Minneapolis 2040 Housing Indicators: Technical Appendix

This technical appendix describes the synthetic control approach that we use to generate counterfactual Minneapolis outcomes and estimated statistical significance (p-value) in the accompanying dashboard. The appendix also describes how we construct each measure featured in the dashboard, including both core indicators (for which we typically produce counterfactuals) and contextual figures.

The Synthetic Control Framework

The Minneapolis 2040 Comprehensive Plan (the Plan) comprises an array of policies that affect different economic and social outcomes in the City of Minneapolis (the City). Consequently, we will track a variety of City outcomes for many years after 2020, when the Plan took effect. Rather than attempt to disentangle the effects of disparate elements of the Plan—from each other and from simultaneous unrelated developments—we assess the Plan in its entirety, focusing on the housing-related impacts.

We do this using the synthetic control method (SCM), a methodology that allows the construction of counterfactual Minneapolis outcomes. The methodology was devised by Alberto Abadie and coauthors, deployed most famously in Abadie, Diamond, and Hainmueller (2010). It has quickly become a staple of program evaluation and an invaluable tool for assessing policy effects when experimental designs are not feasible.

We begin the description of our synthetic control implementation with the standard "potential outcomes" notation used in program evaluation to characterize treatment effects. The effect of the Minneapolis 2040 plan in period t can be described as:

$$\tau_t = Y_t^I - Y_t^N$$

where Y_t^I is the observed outcome for Minneapolis affected by the policy implementation, Y_t^N is the potential outcome in the absence of the policy for Minneapolis, and $t > T_0$ where T_0 is the period before the policy implementation.[1]

[1] t may be of annual or monthly frequency depending on the data availability. For some indicators (e.g. Income Segregation Index), although the data are presented and inputted as annual data, they are calculated using three years or five years of data.

Y_t^N is estimated using a synthetic control, defined as a weighted average of donor unit outcomes. The pool of donor units consists of other cities in the U.S., subject to the following criteria[2]:

1. Total population between 150,000 and 2,000,000
2. No similar policy in effect around the same time (e.g., Portland, OR)
3. No cities that may experience spillover effects of the Minneapolis 2040 plan (e.g., St. Paul, MN)
4. Principal cities only (e.g., excluding Irving, VA)
5. Cities that are a census place during the whole study period (2010-current) (e.g., excluding Urban Honolulu census designated place, HI)

After removing cities that do not meet these criteria, we retain $J = 126$ donor cities across the U.S.

The synthetic control estimator of Y_t^N is then:

$$\hat{Y}_t^N = \sum_{j=1}^{J} w_j Y_{tj} \tag{1}$$

where j denotes each of the donor cities, $w_j \in [0,1]$ is the weight for city j and $\sum_{j=1}^{J} w_j = 1$, and Y_{tj} is the observed indicator for city j at time $t \in (T_0 + 1, \dots, T)$, i.e., during post-treatment periods.

Let $\boldsymbol{W}^* = (w_1^*, w_2^*, \dots, w_J^*)'$ be a vector of the synthetic control weights that minimizes:

$$\|\boldsymbol{X} - \boldsymbol{X_0}\boldsymbol{W}\| = \left(\sum_{h=1}^{k} v_h \left(X_h - w_1 X_{h1} - w_2 X_{h2} - \cdots - w_J X_{hJ}\right)^2\right)^{\frac{1}{2}} \tag{2}$$

$$\text{subject to } \sum w_j = 1 \text{ and } 0 \leq w_j \leq 1$$

where v_h is a given set of non-negative predictor weights, X_h is predictor h for Minneapolis, X_{hj} is predictor h for donor city j, and k is the number of predictors.

Before $\boldsymbol{W}^*$ can be estimated, we need to find $\boldsymbol{V} = (v_1, \dots, v_k)$ such that $\boldsymbol{W}(\boldsymbol{V}) = (w_1(\boldsymbol{V}), \dots, w_J(\boldsymbol{V}))'$ minimizes equation (2) and the weights in $\boldsymbol{W}(\boldsymbol{V})$ are positive and sum to one.

$\boldsymbol{V} \in (0,1)^m$, where m is the total number of predictors, is estimated based on a given set of predictors. In this particular project, the predictors include:

- Average of natural logarithm of total population for each city (Minneapolis and all donor cities) over $\{1, 2, \dots, T_0\}$

---

[2] All estimates are based on ACS 2014-2018 data and cities are Census places.

- Average of percent of owner-occupied housing units for each city over $\{1, 2, \ldots, T_0\}$
- Average of natural logarithm of median housing cost for each city over $\{1, 2, \ldots, T_0\}$
- Average of natural logarithm of median household income for each city over $\{1, 2, \ldots, T_0\}$
- Odd-year lagged values of the observed indicator (e.g. if $\{1, 2, \ldots, T_0\}$ = {2010, 2011, …, 2019}, the values of the indicator at years 2011, 2013, 2015, 2017 and 2019 are included in the algorithm as predictors; if $\{1, 2, \ldots, T_0\}$ is of monthly frequency, then the values of the indicator in December of those odd years would be included as predictors)

These variables are selected such that they are general enough to help predict many of the indicators we consider. The lagged indicator values are particularly important, given that they capture unobserved determinants of a given city's indicator.

We estimate the initial values of $\boldsymbol{V}$ using two methods:

1. Equal weights: $\boldsymbol{V} = (\frac{1}{k}, \ldots, \frac{1}{k})$
2. Regression-based:
   a. Solve for $\boldsymbol{\beta}$: $(X'X)^{-1}X'Z'\boldsymbol{\beta} = \boldsymbol{0}$ where X is a $(J+1) \times (k+1)$ matrix with an intercept column and k columns of predictors as listed above, and Z is a $T_0 \times (J+1)$ matrix where each column is the observed indicator value for Minneapolis and each donor city
   b. Remove the intercept term from $\boldsymbol{\beta}$
   c. Then, $\boldsymbol{V} = \boldsymbol{\beta}\boldsymbol{\beta}'$

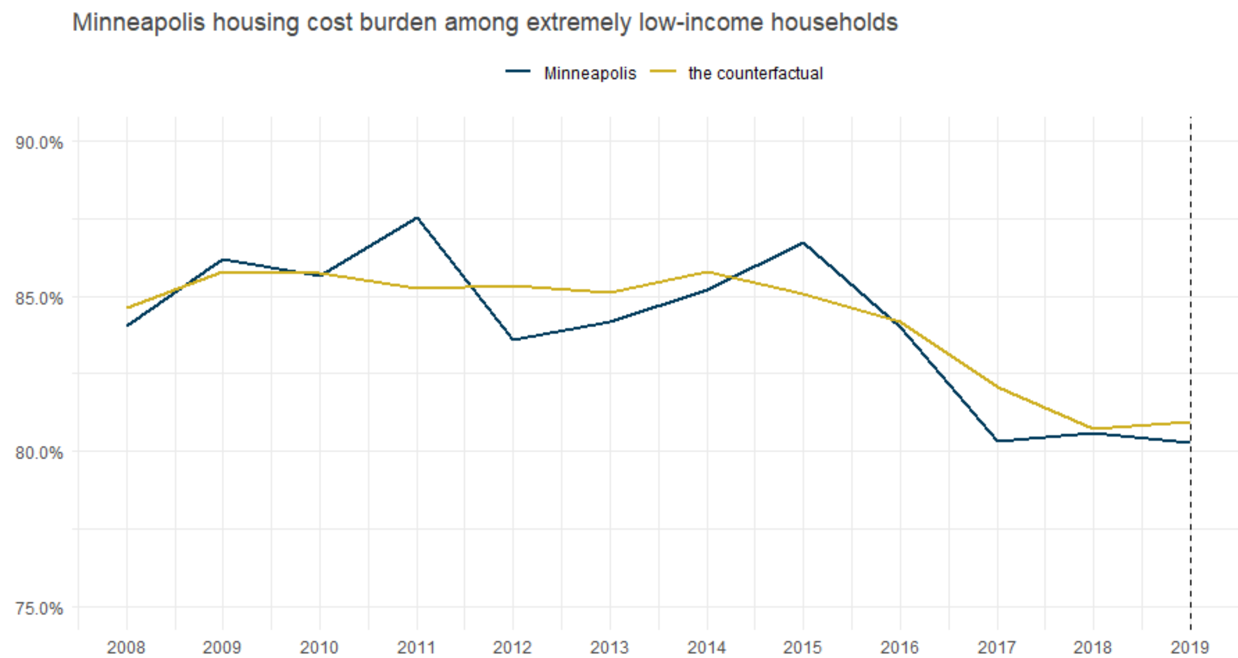
For each of the two potential initial values for $\boldsymbol{V}$, we find the corresponding $\boldsymbol{W}$ by solving the quadratic programming problem:

$$\boldsymbol{W} = \text{argmin} \left\{c'w + \frac{1}{2}w'Hw\right\} \qquad \text{subject to} \qquad \sum w = 1 \text{ and } 0 \leq w \leq 1 \qquad (3)$$

where $c = -X_1' V_{\text{diag}} X_0$ where X_0 is a $k \times J$ matrix containing predictor values for all J donor cities, X_1 is a $1 \times k$ matrix containing predictor values for Minneapolis, and V_{diag} is a diagonal matrix with diagonal values $\boldsymbol{V}$, and $H = X_0{}'V_{\text{diag}}X_0$.

The loss is then calculated as $\frac{1}{|\{1,2,\ldots,T_0\}|} \|\boldsymbol{X} - \boldsymbol{X_0}\boldsymbol{W}\|$. The $\boldsymbol{V}$ that results in the lowest loss—i.e., whichever of methods (1) or (2) performs best—is $\boldsymbol{V}^*$, which then implies a corresponding $\boldsymbol{W}^*$ that solves equation (3).

Finally, the pre- and post-treatment counterfactual outcome of interest can be estimated using equation (1). The figure below shows an example—housing cost burden among extremely low-income households—with observed Minneapolis values in blue and the counterfactual in yellow.



Assessing the Significance of Differences between Actual and Counterfactual Outcomes in the Post-treatment Period

It is important to distinguish genuine impacts of a policy from those differences in actual and counterfactual values that are likely due to chance. However, distinguishing between significant and insignificant effects in an SCM context is somewhat more complicated than it is for simpler econometric methods. The approach we take in this dashboard is to compare the model's fit in the post-treatment period to the fit in the pre-treatment period. Specifically, the root mean squared error (RMSE) of the model can be calculated in the two periods and then the ratio of the post-RMSE to pre-RMSE values can be constructed. Intuitively, the counterfactual should tightly fit the actual series in the pre-treatment period (producing a low value for the RMSE), while the fit for the treated unit in the post-treatment period should be worse if the policy has an effect (yielding a higher value for the RMSE). The resulting ratio of the post-to-pre RMSE values should also increase as the effect of the treatment becomes more pronounced. To construct a p-value for Minneapolis, we implement this process for every donor city in our sample.

For $0 \leq t_1 \leq t_2 \leq T$ and $j = \{1, \ldots, J+1\}$, the RMSE at periods $t_1, \ldots, t_2$ is calculated as:

$$R_j(t_1, t_2) = \left(\frac{1}{|\{t_1, \ldots, t_2\}|} \sum_{t=t_1}^{t_2} \left(Y_{jt} - \hat{Y}_{jt}^N \right)^2 \right)^{\frac{1}{2}}, \quad (4)$$

where $\hat{Y}_{jt}^N$ is the outcome on period t produced by a synthetic control when donor city j is coded as treated and using all other J donor units (excluding Minneapolis) to construct the donor pool. Then, the post-to-pre RMSE ratio is calculated as:

$$r_j = \frac{R_j(T_0 + 1, T)}{R_j(1, T_0)} \tag{5}$$

Because the donor cities have not undergone treatment, we expect the variation between the actual data and the synthetic controls to be random and the resulting RMSE ratios to be smaller than the value obtained for Minneapolis (if the policy has had an effect). The calculated r_j values can then be sorted and the position of Minneapolis, relative to the other donor cities, can be determined and reported as an empirical p-value. Specifically, the p-value based on the permutation distribution of r_j is calculated as:

$$p = \frac{1}{J+1}\sum_{j=1}^{J+1} I_+\big(r_1 - r_j\big), \tag{6}$$

where I_+ returns one if $r_1 \geq r_j$ and zero otherwise, and r_1 is the post-to-pre RMSE ratio for Minneapolis.

For example, if the ratio for Minneapolis was the 10th highest out of the 126 donor cities and itself then the resulting p-value would be 10 / 127 = 0.079. See Abadie, Diamond, and Hainmueller (2010) or Abadie (2020) for more details.

Non-convergence and Overfitting

Synthetic control, even when limiting consideration to its classic Abadie, Diamond, and Hainmueller (2010) version, allows for different implementation choices. We now detail some of those choices and the considerations that informed them.

Lagged dependent variables

The SCM literature is not settled on the question of how much emphasis to place on lagged dependent outcomes. On the one hand, lagged values of the dependent variable capture a range of otherwise unobserved factors that matter for constructing any reasonable counterfactual. This is immediately evident in the within-sample fit improvement that is achieved when including lagged values. However, "excessive" use of lagged values relegates the time-invariant observable factors to a position of little or no influence over the counterfactual. This may lead to undesirable overfitting that reduces the model's out-of-sample accuracy, particularly if the researcher believes that control units with certain key observable characteristics deserve emphasis in the construction of the counterfactual. We therefore aimed to balance these considerations, in line with advice from Abadie and others in their recent publications, by using only odd-year lagged values of the dependent variable.

Numerical optimization

There is nothing about the Abadie, Diamond, and Hainmueller (2010) implementation of SCM that necessitates use of a specific numerical optimization routine for solving the optimization problems (see above for details of those minimization problems) that yield synthetic control weights. Unfortunately, different routines yield different solutions; these differences can be non-trivial in magnitude. (To be clear, this is not a situation that is unique to SCM applications.)

Broadly speaking, numerical routines use some combination of "guess-and-check" simplex methods and gradient-based approaches that necessitate intermediate calculations of numerical derivatives. Again speaking broadly, the former is generally considered more robust to non-smooth problems with multiple local minima, while the latter is more efficient with smoother problems.

We settled on the following approach: solve each problem using both the commonly used gradient-based BFGS method and the venerable Nelder-Mead simplex method, then select the minimum-distance solution. This approach is what the R implementation of Abadie, Diamond, and Hainmueller (2010), written by the authors of that paper, is set up to do by default.

However, in 2021 we explored many potential options for numerical optimization besides the approach described above. Those options included running the numerical methods in sequence, e.g., Nelder-Mead first and then BFGS if it does not converge, or vice versa. Our goals were a) to avoid non-convergence to the extent possible, b) to reliably find global minima, c) to generate "stable" results that did not fluctuate based on economically irrelevant matters like the ordering of the data or negligible changes in variables' values, and d) to minimize computation time.

One note that applies throughout is that while the weights generated by a particular approach may differ visibly from the weights generated with a different approach, this does not necessarily mean that the generated counterfactual will differ meaningfully. There are in general multiple combinations of donor unit weights (particularly in applications like ours with many potential donors) that achieve similar model fit and could be "reached" with different numerical methods. Ex post, different synthetic controls will sometimes have different post-treatment trajectories, but the same identification assumptions must be applied to any estimated counterfactuals in order to interpret post-treatment differences as causal impacts of the treatment. Perhaps most important is that the numerical optimization approach be settled definitively before post-treatment data is observed, eliminating the possibility of "gaming" the approach to reverse-engineer a particular desired counterfactual.

Non-convergence due to extreme values relative to donor cities

As described in more detail above, to calculate p-values associated with post-treatment differences between Minneapolis and its counterfactual, we estimate the counterfactuals for each of the donor cities. While counterfactuals are generated without incident for Minneapolis in each of our indicators, this is not always the case for the donor cities.

In addition to potential non-convergence as described above, non-convergence due to extreme values can be a concern. Ultimately, the output of the SCM procedure is the weighted average of donor cities. Because of how we selected donor cities for Minneapolis—with pre-treatment characteristics that are very roughly similar to that city— Minneapolis outcomes are not at the bottom or top of the distributions of donor cities. The same cannot be said for some of the donor cities, of course. And when constructing counterfactuals for donor city outliers, the SCM procedure can fail to generate corresponding donor city weights. For example, if the housing cost burden for Cincinnati is 10 percent but it is between 15 and 50 percent for other donor cities, then there are no donor city weights that could yield a weighted average of 10 percent. This is because one of the constraints in our optimization is that weights have to add up to 1 and must be non-negative; in other words, the procedure does not permit extrapolation.

While this issue does not arise frequently, it does sometimes occur when we calculate p-values, but not so frequently as to undermine our confidence in the usefulness of the p-values and associated assessment of statistical significance.

Construction of Specific Indicators

1. New Housing

1.1 Core

The New Housing indicator is calculated as the number of units in new residential, multifamily (5 or more units) construction for which the City issues building permits over the preceding 12 months.

Data source: U.S. Census Bureau, Building Permits Survey.

1.2 Contextual

There are two contextual charts for this indicator. One maps new residential, multifamily building permits at the address-level for the most recent four years based on data availability. The other shows the ranking of Minneapolis compared to the donor cities used in the SCM by the most recent available trailing 12-month sum of new multifamily units.

Data sources: Metropolitan Council, Building Permits Survey; U.S. Census Bureau, Building Permits Survey.

2. Housing Composition

2.1 Core

The housing composition indicator is calculated using the Simpson's Reciprocal Diversity Index based on unit counts by structure category (and referred to as a structure diversity index, or *SDI*, in this work). For any city *j,*

$$SDI_j = \frac{N_j(N_j - 1)}{\sum_i n_{ji}(n_{ji} - 1)}$$

where $i \in$ {single family detached, single family attached and 2-4 unit buildings, 5-19 unit buildings, 20 or more unit buildings}, n_{ji} is the number of units in structure category i in city j, and N_{tj} is the total number of units in city j across all four category types. The structure diversity index for any given city will range from 1 to 4. A value of 1 indicates that all of the housing units in the city are of a single category while a value of 4 indicates that the units are equally distributed across all four categories.

Data source: U.S. Census Bureau, American Community Survey 1-year files.

2.2 Contextual

There are three contextual charts for this indicator. The first two show the most recently available structure diversity index at the neighborhood-level in Minneapolis, and the change from five years prior to allow for non-overlapping data. Each neighborhood is constructed based on individual Census block groups. The last contextual chart shows the ranking of Minneapolis compared to the donor cities used in the SCM by the most recent available city-level structure diversity index.

Data sources: U.S. Census Bureau, American Community Survey 5-year and 1-year files.

3. New Affordable Housing

The New Affordable Housing indicator is calculated as the number of new income-restricted housing units at 60 percent of area median income (AMI). Due to the lack of comparable data from other cities, we construct two comparisons to Minneapolis based on the regional data: St. Paul and the combination of other cities in the Twin Cities region that the Metropolitan Council designates as Urban Center and Urban. (Learn more about Community Designations from the Metropolitan Council.)

Data source: Metropolitan Council, Affordable Housing Construction.

4. Stock of Affordable Housing

4.1 Core

The Stock of Affordable Housing indicator is calculated as the share of total rental units where the housing cost is affordable (less than or equal to 30% of income) to households at 50 percent of HUD Area Median Family Income (HAMFI).

Data source: U.S. Department of Housing and Urban Development, Comprehensive Housing Affordability Strategy (CHAS) data.

4.2 Contextual

The only contextual chart for this indicator shows the ranking of Minneapolis compared to the donor cities by the most recent available city-level share of rental units that are affordable to households at 50 percent of HAMFI.

Data source: U.S. Department of Housing and Urban Development, Comprehensive Housing Affordability Strategy (CHAS) data.

5. Housing Cost Burden

5.1 Core

The Housing Cost Burden indicator is calculated as percent of extremely low-income renting households who experience housing cost burden. Extremely low-income household is defined as a household whose household income is less than 30 percent of AMI. A household is cost-burdened if their housing cost-to-household income ratio is at or above 30 percent.

To smooth out the fluctuation due to measurement error from using ACS 1-year files, the three-year lagging average is reported as the final indicator.

Data sources: U.S. Census Bureau, American Community Survey, Public Use Microdata Sample, 1-year files; U.S. Department of Housing and Urban Development.

5.2 Contextual

There are three contextual charts for this indicator. The first one shows racial/ethnic composition of residents among Minneapolis extremely low-income renters, Minneapolis overall, Minnesota overall and U.S. overall. Race/ethnicity of residents is grouped into five categories:

1. Asian, non-Latinx
2. Black or African American, non-Latinx
3. Latinx, inclusive of Hispanic or Latino/a origin regardless of their race
4. White, non-Latinx
5. All other races, inclusive of all residents that do not fall into one of the four categories above. This is done due to the lack of availability of the data for other groups such as Native Americans and individuals reporting two or more races.

The second contextual chart shows the time series of housing cost burden by racial/ethnic group and tenure. The last contextual chart shows the ranking of Minneapolis compared to the donor cities by the most recent available city-level housing cost burden among extremely low-income renters.

Data sources: U.S. Census Bureau, American Community Survey, Public Use Microdata Sample, 5-year and 1-year files; U.S. Department of Housing and Urban Development.

6. Price of Housing

6.1 Core

The Price of Housing indicator is the city-level Median Gross Rent.

Data source: U.S. Census Bureau, American Community Survey, 1-year files.

6.2 Contextual

There are four contextual charts for this indicator. The first one shows the time series of Zillow Observed Rent Index (ZORI) which captures the typical rent charged in Minneapolis. The second chart shows the time series of Zillow Home Value Index (ZHVI) which captures the level and appreciation of home values across Minneapolis. The third contextual chart shows the median rent for the apartment market and the shadow market (non-apartment rentals) by number of bedrooms. The final contextual chart shows the ranking of Minneapolis compared to the donor cities by the median gross rent.

Data sources: Zillow, Inc.; HousingLink, Twin Cities Rental Revue; U.S. Census Bureau, American Community Survey, 1-year files.

7. Housing Choice

7.1 Core

The Housing Choice indicator is calculated as percent of homes in Minneapolis that are affordable to purchase for extremely low-income households. Given household income, we calculate the house buying power as the price of the house the income could afford by assuming:

- 30 percent debt-to-income ratio
- 10 percent down payment
- 30-year term
- 1.36 percent property tax rate, annually
- 1 percent mortgage insurance, annually
- 150 dollars hazard insurance, monthly
- Annual average mortgage interest rate

Extremely low-income household is defined as a household whose household income is less than 30 percent of area median income.

To smooth out the fluctuation due to measurement error from using ACS 1-year files, the three-year lagging average is reported as the final indicator.

Data sources: U.S. Census Bureau, American Community Survey, Public Use Microdata Sample, 1-year files; U.S. Department of Housing and Urban Development; Freddie Mac, 30-Year Fixed Rate Mortgage Average in the United States.

7.2 Contextual

There are three contextual charts for this indicator. The first one shows the most recent available share of homes that are affordable to buy by Minneapolis neighborhood at median household income of homeowners and renters by race and ethnicity. Only single-family homes, townhomes and condominiums are included in this sample to capture the one-unit homes that a household could buy. The second contextual chart shows neighborhoods where median rent is affordable at median household income by race and ethnicity. The data are binary for this chart due to data availability. The final contextual chart shows the ranking of Minneapolis compared to the donor cities by the most recent available city-level share of homes affordable to extremely low-income households. The racial/ethnic group is based on the same definition as in 5.2, and Minneapolis neighborhood is constructed in the same way as in 2.2.

Data sources: MetroGIS Regional Parcel Dataset; U.S. Census Bureau, American Community Survey, Public Use Microdata Sample, 5-year and 1-year files; U.S. Census Bureau, American Community Survey 5-year files; U.S. Department of Housing and Urban Development.

8. Income Segregation Index

8.1 Core

Income segregation is measured using the rank-order information theory index (H^R) as described in Reardon (2011) and implemented via the *OasisR* package in *R*. 5-year ACS data at the Census tract level are used to estimate H^R. Importantly, the household counts are reported in the ACS are for categories of income (less than \$10,000, \$10,000 - \$15,000, \$15,000 - \$20,000, etc.) as opposed to fixed percentiles of the income distribution. Construction of the estimate is done in the following manner:

1. Obtain the counts of households by income group (there are 16 of these in the ACS surveys starting in 2010) for all of the tracts that comprise a city. There are 116 tracts in the city of Minneapolis, so the resulting set of household counts comprises a matrix whose size is 116 x 16. Each column of the matrix corresponds to the number of households in tract (j) that have income in category (k).[3]
2. Convert the matrix of raw counts into one of cumulative counts up to each income category. For example, the third column of the matrix would contain the total number of households with incomes up to \$20,000 (so it would be the sum of the raw counts of columns 1 – 3 in the original matrix.
3. Calculate the traditional *information theory index* of segregation for each the first 15 columns of the matrix in 2 using the following formula:

$$H(k) = 1 - \sum_{j} \frac{t_j * E_j(k)}{T * E(k)}$$

   where T is the population of the entire city (with incomes less than or equal to category k), t_j is the population of Census tract j (with incomes less than or equal to category k), $E(k)$ is the entropy measure for the city, and $E_j(k)$ is the entropy measure for tract j. The entropy measures are given by the following formula:

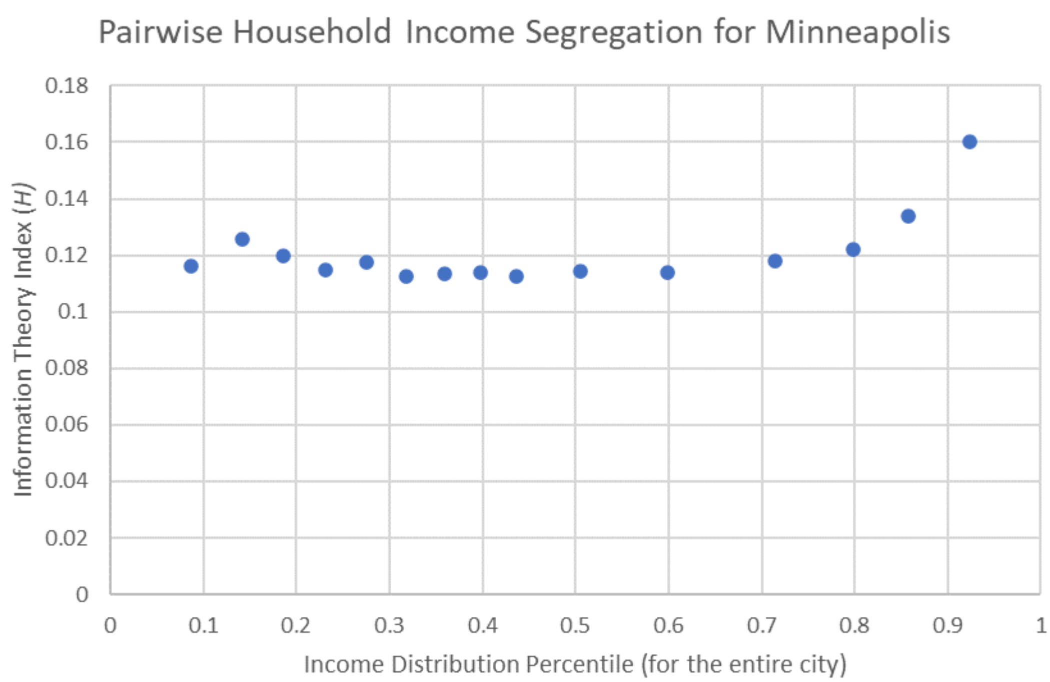
$$E(k) = k * log_2\left(\frac{1}{k}\right) + (1 - k) * log_2\left(\frac{1}{1 - k}\right)$$

[3] Census tract boundaries and census place boundaries do not always align well. To get the tract-to-place mapping to allow us to calculate this indicator for Minneapolis and all the donor cities, we use a 40 percent area overlap between tract and place cutoff. We did a sensitivity check by looking at how different cutoffs affect the indicators and found 40 percent to be a reasonable choice compared to all other options between 25 percent and 50 percent. This mapping is also used in indicators 9 and 10.

The value of k in the above equation corresponds to the percentile of the cumulative income distribution at each income category (measured at the individual Census tract level or for the entire city). For example, if the first Census tract in Minneapolis contained 1,055 households and there were 93 households with income in the first category (less than $10,000), 38 households with income in the second category ($10,000 - $15,000) and 17 households with income in the third category ($15,000 - $20,000) then the value of k for this tract would be 0.139 ([93 + 38 + 17] / 1,055).

The 16th column corresponds to income greater than $200,000. However, this category is redundant since the segregation measures are based on separating the households into two groups at each point (those with incomes below the category cutoff and those with incomes above).

The result from this step will be a vector of 15 values, each of which represents the *information theory index* of segregation at the various income category cutoffs for the city as whole. The chart below shows the result for Minneapolis using the 5-year ACS data from 2014-2018:



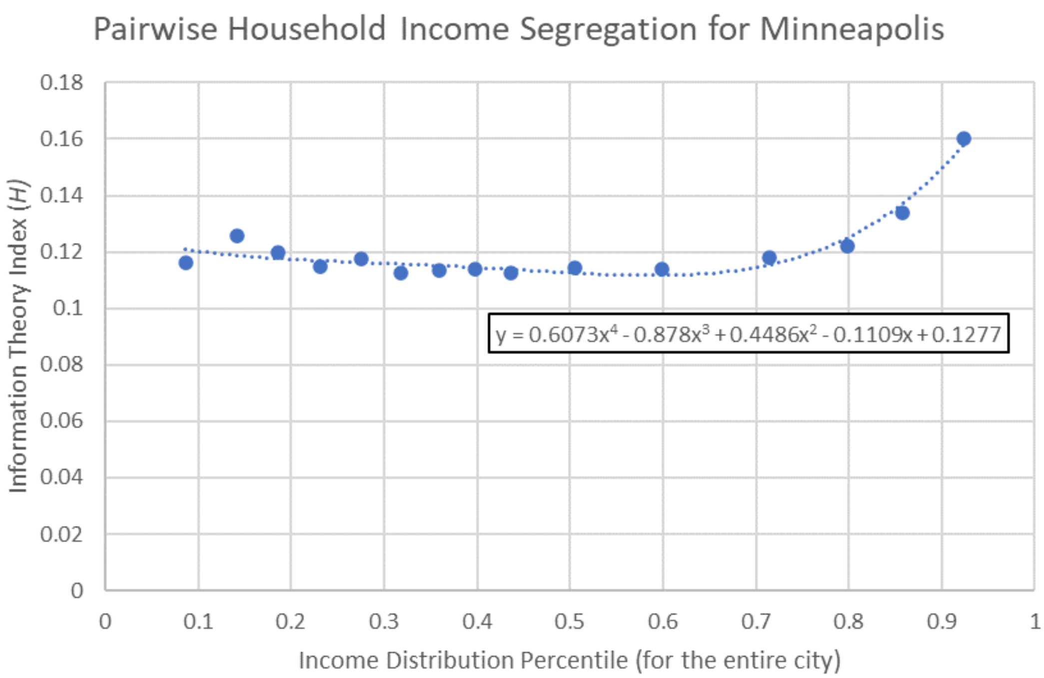
4. Finally, the individual H values across the entire income distribution (i.e., all of the income percentiles) can be aggregated into a single value as the *rank-order information theory index* and denoted by H^R. If the full, continuous H function were available then value of H^R would be given by the following equation:

$$H^R = 2 * \ln(2) \int_0^1 E(p)\ H(p)\ dp$$

However, we only have 15 discrete observations of H as opposed to the full function. Reardon et al. (2006) show that the H^R can be approximated by fitting a 4th order polynomial through the discrete H observations and then using the associated polynomial coefficients in the following equation:

$$\hat{H}^R = \hat{\beta}_0 + 0.5 * \hat{\beta}_1 + 0.305 * \hat{\beta}_2 + 0.208 * \hat{\beta}_3 + \ 0.152 * \hat{\beta}_4$$

The coefficients show below that are associated with the 4th order polynomial for the Minneapolis data produce a value of 0.1189 for $\hat{H}^R$.



Data Source: U.S. Census Bureau, American Community Survey 5-year files.

8.2 Contextual

There are three contextual charts for this indicator. The first two show the income segregation by Minneapolis neighborhood and the change from 5 years prior. The last contextual chart shows the ranking of Minneapolis compared to the donor cities by the most recent available city-level income segregation index.

Minneapolis neighborhood is constructed based on block groups as described in 2.2.

Data source: U.S. Census Bureau, American Community Survey 5-year files.

9. Housing Cost Segregation Index

9.1 Core

This indicator is calculated using the same method as 8. Income Segregation Index, but instead of tract-level household counts by income group, tract-level housing units by housing cost group is used.

Data source: U.S. Census Bureau, American Community Survey 5-year files.

9.2 Contextual

There are three contextual charts for this indicator. The first two show the housing cost segregation by Minneapolis communities and the change from 5 years prior. Minneapolis communities are made up of census tracts. The last contextual chart shows the ranking of Minneapolis compared to the donor cities by the most recent available city-level housing cost segregation index.

Data source: U.S. Census Bureau, American Community Survey 5-year files.

10. Isolation of White Residents

10.1 Core

The isolation index is calculated as:

$$I = \sum_{t \in T} \left(\frac{W_t}{W_R} \times \frac{W_t}{P_t} \right)$$

where T is a set of all census tracts in Minneapolis, W_t is the White, non-Latinx population in tract t, W_R is the total White, non-Latinx population in the city, and P_t is the total population in tract t.

Data source: U.S. Census Bureau, American Community Survey 5-year files.

10.2 Contextual

There are four contextual charts for this indicator. The first two show the isolation index by Minneapolis neighborhood and the change from 5 years prior. Minneapolis neighborhood is constructed based on block groups as described in 2.2.

The third contextual chart shows the actual to predicted number of households of color based on income only. The ratio is calculated as:

$$\text{Actual-to-predicted ratio}_n = \sum_{i \in I} \frac{HOC_{iR}}{T_{iR}} \times T_{in}$$

where Actual-to-predicted ratio$_n$ is the actual-to-predicted number of households of color in neighborhood n, I is the set of all income brackets available in the ACS data, HOC_R is the number of households of color in income bracket i in Minneapolis, T_R is the total households in income bracket i in Minneapolis, and T_{in} is the total households in income bracket i in neighborhood n.

The last contextual chart shows the ranking of Minneapolis compared to the donor cities by the most recent available city-level isolation index.

Data source: U.S. Census Bureau, American Community Survey 5-year files.

References

Abadie 2020, "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects"

Abadie, Diamond and Hainmueller 2010, "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program"

McClelland and Gault 2017, "The Synthetic Control Method as a Tool to Understand State Policy"

Reardon et al. 2006, "A New Approach to Measuring Socio-Spatial Economic Segregation"

Reardon 2011, "Measures of Income Segregation"