

Methodology Report for 2020 Youth Count in Connecticut

Prepared for the Connecticut Coalition to End Homelessness

By Stephen Adair, Ph.D.
Department of Sociology
Central Connecticut State University
New Britain, CT 06050
July 23, 2020
adairs@ccsu.edu
860-558-7475

The estimate of the total number of homeless or unstably housed youth (ages 14 to 24) in Connecticut in 2020 is 7823. This is somewhat less than the estimate from last year of 9303. Table 1 presents the respective figures for each age group.

Table 1: Estimated Number of Homeless and Unstably Housed Youth in Connecticut By Age

Age group	Number of Homeless	Number Unstably Housed	Total
Under 18	696	1748	2444
18-24	1766	3613	5379
Total	2462	5361	7823

Preliminary Observations for the 2020 Estimate

Estimating the number of homeless and unstably housed young people in the state is challenging. In general, estimating the size of a population requires a probability sample in which each person would have an equal chance of being represented in the survey. Over the last couple of years, the Connecticut Coalition to End Homelessness has engaged in a rigorous effort to survey young people to determine both the characteristics and the number of homeless and unstably housed young people. Particular youth are targeted for the survey, and there are no established statistical protocols or documented methods to determine a population size from a non-probability sample based on an anonymous survey.

The data for the 2020 count were collected before the outbreak of COVID-19 in the United States. Table 2 compares the overall results over the last two youth counts.

Table 2: Overall Results from Youth Survey in 2019 and 2020

Year	Total Surveys	Stably Housed	Unstably Housed (Youth Survey)	Homeless (Youth Survey)	Homeless from HMIS or unsheltered PIT data
2019	4751	3892 (81.9%)	675 (14.2%)	184 (3.9%)	152
2020	2156	1318 (61.1%)	579 (26.9%)	249 (12%)	339

Four observations:

- HMIS data and the annual point-in-time count only capture a fraction of the number of homeless and unstably housed youth in the state of Connecticut.
- Housing instability is two to three times as common as homelessness.
- In 2020, less than half the total number of 2019 surveys were completed (4751 compared to 2156), yet the relative prevalence of unstably housed and homeless increased. Given the higher prevalence, if the same methodology was used as last year, the estimate of both unstably housed and homeless youths would increase dramatically. Some changes in data collection likely explain at least a portion of the increase. For example, in 2019, several schools were canvassed widely to capture a large portion of students, rather than the targeted efforts to capture vulnerable youth.
- If every young person (age 14 to 24) had an equal chance of being represented in the survey, i.e. the survey was generally representative of the population, then the estimate for 2020 would be

that nearly 58,000 young people would be homeless (12 percent of 482,140) and 130,000 would be unstably housed. Beyond question, these figures greatly exceed the reality. The obvious assumption is that the canvassers can target vulnerable populations “successfully.”

Estimating the prevalence of homelessness and housing instability among young people therefore depends on developing a measure of the relative rate of “success” of the canvassers in targeting people for the survey. Some evidence for this measure is provided by comparing the results from 2019 with 2020. In administering fewer surveys in 2020, canvassers were more likely to be “successful” in finding homeless and unstably housed youth. As the size of the sample increases, the rate of “success” tends to decline.

In 2019, distinct estimating strategies were used for the 14-to-17-year-olds, than for those 18 to 24. For the younger age group, the school-based surveys were the basis for the estimate, which is not an option for the 2020 data as fewer school-based surveys were conducted. Table 3 includes all age groups and compares the data from 2019 and 2020 in selected cities and towns.

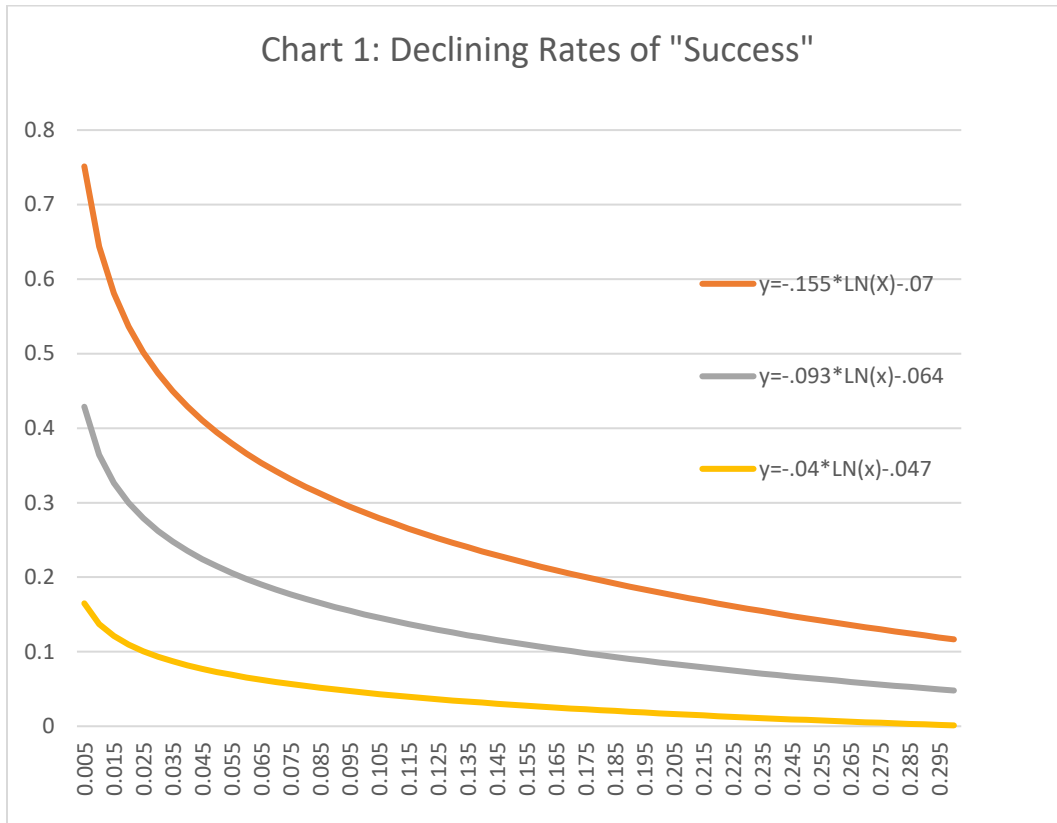
Table 3: Rates of “Success” for Selected Cities and Towns in Connecticut in 2019 and 2020

City or Town	Unstable or Homeless	2019		Unstable or Homeless	2020	
		Total Surveys	Percent “Successful”		Total Surveys	Percent “Successful”
Bridgeport	124	306	40.5	61	131	46.6
Danbury	134	167	80.2	76	89	85.4
Hartford	86	344	25.0	129	306	42.2
Killingly	47	576	8.2	63	196	32.1
Manchester	22	312	7.1	4	60	6.7
Mansfield	17	306	5.6	1	11	9.1
Meridan	18	237	7.6	8	19	42.1
Middletown	17	185	9.2	30	55	54.5
Naugatuck	8	56	14.3	11	26	42.3
New Britain	19	162	11.7	38	88	43.2
New Haven	72	110	65.5	38	51	74.5
New London	21	113	18.6	16	38	42.1
Norwalk	29	134	21.6	37	77	48.1
Norwich	61	502	12.2	71	151	47.0
Stamford	17	216	7.9	25	74	33.8
Thompson	8	105	7.6	14	44	31.8
Torrington	18	128	14.1	15	97	15.5
Waterbury	36	287	12.5	34	84	40.5
Windham	25	164	15.2	23	53	43.4

For the 19 cities and towns presented in Table 3, only Manchester had a lower rate of “success” in targeting unstably housed or homeless young people in 2020 than in 2019, and that decline was very small. In many cities, i.e. Meridan, New Britain, Norwich, Waterbury, Windham, the rate was several times higher in 2020 than it was in 2019. Danbury is an outlier with a surprising high rate of “success,” as it also had in each of the last four years of the youth count.

Graphing the change in “success” rate from 2019 to 2020 suggests a declining curve with an x-axis determined by the proportion of the population between the ages of 15 to 24 (based on census data) that was sampled.

The curves below are based on the natural log of the proportion of the population sampled. The curves are similar to graphs used to measure declining levels of productivity. The respective formulas for the three curves are presented in the legend.



Cities and towns represented by the top orange/red line would have the highest rates of housing instability. Canvassers would have a relatively easy time targeting individuals, such that their rate of success would be relatively high through the first few percent of a population that were sampled, and their success rate would tail off slowly as conditions of housing instability would be relatively pervasive across the population. The orange line corresponds to an overall rate of about 5.0 percent of the youth population being unstably housed or homeless.

Cities and towns represented by the yellow line would have relatively low rates of housing instability. While canvassers might be able to have some modest success in their initial targeting efforts, finding housing instability grows more difficult as they sample large portions of the population. The yellow line corresponds to an overall rate of about .55 percent of the youth population being unstably housed or homeless.

Although the graph presents three discrete lines, these might be envisioned as three instances of a three-dimensional curve with a surface that extends toward the viewer from the orange line to the yellow and beyond.

Methods and Calculations

Table 4 (below) presents the methodology for the estimates for the youth count. Although included in Table 3, Naugatuck and Mansfield have been dropped from the table calculations due to the small sample sizes for these locations in the 2020 count.

The first column is a relative rate of homelessness for each city and town. These rates are based on HMIS data for all people who entered a homeless shelter in Connecticut in 2015 and 2016. Of the roughly 17,000 people who spent at least one night in a shelter in Connecticut, about 14,000 people reported a last address in a city or town in Connecticut. The population size for all the cities and towns were identified from census data, so that a relative rate of each city and town's contribution to the homeless population per 1000 residents could be established. These figures are only used to select the curve that is applied for each city.

The next two columns list the number of 15-19-year-olds and 20-24-year-olds in these cities based on data in the US Census American Community Survey.

The next three columns present the results from the survey. The counts of both homeless and unstably housed youth and the total number of surveys conducted in each community are listed. The next column calculates a rate of "success" by taking the number of homeless and unstably housed counted in the survey as a portion of the number of people who completed a survey.

The next column identifies the portion of the population surveyed in each community. This is determined by taking the total number of surveys completed in each community and dividing it by the total number of 15-to24-year-olds in the community times 1.11, which is to expand the total population slightly to include an estimate of the number of 14-year-olds.

The three columns labeled p. orange, p. blue, and p. yellow predict relative success rates using the respective formulas for the three curves. The x in the formula is the proportion of the population that was sampled and listed in the previous column. Because the curves predicting rates of success increase rapidly when the proportion is less than .005, this was set as a minimum in the proportion of the population that was sampled. This limit was only imposed in New Haven and the remainder of state calculations.

The predicted success rate is based on a simple, proportional interpolation based on the homeless rate to determine a location on the imaginary three-dimensional curve. The orange curve corresponded to a homeless rate of 9.6 per thousand, the blue curve to a rate of 4.5 per thousand, and the yellow curve to a rate of 1.1 per thousand. These values were selected because they corresponded to the proportional differences in the total counts that would be estimated from a population for the three curves, and because they approximated the range of rates described in the first column.

For example, Bridgeport has a homelessness rate of 5.52. This is roughly 1 above the 4.5 rate for the blue curve, so in Bridgeport the predicted success rate ought to be a bit above 40.79. The bit is estimated as $1/(9.6-4.5)$, which is roughly 1/5 of the way between the blue and the orange curves, times the difference between the respective, predicted success rates: $71.65-40.79$. This yields a predicted "success" rate of 46.84, which is remarkably close to the observed rate of 46.56.

I did not interpolate above the orange curve. For the two cities with homeless rates above 9.6, Hartford and New London, the predicted rate from the orange curve was used.

In general, the predicted rate was very accurate for Bridgeport, Meridan, New Britain, New Haven, and Stamford. There were roughly as many as cities with actual “success” rates above the prediction as below, which suggests that the predicted rates might be a reasonable estimate. The predicted success rate was an especially poor predictor of the “success” rate in Danbury, the consistent outlier in the state.

On a separate Excel calculation, the “success” rates from the curves were converted into proportions of the total population that ought to have been measured based on the proportion of the population that was sampled. The column labeled “estimate” is the total number of “successes” one would expect given the predicted success rate and the proportion of the population that was measured. Because the prediction for Bridgeport (and Meridan, New Britain, New Haven, and Stamford) was so accurate, the estimate is nearly identical with the number actually observed (61.4 compared to 61).

The next two columns make estimates of the total number of homeless and unstably housed young people in each city or town. The first column uses the estimated number of successes. The second uses the observed number. For Bridgeport, this is determined by $61.4/.064$ and $61/.064$, which results in 959 and 953 respectively.

The difference between these figures is especially dramatic for Danbury, i.e. 216 and 1056.

A reasonable argument could be made regarding the level of confidence one might have for either figure. The estimated values represent a type of regression that moderates inconsistencies in the ability of the canvassers to find their targeted population. The actual values are based on the actual observations and the data collected. In the case of Danbury, using the actual value would suggest that Danbury has the highest rate of homelessness and unstable housing among young people in the state. However likely or unlikely this possibility may be, glossing it over by hiding it behind a purely predicted estimate may also be a disservice.

The last column represents a simple average of the two previous columns and is presented here as the recommended estimate. The total estimate of homeless and unstably housed people between the ages of 14 to 24 is 7823.

Age Groups and Distinguishing the Homeless from the Unstably Housed

Table 5 contains the same first four columns as Table 4. The rest of the table presents the survey results by age group and disaggregates the homeless from the unstably housed.

In the 2020, 1080 surveys were collected from people under 18 and 1076 were collected from those 18 to 24. The younger age (nearly all between 14 and 17) is more heavily represented than their proportion of the population would warrant. Since the objective is distribute the 7823 estimated number of homeless and unstably housed across the age groups to get an estimate for each group, the totals were adjusted to be proportional.

The values across the bottom row represent the estimates for the respective age groups