An Equitable Distribution Model of Cook County's CARES Act Funding to Suburban Municipalities

White Paper on the Subgrantee Allocation Process

Authored By:
Dean A Constantinou, Manager of FP&A, Cook County Government
Chicago Metropolitan Agency for Planning (CMAP) Staff
Lauren (Ché) Anderson-Justice, Financial Analyst, Cook County Government

COOK COUNTY GOVERNMENT

Published July 14, 2020

Abstract

On April 24, 2020, Cook County received $428.6 million in Coronavirus Relief Funding (CRF) under the Coronavirus Aid, Relief, and Economic Security (CARES) Act to help offset the County's costs related to COVID-19. Under guidance from the US Treasury's Office, Cook County was allowed, though not required, to allocate some of the funding to suburban municipalities within Cook County. The County opted to provide about $51 million, or 12% of the $428.6 million, to the suburban municipalities. In working with its regional planning partner, the Chicago Metropolitan Agency for Planning (CMAP), the County developed an equitable funding allocation method that utilized publicly available third party data that could be independently verified. The allocation method considered the rising COVID-19 related costs experienced by all municipalities, as well as the capacity of various municipalities within the County to respond to those costs in light of unequal historic and current investment in those communities. The following document provides a description of the funding allocation model developed by Cook County and CMAP, along with a description of the decision-making process used to develop the Sub-Grantee Allocation Method.

Overview of the Method

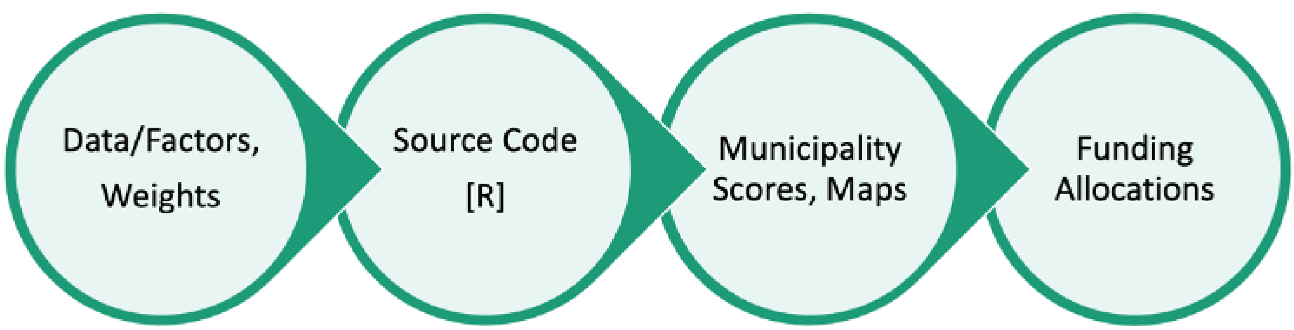


Figure 1. Conceptual representation of the Sub-grantee Allocation Method

The Sub-Grantee Allocation Method involved a 4-step process.

1) The data points were selected and gathered, and weights were applied to each data set.
2) The data and weights from step 1 were run through the CMAP Model (using R)[1].
3) From that process, several outputs were created. These outputs included various heat maps and visualizations of the data, as well as an individual score for each municipality.
4) The score for each municipality was used to develop an individual allocation to each municipality.

[1] "R" is an open source statistical analysis software.

Description of CMAP's Model Data

Table 1. Municipal data factors included in the model.

Factor	Data Source
Population	U.S. Census Bureau's Population Estimates Program (vintage: 2019 estimate from the 2010-2019 City and Town Population Totals table)
Median household income	U.S. Census Bureau's American Community Survey (vintage: 2014-2018 ACS 5-year estimates, table B19013)
Tax base per capita	Illinois Department of Revenue total equalized assessed value (2018) plus total retail sales (2019), normalized by population factor (above)
Percent of population located in an economically disconnected or disinvested area (EDA)	CMAP's 2015 Parcel-Based Housing Inventory and EDA boundaries

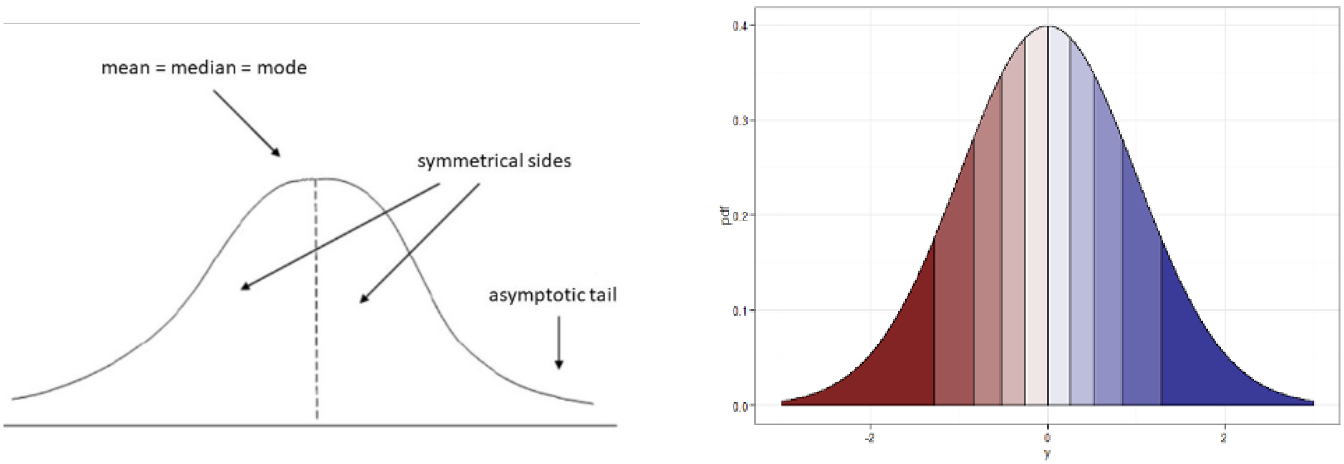
Methodology

In developing our model, the County adapted CMAP's Community Cohort Evaluation Tool[2] to fit our needs. The model has three components: input data, source code, and output files (Figure 1). The inputs consist of a single Excel file containing two sheets. The first sheet holds the factors/data for each of the 134 Suburban Cook municipalities, while the second defines the weights assigned to each factor. The Excel file can be modified to change the datasets and weights without altering the source code. The source code is a sequence of computational statements written in R. The primary output of the model consists of a CSV file containing the final scores for each municipality. The model also generates several charts and maps to visualize the results.

- **Assumptions:** The only assumption is that the datasets (i.e. factors) should be approximately normally distributed. A normal distribution has a bell-shaped density curve that allows a description of a dataset by its mean and standard deviation (Figure 2a). The normal distribution curve is symmetrical, centered about its mean/median, and its spread is determined by its standard deviation. With perfectly normally distributed data, the individual data points can be organized into ten equally-sized groups (deciles) based solely on the median and specific multiples of the standard deviation (Figure 2b).

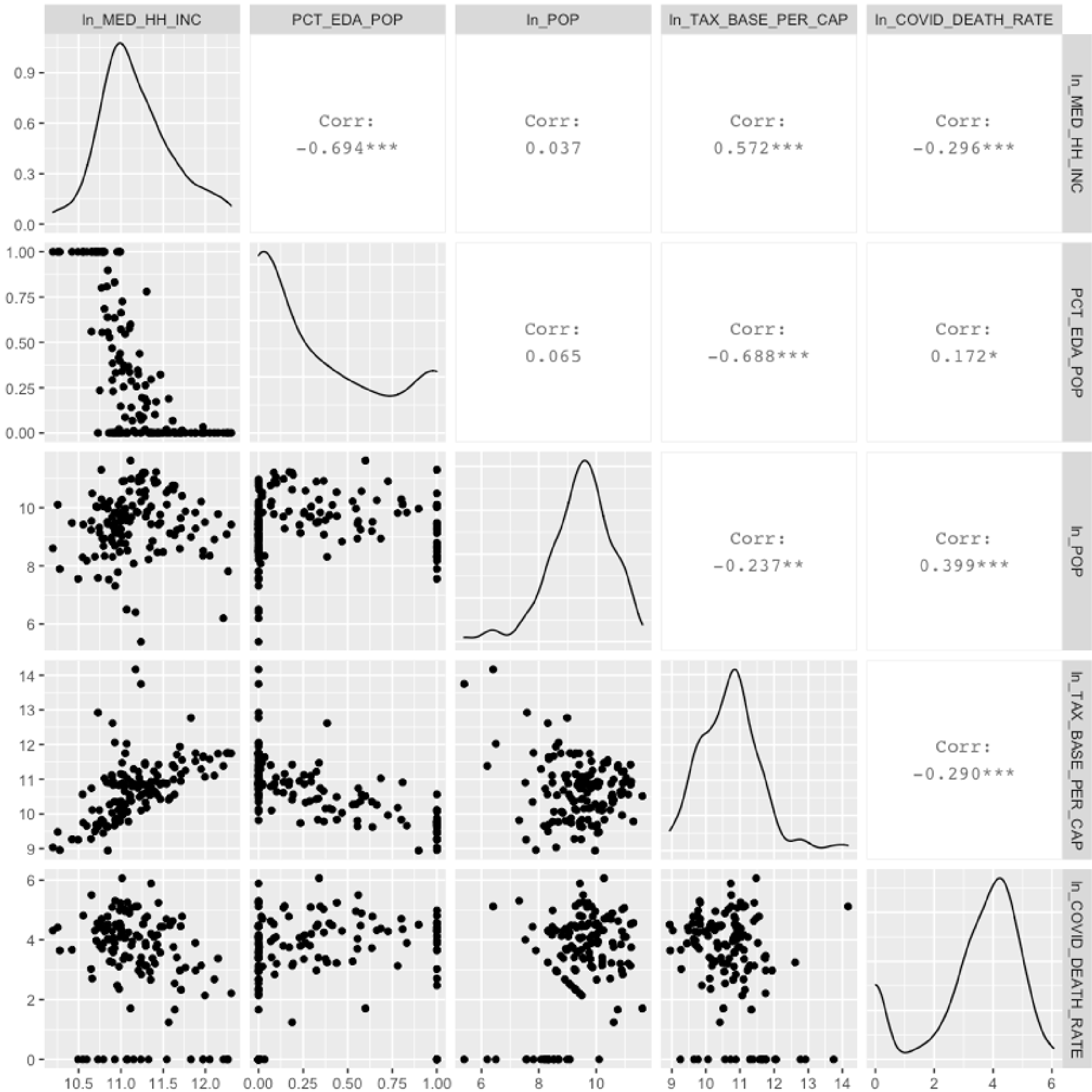
[2] https://www.cmap.illinois.gov/documents/10180/997259/CCET+Whitepaper+2020-02-10_web.pdf

Figures 2a-b. A conceptual representation of a bell-shaped normal distribution's probability density function (left), divided into deciles based on median and standard deviation (right).



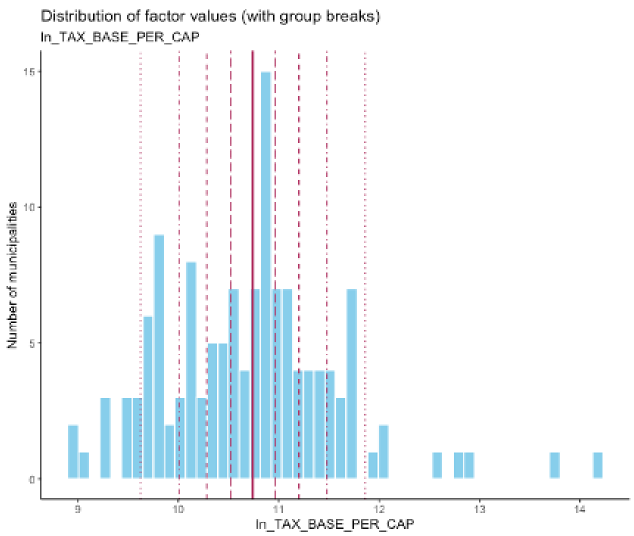
- **Normal distribution test:** The input factor distributions were evaluated before inclusion in the model. Population, median household income, tax base per capita, and COVID-19 death rate all had natural log transformations applied to better approximate normal distributions. The percent of population located in EDAs factor is not at all normally distributed Instead, this factor has a bimodal distribution, with the vast majority of communities having either 0% or 100% of their population in EDAs. A correlation test was also performed to determine whether any factors were highly correlated to each other (Figure 3). Correlation tests show how strongly related two series of data are, and is normally expressed using the correlation coefficient of "r" which ranges between -1 and 1. If r is closer to 0 it means that there is no relationship. The closer r is to +1 or -1, the more closely the two series of data are related. Any correlation coefficient 0.7 or higher was considered highly correlated in this analysis, and no pair of included factors has a correlation coefficient exceeding that threshold. The percent of population in EDAs factor was kept in the analysis despite its non-normal distribution due to the importance of this factor in creating an equitable allocation.

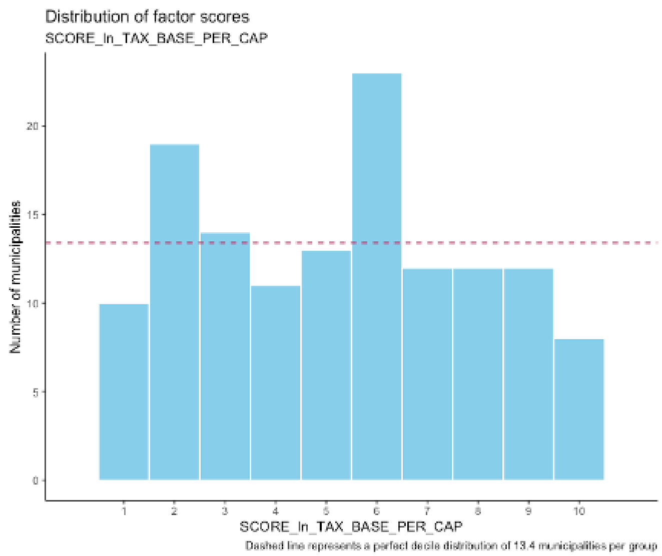
Figure 3. Data factor distributions (diagonal), correlations of factor pairs (upper-right), and scatterplots of factor pairs (lower-right).



- **Determination of factor-specific scores:** The decile "z-scores" (multiples of standard deviation) for a perfectly normal distribution were used to assign data into five groups at each side of the median value for each of the factor distributions. The analysis used median value instead of mean for the midpoint so that, for each factor, exactly 50% of communities would receive a score of 1-5 and the other 50% would receive a score of 6-10 (see Figures 4a-b).

Figures 4a-b. Left: an example of the log-transformed tax base per capita factor's distribution, with group thresholds (maroon lines) assigned by median (thickest line) and z-score to approximate deciles. Right: the resultant number of municipalities receiving each score, based on the approximated deciles.





A score of 1-10 was assigned to the groups with 1 being the lowest score and 10 being the highest. A score of 1 was assigned to the group at the far left side of the bell curve, while a score of 10 was assigned to the group at the far right of the bell curve. This score ordering would be reversed for factors with a negative weight. The group thresholds were defined as follows:

- Group 1 < median - 1.2816 standard deviations
- Group 2 < median - 0.8416 standard deviations
- Group 3 < median - 0.5244 standard deviations
- Group 4 < median - 0.2533 standard deviations
- Group 5 < median + 0 standard deviations
- Group 6 < median + 0.2533 standard deviations
- Group 7 < median + 0.5244 standard deviations
- Group 8 < median + 0.8416 standard deviations
- Group 9 < median + 1.2816 standard deviations
- Group 10 >= median + 1.2816 standard deviations

Due to its unique distribution, the percent of population in EDAs factor was grouped differently. Communities in which 90-100% of the population lived in EDAs were assigned to Group 1, communities with 80-89.99% of their population in EDAs were assigned to Group 2, communities with 70-79.99% of their population in EDAs were assigned to Group 3, and so on. Group 10, which consisted of municipalities with 0-9.99% of their population living in EDAs, contained 63 of the 134 suburban Cook communities. 48 of the municipalities in Group 10 had 0% of their population living in EDAs.

For the COVID-19 death rate factor, Group 1 was redefined to only include the municipalities with zero deaths (19 municipalities, as of 6/21/2020). Some municipalities were initially assigned to Group 1 based on their z-score, but had at least one COVID-19 death. These municipalities were re-assigned to Group 2.

- **Calculation of overall scores:** The overall score for each municipality was calculated by multiplying each factor's score by the absolute magnitude of its weight (Table 2), then summing the weighted scores. The overall scores were then dynamically rescaled so that the total range of theoretically achievable scores would be 0-100, regardless of how many factors were included or what weights were used. The value was then subtracted from 100 so that higher scores would indicate a higher level of financial need.

Table 2. Factor weights in the base scenario (equal weighting). Negative weights indicate that the lowest data values should get the highest factor scores, instead of the lowest factor scores.

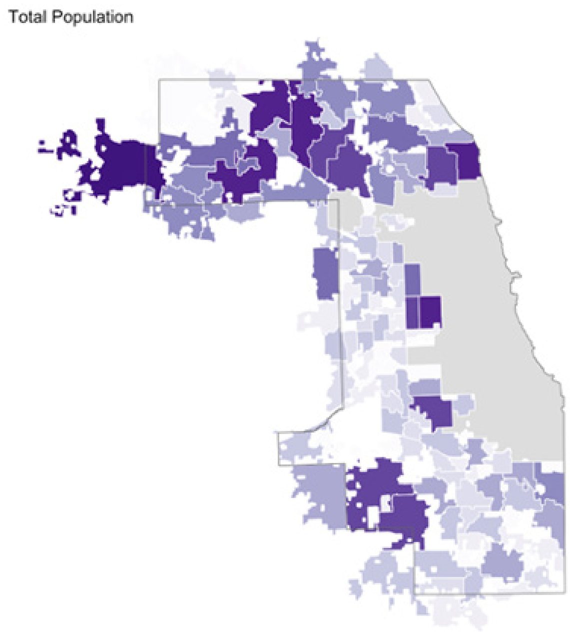
FACTOR NAME	WEIGHT
Population	-20
Tax base per capita	20
Median household income	20
Percent of population located in an economically disconnected or disinvested area (EDA)	-20
COVID-19 deaths per 100,000 population	-20

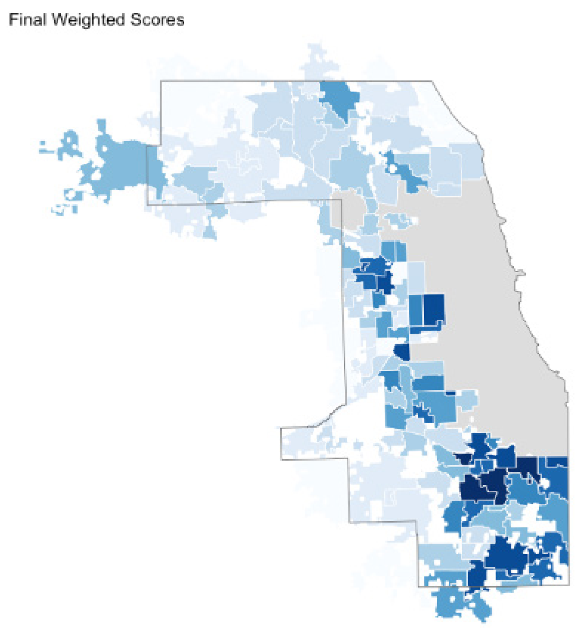
- **Output:** The output of the model consists of a CSV file (Table 3) containing each municipality's total population, overall score (0-100) and factor-specific scores (1-10), as well as a map showing the geographic distribution of the overall scores (Figures 5a-b).

Table 3. Sample of the output CSV file for the base (equal weights) scenario.

GEOID	MUNI	SCORE_ln_POP	SCORE_ln_TAX_BASE_PER_CAP	SCORE_ln_MED_HH_INC	SCORE_PCT_EDA_POP	SCORE_ln_COVID_DEATH_RATE	SCORE_OVERALL_SCALED	Allocation Score
1701010	Alsip	4	6	4	10	9	73.33	26.67
1702154	Arlington Heights	1	6	8	9	7	77.78	22.22
1703844	Barrington	7	9	10	10	8	94.44	5.56
1703883	Barrington Hills	9	9	10	10	10	98.89	1.11
1704013	Bartlett	2	4	9	9	9	83.33	16.67
1704572	Bedford Park	10	10	6	10	2	68.89	31.11
1704975	Bellwood	4	2	4	2	6	26.67	73.33
1705248	Bensenville	4	7	5	8	8	66.67	33.33
1705404	Berkeley	9	3	5	10	10	75.56	24.44
1705573	Berwyn	2	2	5	3	4	30.00	70.00

Figure 5a-b. Left: a heatmap showing the total population of each suburban Cook County municipality. Right: a heatmap showing the final score for each municipality, based on the weights assigned to each factor a darker color indicates a greater relative allocation.





Determining the Weights: The Equity Lens

One major aspect of our analysis was the weights applied to each factor. The weights were determined within the broader context of the County's allocation of the full $428.6 million of CRF CARES Act Funding. Under the governance structure within Cook County, a subcommittee was established specifically for the allocation and reimbursement of CARES act funding to Cook County's suburban municipalities.

After an initial review of the model by the project management oversight group, many outputs from the CMAP model were presented. These outputs included heat maps various other data visualizations. The scenarios shown in Table 4 were presented to the group for consideration, in order to demonstrate the impact of weighting one element over another in the CMAP model.

For each scenario beyond the baseline scenario, the following charts were presented to the group for consideration.

Table 4: The scenarios presented to the project management oversight group, with weights for each factor.

SCENARIO	POPULATION	TAX BASE PER CAPITA	MEDIAN INCOME	% POP. IN EDA	COVID-19 DEATHS PER 100K
BASE	20%	20%	20%	20%	20%
POPULATION	40%	15%	15%	15%	15%
TAX BASE	15%	40%	15%	15%	15%
MEDIAN INCOME	15%	15%	15%	15%	15%
% POP. IN EDA	15%	15%	40%	40%	15%
COVID-19 DEATHS PER 100K	15%	15%	15%	15%	40%

In order to facilitate discussion, the CMAP model and its inputs were explained to members of the sub-grantee committee, as follows:

The CMAP model looks at each of the below metrics and fits the data to a normal distribution (bell curve). Once adjusted, each municipality is given a score based on which decile of the bell curve they land in (ranging from 1 to 10). **The municipalities with the lowest scores will receive a greater allocation because they have the highest need.**

Population: The municipalities with the highest populations have the lowest scores because they serve more people and therefore have a greater need and would therefore be eligible for a higher allocation.

Median income: Municipalities with the highest median income have the highest scores, because the people in these municipalities have more resources and would therefore be eligible for a lower allocation.

Percent Population in Economically Disinvested Area: Municipalities with the highest percent population in an economically disinvested area have the lowest scores because they have more vulnerable populations, and would therefore be eligible for a higher allocation.

Covid-19 deaths per 100K of population: Municipalities with the highest number of deaths have the lowest score because they have suffered the most from Covid-19 and would therefore be eligible for a higher allocation.

Tax Base Per Capita: Municipalities with the highest tax base per capita have the highest scores, because the residents enjoy a greater level of service per person and would therefore be eligible for a lower allocation

Depending on the weights we decide on, an aggregate score is produced. Again, the municipalities with the lowest scores will receive a greater relative allocation.

The outputs from CMAP's model consist of four charts for each scenario. For the base scenario, only two charts are produced.

The scenario weighted toward population was used as an example to explain the outputs.

Figure 6, to the right, shows the distribution of scores if we weight the population scores at 40% and the rest at 15%.

Figure 7, to the right, shows a "heat map" demonstrating the scores in the heavily weighted population scenario. Darker colored municipalities have lower scores and therefore would receive a greater allocation than lighter municipalities.

Figure 6: Distribution of overall scores in the Population scenario.

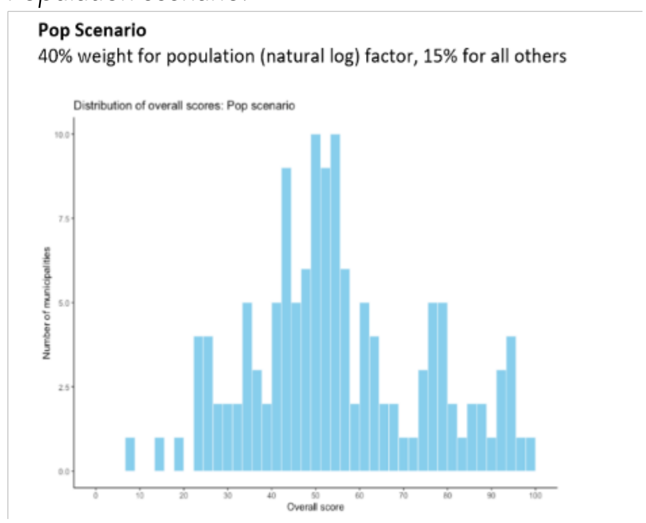
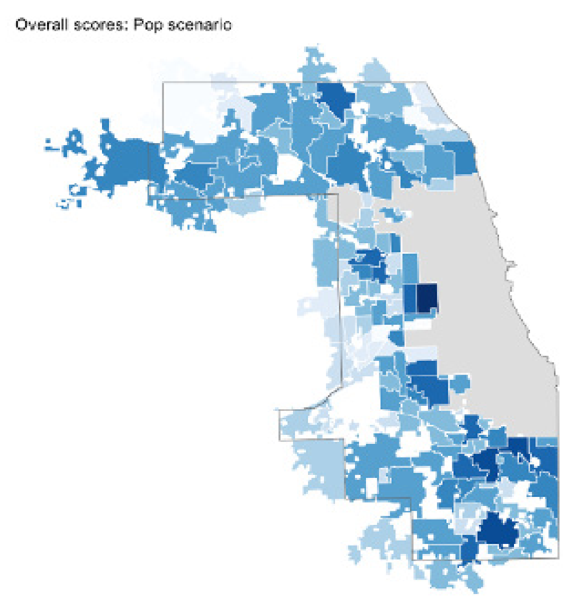
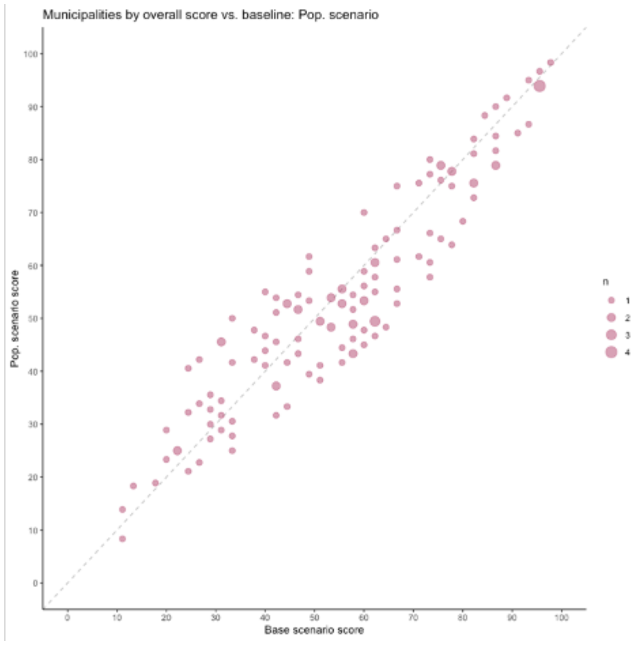


Figure 7: A heat map showing scores for each municipality in the Population scenario.



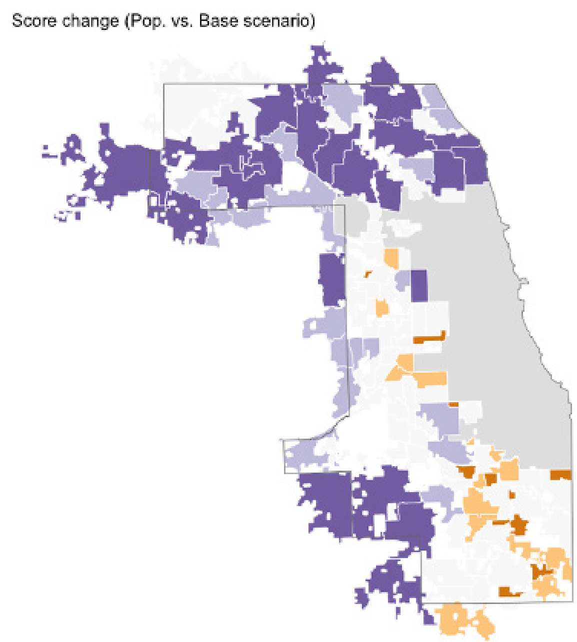
The third graphic, Figure 8, is available for all scenarios except the baseline scenario because it shows a comparison of the baseline scenario to the alternate scenarios (population in this case). This chart plots the scores of each municipality in both scenarios with the baseline scenario score on the x axis and the population scenario on the y axis. A good way to interpret this chart is that all of the municipalities below the dotted line would prefer the population scenario to the baseline scenario, as they would receive a greater relative allocation. We encourage the readers to review this chart for the "Scenario % pop in EDA" in the pdf.

Figure 8: A graph comparing scores for each municipality in the baseline scenario vs population scenario.



The fourth graphic, Figure 9, also compares the population scenario to the baseline scenario and provides a heat map that shows how the population scenario would impact specific municipalities. In this map, the darker purple municipalities would see an increased allocation under the population scenario, while the darker orange municipalities would see a decreased allocation.

Figure 9: A heat map showing the score change for municipalities in the population vs baseline scenario.



A preferred weighting strategy was presented and after much discussion the County settled on the following weights for each input:

40% for percent population in Economically Disadvantaged Areas

30% for median income

20% for COVID-19 deaths per capita

10% for tax base per capita

The population element was removed from the CMAP Model in favor of allocating a portion of the available funding on a per capita basis. The final heat map is presented in Figure 5b, shown on page 8 of this document.

Using the Scores for Allocation

In developing an allocation technique, we considered many different options. We believe that the final technique provides a base level of allocation to each municipality according to their population, while also allowing us to take into consideration the uneven capacity of different municipalities to respond to COVID-19 due to historic disinvestment. This technique utilizes the scores from CMAP's model in a way that accurately reflects the differences in need and capacity between municipalities.

$$\boldsymbol{Allocation\ Per\ Municipality} = ((\boldsymbol{F} * (\boldsymbol{p}/(\textstyle\sum \boldsymbol{p})) + \boldsymbol{S} * ((\mathbf{100} - \boldsymbol{SCORE}))/(\textstyle\sum - \boldsymbol{SCORE})))$$

Whereas,

$F = \ First\ Allocation$

$S = \ Second\ Allocation$

$p = \ Municipal\ population$

$s_c \ = \ Covid19\ deaths\ per\ 100K\ score$

$s_t = \ Tax\ base\ score$

$s_i = \ Median\ household\ income\ score$

$s_e = \ \%\ of\ population\ in\ EDA\ score$

$w_c, w_t, w_i, w_e = \ Factor\ weights$

$$SCORE = Weighted\ CMAP\ SCORE$$

$$= 100 * \frac{\left((s_c * |w_c| + s_t * |w_t| + s_i * |w_i| + s_e * |w_e|) - (|w_c| + |w_t| + |w_i| + |w_e|)\right)}{9 * \ (|w_c| + |w_t| + |w_i| + |w_e|)}$$

In the Allocation Per Municipality formula above, "F" was determined by reviewing a survey issued to each of the municipalities. The survey was sent on May 14, 2020 and representatives from each municipality was asked to respond by May 18, 2020. Based on the responses, we developed an estimate of the County's total needs for direct and immediate expenses, such as Personal Protective Equipment, increased technology, additional labor, and teleworking capabilities. The estimated value was then allocated based on each municipality's population.

The remaining amount "S" was then allocated to each municipality based on their score. Within the model, a lower score would mean a higher level of eligibility. Since the scores are distributed along a normal distribution between 0 and 100, subtracting the score from 100 allowed us to modify the score to reflect the higher eligibility for lower scoring municipalities. The scores modified in this way were then summed, and the remaining value from the first allocation was then allocated to each municipality depending on the proportion of their modified score in comparison to the aggregate.

As a final step, municipalities that had populations within both Cook County and another county had their combined allocation allocated to them based on the % of their population within Cook County.

Conclusion

The impact of Covid-19 on the region represents a clear and present danger from a public health perspective and a long-term potential threat to the economic vitality of all suburban municipalities. Historic disinvestment in certain communities within the region has resulted in an unequal capacity for all municipalities to respond to the challenges that Covid-19 presents. By working with CMAP (our regional planning body) and committing to an equitable distribution strategy, the County hopes to prevent a widening disparity gap, as experienced during the last economic recession. We hope that this method can serve as an example for future allocation strategies that utilize statistical analysis, demographic data, and regional cooperation, and that an equity lens will continue to be used in future funding allocation endeavors.