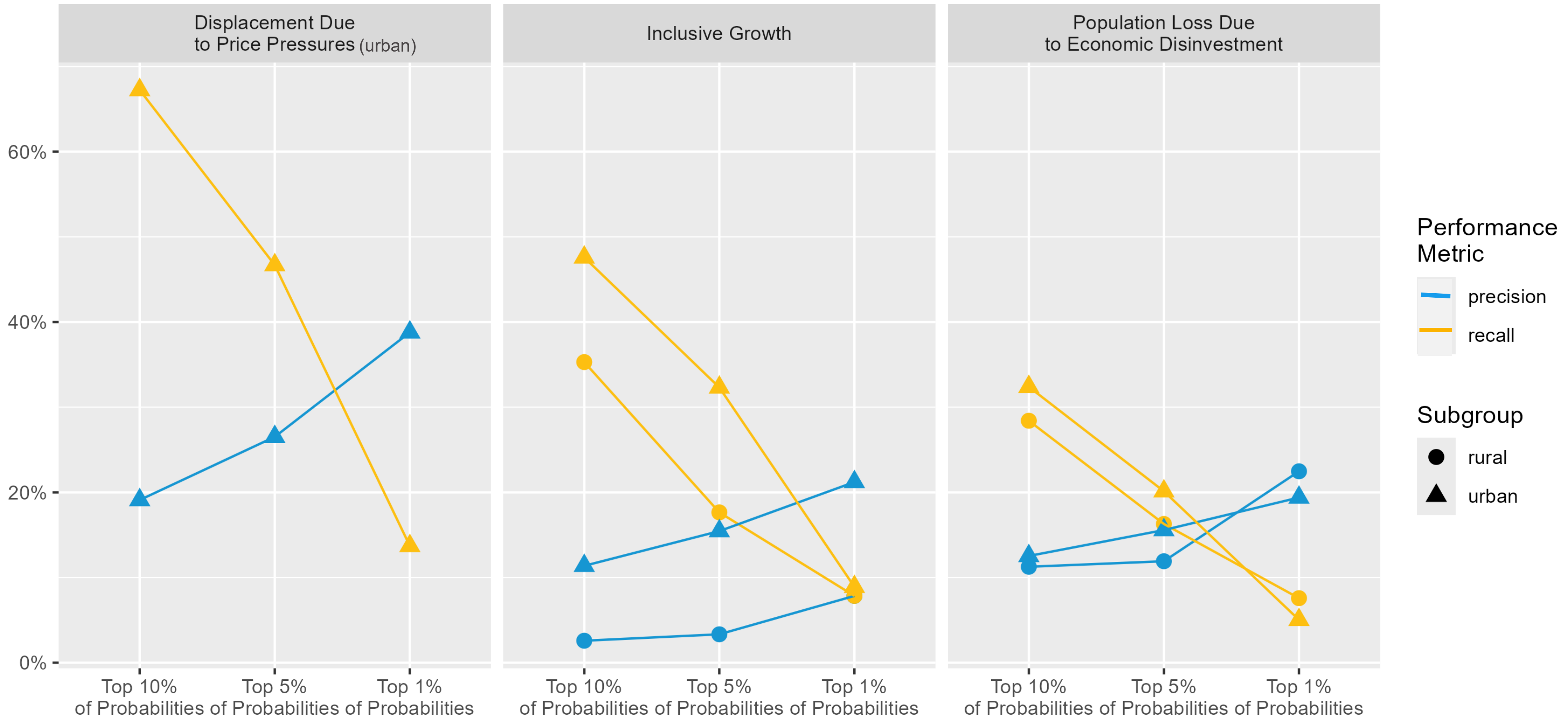
Period of Change Predicted	Date Prediction Made	Most Recent ACS 5-year Available for Features
2017-2022	December 2021	2020
2016-2021	December 2020	2019
2015-2020	December 2019	2018
2014-2019	December 2018	2017
2013-2018	December 2017	2016

Results and Machine Learning Models

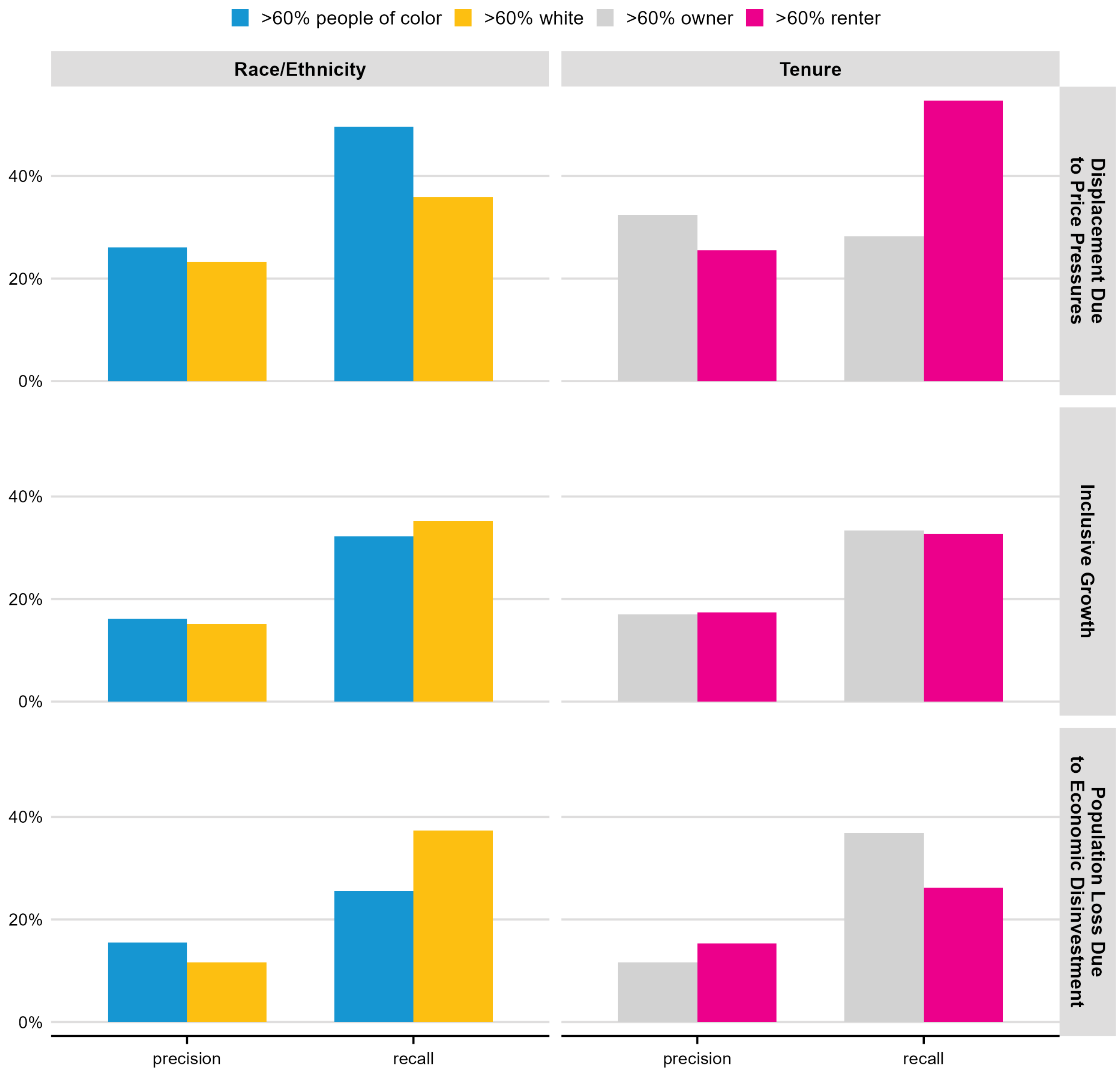
Modeling Approach

- Randomly split tracts into training and testing sets (within each subgroup)
- Try many combinations of model types and specifications on the training data for each subgroup
- Evaluate the performance and fairness of the models
- Select the best-performing model for each subgroup and apply to the test data

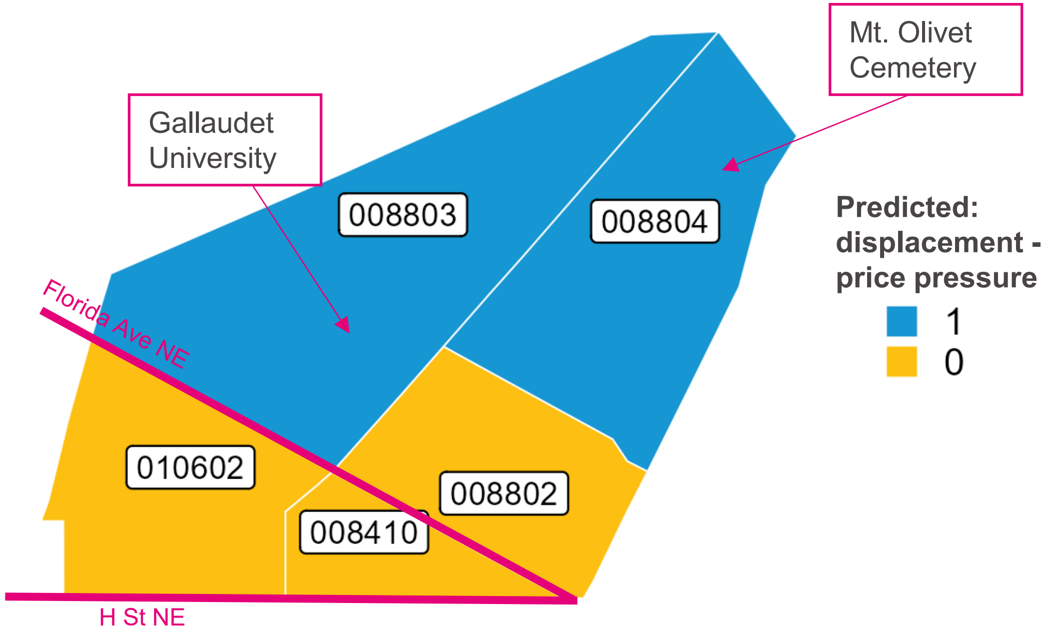
Modeling Results on Test Data for 2022

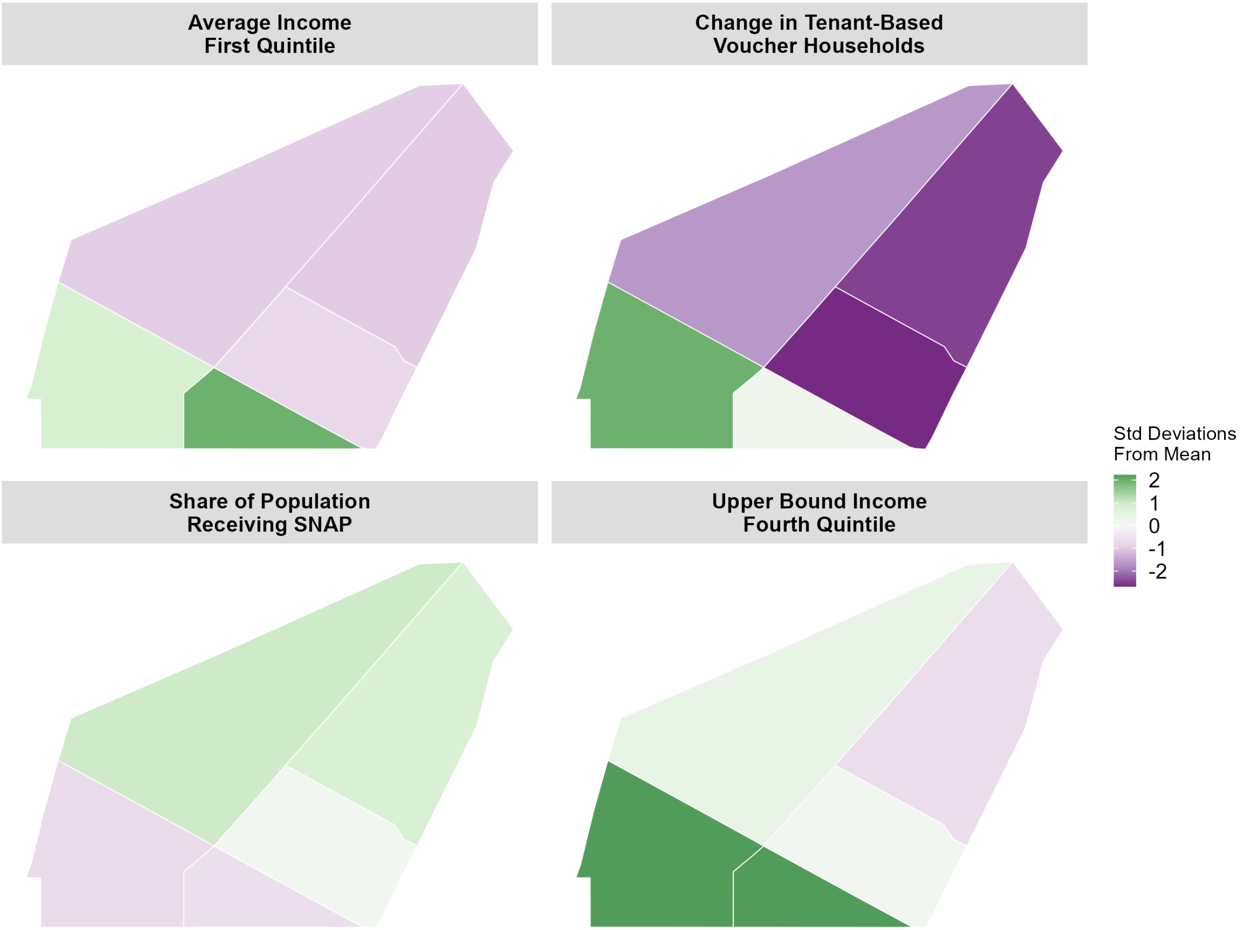


Fairness Metrics on Test Data for Urban Tracts in 2022



Neighborhood Change Classifications





*At top 5% threshold

Limitations & Next Steps

- Limitations
  - National focus excluded useful local data
  - Include more datasets
  - Test more model variations
- User guide
  - Conversation-starter
  - How to use the information on neighborhood change

Recommendations

- Communicate the change type & the probability of change
- Show multiple years
- Allow user feedback
- Overlay more data for deeper context (e.g., climate change, transportation, rural neighborhood assets, R/ECAP)

Areas for Further Research

- Explore other modeling set-ups to enhance model's ability to target changing tracts (e.g., continuous change, change relative to county, only include "eligible" tracts, binary classification)
- Test other modeling approaches and specifications (e.g., deep learning models)
- Explore other subgroup divisions (e.g., rural, suburban, urban)
- Add more datasets &variables tailored to subgroups and types of change
- Test prediction of change further out than 1 year
- Explore persistently unchanging neighborhoods

Audience Thoughts and Q&A

Appendix

Displacement due to price pressures (only urban areas)

- Median household income is below county median household income in the starting year, suggesting a population at risk of displacement.
- Monthly median housing costs as a share of mean second quintile income in the start year increases by at least 10 percent, suggesting housing price pressures.
- The proportion of the tract's population using public benefits decreases, suggesting a displacement of the lowest income community members.
- The tract's median household income increases, suggesting a displacement of lower income households by higher income households.

Population loss due to economic disinvestment

- The proportion of the population with a bachelor's degree or higher declines, suggesting a loss of the more highly educated population.
- Median household income declines by at least 5 percent, suggesting a decline in economic outcomes.
- A decrease in the number of households, suggesting population loss.

Inclusive growth

- Median household income is below the county median household income in the starting year, suggesting a sufficient low-income at the start point.
- The inflation-adjusted (to 2022 dollars) income for the first and second quintile of residents both increase, suggesting that income increases are shared by the lower quintiles of the income distribution.
- Monthly median housing costs as a share of mean second quintile income in the start year increases by less than 5 percent, suggesting limited price pressures.
- The number of tenant-based voucher holders does not decline during the period, suggesting that landlords are not refusing to rent to voucher holders.
- The number of households grows during the period, indicating population growth.

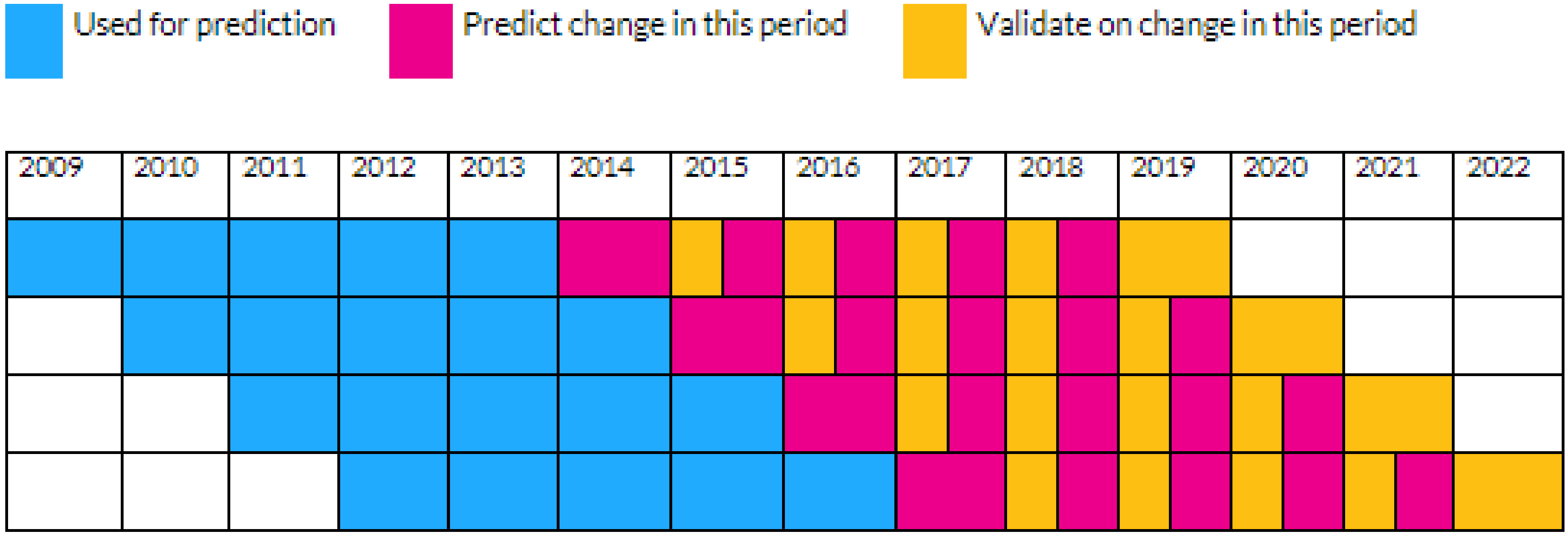
Why we determined these definitions?

- Heavily informed by the literature
  - But, the literature is inconsistent in how it operationalizes neighborhood change types
- Wanted to focus on the process rather than loaded phenomena that can add complexity and nuance
  - Definitions can vary by locality but ours needs to apply more broadly across the county.
- Have simple definitions that are understandable

Prevalence of Neighborhood Change by Type

Change Type	2018	2019	2020	2021	2022
Displacement Due to Price Pressures	1.2%	1.9%	2.4%	2.4%	2.2%
Population Loss Due to Disinvestment	4.2%	3.3%	4.0%	3.9%	4.2%
Inclusive Growth	2.4%	2.4%	2.2%	2.0%	2.2%

Model Training



Input Data Sources

Neighborhood Change Factor	Relevant Datasets (Variables Measuring Factors)
Demographic composition	▪ American Community Survey (race, ethnicity, age, foreign born, languages spoken)
Income and education composition	▪ American Community Survey (household income, educational attainment, health insurance coverage)
Land and dwelling use	▪ American Community Survey (tenure, units in structure) ▪ HUD USPS vacancy data (residential and business vacancy, no-stat, and active addresses) ▪ HUD administrative data (housing choice voucher tenants and projects)
Cultural and institutional characteristics	▪ IMLS Public Libraries Survey (library openings and closures) ▪ National Register of Historic Places (historic preservation designation)

Input Data Sources

Neighborhood Change Factor	Relevant Datasets (Variables Measuring Factors)
Quality of housing stock	▪ Home Mortgage Disclosure Act (mortgage activity, home prices). ▪ American Community Survey (median gross rent, home value, number of bedrooms, ratio income to home value) ▪ Comprehensive Housing Affordability Strategy (housing problems). ▪ FEMA Disaster Data (disaster shocks)
Economic Investment	▪ LEHD Origin-Destination Employment Statistics (Jobs/workers at different income points). ▪ American Community Survey (broadband access, commute time, employment rate)

Input Data Source Publication Lag

Data Source	Earliest Year Available	Latest Year Available*	Publication Lag (Years)
ACS 5-Year	2013	2022	1
HUD USPS	2008	2023	0
HUD Administrative	2008	2023	0
HMDA	2013	2022	1
FEMA	2009	2022	1
LODES	2011	2021	2
IMLS PLS	2012	2022	1
HUD CHAS	2011	2020	3
National Register of Historic Places	2014	2022	1

***As of December 2023**

Most important features: urban model

- **Previous neighborhood change:**
  - inclusive growth, displacement due to price pressures, depopulation due to disinvestment
- **Income**
  - 80th percentile income, previous change in household income, mean first quintile income, 100th percentile income, 2nd quintile share of aggregate income
- **Education**
  - Change in the share with a bachelor's degree
- **Housing**
  - median mortgage loan amount, previous change in number of households, number of owner-occupied units, change in housing costs, proportion of owner-occupied units that are in 1-unit buildings
- **Public assistance**
  - change in proportion receiving public assistance, change in the number of tenant-based voucher households, share of population receiving SNAP

Most important features: rural model

- **Previous neighborhood change:**
  - inclusive growth, depopulation due to disinvestment
- **Income**
  - Whether the tract is below county median income, change in the share below the federal poverty limit, change in median household income, mean and upper bound of first income quintile, share of total income held by third quintile, 100^{th} percentile income
- **Education:**
  - Share with education more than a bachelor's degree, share with a bachelor's degree
- **Housing**
  - median housing cost, previous change in number of households, share of units built before 1960, change in housing costs, proportion of owner-occupied units that are in 1-unit buildings, count of active residential addresses
- **Economic:**
  - Number of Hispanic workers