



Seattle kids: changing demographics and the implications for public schools

Laura Minter

December 2021

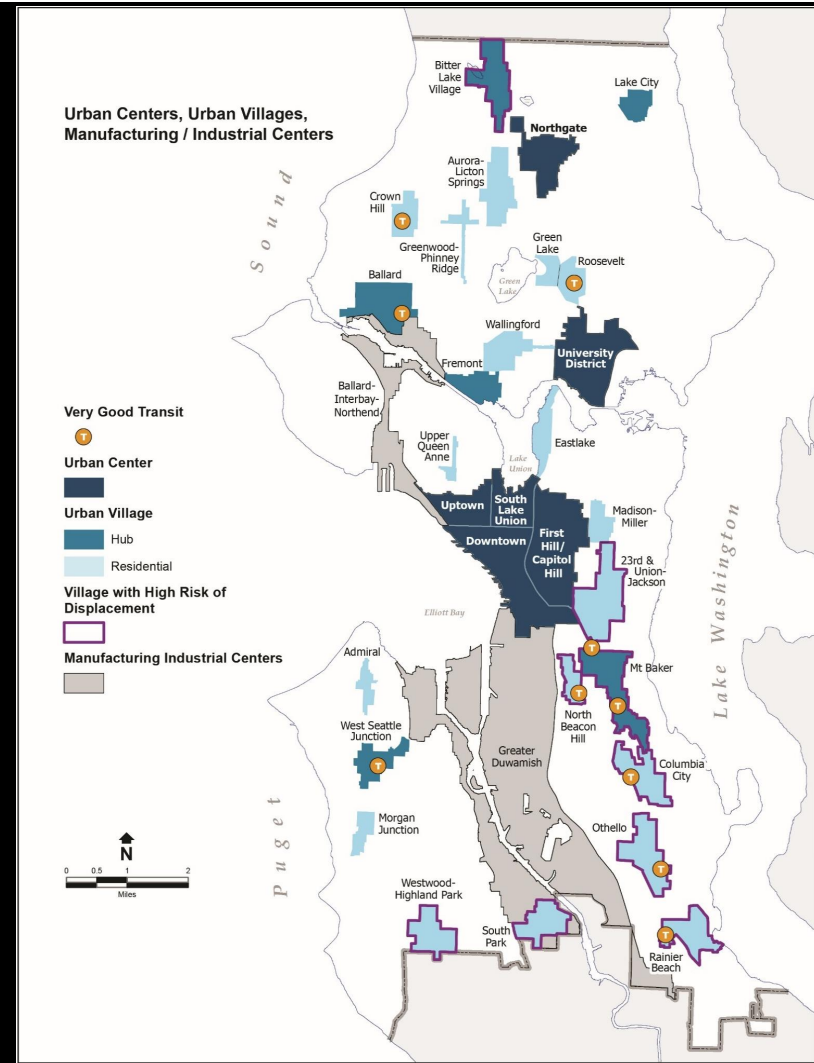
Background

Explosive population growth
and expensive real estate

Urban Centers and Villages
development plan

Debates over children and
families in urban environments

Background



Problem statement

What role does density
play in the distribution
of children in Seattle's
Urban Centers and
Villages?

What role do Urban
Centers and Villages
play in changing
enrollments in Seattle
Public Schools?

Can we model
enrollment using
current population data?

Data Collection



Census data
aggregated by UCV
and by region

Source: City of
Seattle
Format: CSV files



Enrollment data by
individual school

Source: Seattle
Public Schools
Format: pdf reports



Geographical
correspondence

Source: Created by
comparing maps
from city and
schools
Format : Excel file

Data Dashboards

- [Click me \(I'm not sketchy at all!\)](#)

Exploratory Data Analysis: Key Findings



Centrally located UCVs are experiencing the steepest rise in child population

Very urban, essentially no single-family housing



Inconsistencies in UCV boundaries leads to interesting density trends

Uptown (Lower Queen Anne) is low density while Upper Queen Anne has 2nd highest density



UCVs split to balance school enrollments

Othello split between 6 elementary schools
Schools do not serve them as a whole

Modeling: enrollment



Enrollment Data

By school, grade



Population Data

Local (UCV) and
citywide
Time lags



Annual Timescale

Linear modeling

Challenges

01

Reworked school boundaries with UCVs split across schools

02

Different reporting frequencies

03

Aggregated by age:
Over/Under 18 only

Challenges

Feasible approach:
Model total enrollment using
interpolated population data

01

Reworked school
boundaries with
UCVs split across
schools

02

Different reporting
frequencies

03

Aggregated by age:
Over/Under 18 only

Modeling enrollment

Simple Linear Regression

- Total and grade level enrollment
- Population data

Multiple Linear Regression

- Total and grade level enrollment
- Population data
- Time lagged population data

Null Model

- Enrollment average

Results: how much variation can our models explain?



Metric: R2 score

What fraction of variation is explained?



Best total enrollment model:

total enrollment

score: 91.3%

population and 5-year lag



Best grade level model:

middle school enrollment

score: 98.5%

population and 10-year lag

Recommendations

- Re-examine how we define school assignments in relation to urban centers and villages, particularly Othello
- Consider rising population of children in urban centers and villages when determining how to manage school assignments and plan for capital investments

Next steps

Housing and \$

Housing and economic indicators

- American Community Survey data
- Would need to be aggregated by UCV

Age breakdown

More detailed information on age distribution

- Get breakdown by age not aggregated as child/adult
- Aggregate census tracts to get UCV data

Distance

Distance data

- How long do students commute?
- What does the commute look like based on neighborhood?

Acknowledgments

Thanks to DSI-927!
It has been a wild ride!

Thanks to my husband and
daughter for letting me obsess
over this project and listening
to my endless hypothesizing!