



WEIGHTING DOCUMENTATION

This chapter describes the weighting methodology used for the Chicago Household Travel Tracker survey. Weighting of the survey data is needed to develop estimates of population parameters and more generally to draw inferences about the population that was sampled. Without the use of analytic weights, population estimates are subject to biases of unknown (possibly large) magnitude. A household level analytic weight will be developed in this study. The common components of the analytic weights are as follows:

- Sampling weights – to adjust for probabilities of selection of a phone number or an address from the sampling frame,
- Adjustment for unknown eligibility and nonresponse– to compensate for differing patterns of eligibility and response,
- Dual frame sampling weight – to adjust for bias associated with combining the RDD and address-based sampling frame,
- Raking Adjustment – to align the weighted sample to known population distribution from 2000 Census data.

These weights adjust the relative importance of responses to reflect the different probabilities of selection of respondents, reduce bias in survey estimates from differing patterns of eligibility and response, adjust for bias associated with combining two sampling frames, and align the sample distributions to population distributions thereby improving coverage and precision. This chapter discusses the components of the household weight in detail.

SAMPLING WEIGHT

The sampling weight reflects the probability of selection of a telephone number or an address from the sampling frame. Considering the dual sampling framework employed in this study, separate sampling weights were calculated for the RDD and the address-based sampling frame¹. Specifically, the sampling weight for a sampling unit j in the sampling frame, selected from a stratum i , denoted as $W_{ij, SampFr}$, is simply the reciprocal of the selection probability of the sampling unit for the corresponding sampling stratum.

$$W_{ij, SampFr} = \frac{1}{\text{Prob}_{ij, SampFr}}$$

where,

Sampling unit j is a telephone number in the RDD sampling frame, and an address in the address-based sampling frame,

Sampling frame $SampFr$ is either the RDD sampling frame or the address-based sampling frame,

Stratum i is defined by cross-classifying the pre-defined strata and county of residence.

¹ It is important to note that the phone numbers were disproportionately drawn from the RDD and address-based sampling frame based on pre-defined strata. The pre-defined strata were defined by computing a composite measure that captures the population and job densities (with higher densities reflecting the more urbanized portions of the region) and the level of transit services and have five levels, with level 1 having the lowest levels of densities and transit service (both bus and rail) and level 5 having the highest available (see sampling plan for details).

Sampling Weight for Chicago Study area

Table 1 presents the sampling weights for the RDD sampling frame. A comparison of the percentage of phone numbers in the population and sample in Table 1 indicates an over-sampling of telephone numbers in areas within strata 3, 4 and 5.

Strata 3, 4 and 5 have higher levels of population and employment density, transit service availability and access. In addition, strata 3 and 4 have higher densities of minorities including African-American households (80% or higher), Hispanic households (60% or higher), and low-income households with household income less than \$25,000 (60% or higher). Further, strata 4 and 5 have the highest density of young households with age of the householder less than 25 years of age (10% or higher).

The sampling weights for the Address-based sampling frame are presented in Table 2. The table indicates oversampling of areas in strata 3 and 4 that have higher densities of minorities described above. In addition, these strata have the higher levels of population and employment density, transit service availability and access.

The sampling weights adjust for the bias associated with high probability of selection of phone numbers or addresses in over-sampled areas, and low probability of selection of phone numbers in under-sampled areas.

TABLE 1: SAMPLING WEIGHTS FOR RDD SAMPLING FRAME - CHICAGO STUDY AREA²

Stratum	County	Telephone numbers in Population ³		Telephone numbers in Sample		Sampling Weight
		Frequency	Percent	Frequency	Percent	
1	Cook	203,779	7%	1,658	6%	122.91
	DuPage	94,190	3%	845	3%	111.47
	Grundy	17,315	1%	109	<1%	158.85
	Kane	94,231	3%	940	3%	100.25
	Kendall	35,432	1%	150	1%	236.21
	Lake	126,428	4%	1,003	3%	126.05
	McHenry	87,433	3%	717	2%	121.94
	Will	168,892	6%	1,166	4%	144.85
2	Cook	441,036	15%	3,483	12%	126.63
	DuPage	167,292	6%	1,836	6%	91.12
	Kane	35,253	1%	366	1%	96.32
	Lake	84,298	3%	863	3%	97.68
	McHenry	20,914	1%	115	<1%	181.86
	Will	45,362	2%	315	1%	144.01
3	Cook	633,539	22%	7,133	24%	88.82
	DuPage	74,538	3%	641	2%	116.28
	Kane	31,979	1%	103	<1%	310.48
	Lake	32,491	1%	441	1%	73.68
	McHenry	1,632	<1%	0	0%	0
	Will	11,050	<1%	67	<1%	164.93
4	Cook	286,273	10%	4,000	13%	71.57
	DuPage	7,664	<1%	68	<1%	112.71
	Kane	357	<1%	0	0%	0
	Lake	1,903	<1%	25	<1%	76.12
5	Cook	237,263	8%	3,809	13%	62.29
	Kane	879	<1%	0	0%	0
	Will	800	<1%	0	0%	0
Total		2,942,223	100%	29,853	100%	

² It is important to note that the sample was drawn (proportional to the total number of households) by strata and not by a combination of strata and county. Hence, certain counties such as McHenry county in stratum 3, Kane county in stratum 4, and Kane and Lake counties in stratum 5, that have lower number of telephone numbers in the population compared to other areas do not have any sample drawn from the population.

³ The population refers to the total working residential numbers in the RDD sampling frame in the Chicago study area.

TABLE 2: SAMPLING WEIGHTS FOR ADDRESS-BASED SAMPLING FRAME - CHICAGO STUDY AREA

Stratum	County	Addresses in Population ⁴		Addresses in Sample		Sampling Weight
		Frequency	Percent	Frequency	Percent	
1	Cook	112,975	7%	18,988	6%	5.95
	DuPage	50,837	3%	5,026	1%	10.11
	Grundy	11,253	1%	1,895	1%	5.94
	Kane	53,446	3%	5,278	2%	10.13
	Kendall	19,270	1%	1,743	1%	11.06
	Lake	75,902	5%	9,922	3%	7.65
	McHenry	51,931	3%	5,887	2%	8.82
	Will	91,392	6%	11,548	3%	7.91
2	Cook	242,049	15%	55,702	16%	4.35
	DuPage	87,196	5%	11,522	3%	7.57
	Kane	17,970	1%	2,705	1%	6.64
	Lake	48,904	3%	8,460	2%	5.78
	McHenry	12,816	1%	1,987	1%	6.45
	Will	21,044	1%	4,690	1%	4.49
3	Cook	360,678	22%	121,521	35%	2.97
	DuPage	39,654	2%	6,309	2%	6.29
	Kane	16,310	1%	3,169	1%	5.15
	Lake	18,544	1%	3,905	1%	4.75
	McHenry	1,013	0%	188	0%	5.39
	Will	6,098	0%	1,314	0%	4.64
4	Cook	157,359	10%	46,768	14%	3.36
	DuPage	4,718	0%	742	0%	6.36
	Kane	97	0%	16	0%	6.06
	Lake	884	0%	385	0%	2.3
5	Cook	107,705	7%	14,566	4%	7.39
	Kane	298	0%	58	0%	5.14
	Will	424	0%	82	0%	5.17
Total		1,610,767	100%	344,376	100%	

⁴ The population refers to the total residential addresses in the Address-based sampling frame in the Chicago study area.

Sampling Weight for NIRPC Study area

Table 3 presents the sampling weights for the RDD sampling frame. A comparison of the percentage of phone numbers in the population and sample in Table 3 indicates an over-sampling of telephone numbers in strata 3, 4 and 5.

These strata have higher levels of population and employment density, transit service availability and access. In addition, strata 3 and 4 have higher densities of African-American households (72% or higher), and Hispanic households (10% or higher). Stratum 4 has the highest density of low-income households with household income less than \$25,000 (44% or higher), while strata 3 and 5 have the highest density of young households with age of the householder less than 25 years of age (7% or higher).

The sampling weights for the Address-based sampling frame are presented in Table 4. The table indicates oversampling of areas in strata 4 and 5 that have higher densities of minorities described above. In addition, these strata have the higher levels of population and employment density, transit service availability and access.

TABLE 3: SAMPLING WEIGHTS FOR RDD SAMPLING FRAME - NIRPC STUDY AREA

Stratum	County	Telephone Numbers in Population ⁵		Telephone Numbers in Sample		Sampling Weight
		Frequency	Percent	Frequency	Percent	
1	Lake	27,804	10%	434	9%	64.06
	La Porte	13,070	4%	276	6%	47.36
	Porter	18,041	6%	240	5%	75.17
2	Lake	30,177	10%	695	14%	43.42
	La Porte	8,553	3%	16	0%	534.56
	Porter	14,019	5%	141	3%	99.43
3	Lake	52,126	18%	469	10%	111.14
	La Porte	10,055	3%	393	8%	25.59
	Porter	28,692	10%	691	14%	41.52
4	Lake	51,555	18%	984	20%	52.39
	La Porte	7,932	3%	0	0%	0
	Porter	1,763	1%	0	0%	0
5	Lake	24,921	9%	457	9%	54.53
	La Porte	3,585	1%	95	2%	37.74
Total		292,293	100%	4,891	100%	

⁵ The population refers to the total working residential numbers in the RDD sampling frame in the NIRPC study area.

TABLE 4: SAMPLING WEIGHTS FOR ADDRESS-BASED SAMPLING FRAME - NIRPC STUDY AREA

Stratum	County	Addresses in Population ⁶		Addresses in Sample		Sampling Weight
		Frequency	Percent	Frequency	Percent	
1	Lake	18,956	11%	6,019	8%	3.15
	La Porte	8,472	5%	2,215	3%	3.82
	Porter	12,104	7%	2,914	4%	4.15
2	Lake	20,166	11%	6,955	9%	2.9
	La Porte	5,846	3%	2,081	3%	2.81
	Porter	8,472	5%	3,034	4%	2.79
3	Lake	30,068	17%	11,838	15%	2.54
	La Porte	6,645	4%	3,234	4%	2.05
	Porter	17,017	10%	6,501	8%	2.62
4	Lake	29,633	17%	19,001	24%	1.56
	La Porte	4,861	3%	2,552	3%	1.9
	Porter	1,054	1%	649	1%	1.62
5	Lake	13,791	8%	11,467	14%	1.2
	La Porte	2,018	1%	1,771	2%	1.14
Total		179,103	100%	80,231	100%	

ADJUSTMENT FOR UNKNOWN ELIGIBILITY AND NON-RESPONSE

Weight adjustments are necessary to minimize the potential bias due to unknown eligibility of telephone numbers/addresses and nonresponse. In order to apply the adjustments due to unknown eligibility and non-response, the sample was first divided into four categories based on the sample dispositions⁷. These categories for the RDD sampling frame are:

1. Eligible respondents – All telephone numbers known to be eligible households (i.e. residing within the study area) that participated in the survey.
2. Eligible non-respondents – All telephone numbers known to be eligible households that did not participate in the survey.
3. Ineligible – All ineligible telephone numbers (such as disconnected numbers, numbers belonging to government agencies, modem/fax numbers etc.)
4. Phone numbers with unknown eligibility – All telephone numbers with undetermined eligibility status (such as ring that goes to answering machine, ring with no answer on repeated calls etc)

Next, adjustments for unknown eligibility and nonresponse were calculated in two stages. In the first stage, the weighted telephone numbers with unknown eligibility status was distributed proportionately over the weighted telephone numbers with known eligibility status (including eligible respondents, eligible non-respondents and ineligible numbers). Thus, for each cell h , the adjustment for unknown eligibility, denoted as $NR1_h$ was calculated as ratio of the total sum of weighted cases (representing both eligible and ineligible cases) to the weighted sum of cases with known eligibility.

⁶ The population refers to the total residential addresses in the Address-based sampling frame in the NIRPC study area.

⁷ These categories were based on the call dispositions for RDD sampling frame and mail and/or call dispositions for the Address-based sampling frame.

$$NR1_h = \frac{\sum_{ER} W_{jh} + \sum_{ENR} W_{jh} + \sum_{IE} W_{jh} + \sum_{UK} W_{jh}}{\sum_{ER} W_{jh} + \sum_{ENR} W_{jh} + \sum_{IE} W_{jh}}$$

Where,

h	Adjustment cell (<i>i.e.</i> Sampling frame in this case)
W_j	Sampling Weight for telephone number j
ER	Eligible Respondent
ENR	Eligible Non-Respondent
IE	Ineligible Telephone number
UK	Phone Number with Unknown Eligibility

Similarly, the sample drawn from the address-based frame was first divided into four categories based on mail and/or call dispositions, following which the adjustments for unknown eligibility and nonresponse were calculated. Table 5 presents the adjustments for unknown eligibility by study area and sampling frame. The table shows that when the adjustments are applied, the weights of the telephone numbers/addresses with unknown eligibility status are distributed proportionally to the other three categories. This adjusts for the bias associated with failure to account for the cases with unknown eligibility status.

TABLE 5: ADJUSTMENTS FOR UNKNOWN ELIGIBILITY

Study Area	Sampling Frame	Eligibility Status	Weighted Sample ⁸	Adjustment for Unknown Eligibility	Adjusted Sample
Chicago	RDD	Eligible Respondents	28,485	2.24395	63,918
		Eligible Non-respondents	153,120	2.24395	343,593
		Ineligible	1,127,941	2.24395	2,531,044
		Unknown Eligibility Status	1,629,010	0	-
		Total	2,938,556		2,938,555
	Address-based	Eligible Respondents	52,488	3.96400	208,063
		Eligible Non-respondents	161,457	3.96400	640,017
		Ineligible	192,404	3.96400	762,688
		Unknown Eligibility Status	1,204,418	0	-
		Total	1,610,767		1,610,768
NIRPC	RDD	Eligible Respondents	4,955	1.91665	9,497
		Eligible Non-respondents	24,404	1.91665	46,775
		Ineligible	118,084	1.91665	226,327
		Unknown Eligibility Status	135,154	0	-
		Total	282,597		282,599
	Address-based	Eligible Respondents	9,141	3.16312	28,914
		Eligible Non-respondents	25,592	3.16312	80,952
		Ineligible	21,889	3.16312	69,238
		Unknown Eligibility Status	122,480	0	-
		Total	179,102		179,104

⁸ The sample cases were weighted by the 'sampling weight'.

In the second stage, we adjusted for the bias associated with non-response by applying an adjustment factor, denoted as $NR2_h$ that is the ratio of sampled eligible cases (including responding and non-responding cases) to those that completed the survey, as shown in the formula below:

$$NR2_h = \frac{\sum_{ER} (W_{jh} * NR1_h) + \sum_{ENR} (W_{jh} * NR1_h)}{\sum_{ER} (W_{jh} * NR1_h)}$$

Where,

h	Adjustment cell (<i>i.e.</i> Sampling frame in this case)
W_j	Sampling Weight for telephone number/address j
$NR1_h$	Adjustment for unknown eligibility for adjustment cell, h
ER	Eligible Respondent
ENR	Eligible Non-Respondent

Table 6 presents the adjustments for non-response by study area and sampling frame. The application of the adjustment factor increases the number of eligible respondents in the RDD and address-based sampling frames respectively.

TABLE 6: ADJUSTMENTS FOR NONRESPONSE

Study Area	Sampling Frame	Eligibility Status	Weighted Sample ⁹	Adjustment for Non-response	Adjusted Sample
Chicago	RDD	Eligible Respondents	63,918	6.37553	407,511
		Eligible Non-respondents	343,593	0	-
		Ineligible	2,531,044	1.00000	2,531,044
		Unknown Eligibility Status	-	-	-
		Total	2,938,555		2,938,555
	Address-based Sampling Frame	Eligible Respondents	208,063	4.07607	848,080
		Eligible Non-respondents	640,017	0	-
		Ineligible	762,688	1.00000	762,688
		Unknown Eligibility Status			
		Total	1,610,768		1,610,768
NIRPC	RDD	Eligible Respondents	9,497	5.92524	56,272
		Eligible Non-respondents	46,775	0	-
		Ineligible	226,327	1.00000	226,327
		Unknown Eligibility Status			
		Total	282,599		282,599
	Address-based Sampling Frame	Eligible Respondents	28,914	3.79975	109,866
		Eligible Non-respondents	80,952	0	-
		Ineligible	69,238	1.00000	69,238
		Unknown Eligibility Status			
		Total	179,104		179,104

⁹ The sample cases were weighted by the ‘sampling weight’ and ‘adjustment for unknown eligibility’.

ADJUSTMENT FOR MULTIPLE PHONE NUMBERS

The adjustment for multiple phone numbers adjusts for the high probability of selection associated with households with more than one landline. This adjustment factor, applicable to the RDD sampling frame only, is the reciprocal of the number of landlines owned by the household.

Typically, a weighting factor of 1 is assigned to households reporting only one landline in the household, and an adjustment factor of $\frac{1}{2}$ is assigned to households with more than one telephone number. However, considering the high percentage of households that own three or more phone numbers (about 7%), adjustment factors were assigned depending upon the number of landlines owned by the household. In particular, an adjustment factor of 1 was assigned to household reporting one landline, $\frac{1}{2}$ to households reporting two landlines, and $\frac{1}{3}$ to household reporting three or more landlines.

The application of the adjustment factor decreases the number of eligible respondents in the RDD sampling frame to 352,085 and 48,048 for Chicago and NIRPC study area respectively.

DUAL SAMPLING WEIGHT

The Dual Sampling Frame weight is calculated to adjust for households that have a higher probability of being selected in both the RDD and address-based sampling frame. In particular, there is a high probability that households that meet the following scenarios are present in both frames: (1) Households with more than one listed landline in address-based frame, (2) Household with one or more unlisted landlines in the address-based frame, and (3) Households with more than one landline in the RDD frame. To illustrate, suppose a household from an address-based frame when matched with a telephone directory shows that the household owns a listed landline, while the survey data indicates that they own two landlines. Clearly, the other landline is unlisted and there is a high probability of this household getting selected in the RDD sampling frame (that includes both listed and unlisted landlines). Thus, while we can check for duplicates between RDD and Address-based sampling frame based on listed landline numbers, we cannot account for the unlisted landlines (due to lack of information on the unlisted landline number). Following the weighting approach used by Brick *et. al.* (2006), a simple composite dual sampling weight of 0.5 was applied to the households that meet one of the aforementioned scenarios.

RAKING ADJUSTMENT

Raking improves the reliability of the survey estimates. Hence, raking adjustments were used to align the weighted sample to the 2000 Census data using raking variables. In particular, the aforementioned dual sampling composite weights were adjusted so that the sums of the adjusted weights are equal to known population totals from Census for certain subgroups of the population defined by demographic characteristics and geographic variables. The variables used for raking are as follows:

- Household size (one, two, three, four, five, six or more)
- Household Income (6 categories)
- Ethnicity (Hispanic, non-Hispanic)
- Race (White, African-Americans, Other)
- Age of the householder (less than 20 years, 20 – 24 years, 25 – 54 years, 55 – 64 years, 65 years or older)

- Strata (including the 5 pre-determined strata)
- County of residence

The above variables were chosen as the raking variables due to significant differences in the coverage by categories of these variables, and hence maximum bias reduction would be achieved using these variables. It is important to note that to calculate the raking adjustments, the missing values in the raking variables were imputed.¹⁰

The raking procedure is based on an iterative proportional fitting procedure and involves simultaneous ratio adjustments to two or more marginal distributions of the population counts. The raking procedure is carried out in a sequence of adjustments. First, the base weights are adjusted to one marginal distribution and then to the second marginal distribution, and so on. One sequence of adjustments to the marginal distributions is known as a cycle or iteration. The procedure is repeated until convergence is achieved.

Following the raking procedure, the inordinately large weights, a by-product of raking, were trimmed. These very large weights tend to substantially increase sampling errors. Thus, by not allowing weights to get too large, sampling errors are reduced although there is some loss in the bias reduction due to nonresponse adjustment and raking. It is important to note that trimming is used to reduce very large weights only and does not edit the data in any way. In particular, the ‘very large’ weights were trimmed to equal a maximum of twelve times the mean weight. After trimming the large weights, the raking process was repeated to align the survey estimates to the control totals.

Table 7 and 8 shows the sample and population distribution by demographic and geographic raking variables for the Chicago and NIRPC study area respectively. A comparison of the unweighted difference and weighted difference between the survey data and the census indicates that the raking procedure has aligned the sample statistics to the population statistics.

¹⁰ The missing values in the household income variable were imputed by taking the average of the income of households with similar demographic characteristics i.e. household size, household vehicle ownership, and home ownership status, and similar area of residence i.e. strata and county. Similarly, the missing values of age of the householder were imputed based on employment status, retirement status, presence of children in the household, and education status of similar households.

TABLE 7: RAKING VARIABLES FOR CHICAGO STUDY AREA

Post-Stratification Variables	Unweighted Data¹¹	Weighted Data	Census	Unweighted Difference	Weighted Difference
Household Size					
1	32.23%	26.38%	26.39%	5.84%	-0.01%
2	36.83%	28.71%	28.72%	8.11%	-0.01%
3	13.29%	16.00%	16.01%	-2.72%	-0.01%
4	11.36%	14.86%	14.85%	-3.49%	0.01%
5	4.61%	7.97%	7.97%	-3.36%	-0.00%
6+	1.68%	6.08%	6.05%	-4.37%	0.03%
Total	100.00%	100.00%	100.00%		
Household Income					
\$0-\$19,999	8.71%	16.83%	16.90%	-8.19%	-0.07%
\$20-\$34,999	9.75%	15.73%	15.70%	-5.95%	0.03%
\$35-\$49,999	12.93%	15.28%	15.30%	-2.37%	-0.02%
\$50-\$74,999	21.78%	20.90%	20.90%	0.88%	0.00%
\$75-\$99,999	19.81%	12.94%	12.90%	6.91%	0.04%
\$100k +	27.01%	18.31%	18.30%	8.71%	0.01%
Total	100.00%	100.00%	100.00%		
Ethnicity					
White	78.24%	65.42%	65.50%	12.74%	-0.08%
African American/Black	14.49%	18.93%	18.90%	-4.41%	0.03%
Other	7.27%	15.66%	15.60%	-8.33%	0.06%
Total	100.00%	100.00%	100.00%		
Hispanic					
Yes	4.87%	17.63%	17.20%	-12.33%	0.43%
No	95.13%	82.37%	82.80%	12.33%	-0.43%
Total	100.00%	100.00%	100.00%		
Respondent Age					
<20	0.17%	29.88%	29.50%	-29.33%	0.38%
20 - 24	0.98%	6.69%	6.70%	-5.72%	-0.01%
25 - 54	49.30%	44.90%	45.10%	4.20%	-0.20%
55 – 64	22.84%	7.93%	8.00%	14.84%	-0.07%
65 +	26.71%	10.60%	10.70%	16.01%	-0.10%
Total	100.00%	100.00%	100.00%		
Strata					
1	22.20%	22.68%	22.70%	-0.50%	-0.02%
2	25.81%	26.28%	26.30%	-0.49%	-0.02%
3	32.17%	30.64%	30.60%	1.57%	0.04%
4	13.30%	12.49%	12.50%	0.80%	-0.01%
5	6.53%	7.91%	7.90%	-1.37%	0.01%
Total	100.00%	100.00%	100.00%		
County					
Cook	66.20%	67.23%	67.15%	-0.95%	0.08%

¹¹ This unweighted statistics are based on the imputed data.

DuPage	9.41%	11.05%	11.08%	-1.66%	-0.03%
Grundy	0.64%	0.48%	0.49%	0.15%	0.01%
Kane	4.38%	4.54%	4.55%	-0.17%	-0.01%
Kendall	0.69%	0.64%	0.64%	0.05%	0.00%
Lake	9.39%	7.35%	7.36%	2.03%	-0.01%
McHenry	3.50%	3.03%	3.04%	0.46%	-0.01%
Will	5.78%	5.68%	5.70%	0.08%	-0.02%
Total	100.00%	100.00%	100.00%		

TABLE 8: RAKING VARIABLES FOR NIRPC STUDY AREA

Post-Stratification Variables	Raw Data	Weighted	Census	Raw Difference	Weighted Difference
Household Size					
1	28.22%	24.98%	24.99%	3.23%	-0.01%
2	43.98%	31.83%	31.83%	12.15%	0.00%
3	12.25%	17.40%	17.41%	-5.16%	-0.01%
4	9.98%	14.76%	14.75%	-4.77%	0.01%
5	4.01%	7.00%	7.00%	-2.99%	0.00%
6+	1.56%	4.03%	4.03%	-2.47%	0.00%
Total	100.00%	100.00%	100.00%		
Household Income					
\$0-\$19,999	11.46%	20.90%	20.90%	-9.44%	0.00%
\$20-\$34,999	12.87%	18.60%	18.60%	-5.73%	0.00%
\$35-\$49,999	15.95%	16.70%	16.70%	-0.75%	0.00%
\$50-\$74,999	26.32%	22.40%	22.40%	3.92%	0.00%
\$75-\$99,999	18.13%	11.70%	11.70%	6.43%	0.00%
\$100k +	15.27%	9.70%	9.70%	5.57%	0.00%
Total	100.00%	100.00%	100.00%		
Ethnicity					
White	87.39%	75.30%	75.30%	12.09%	0.00%
African American/Black	7.58%	18.20%	18.20%	-10.62%	0.00%
Other	5.03%	6.50%	6.50%	-1.47%	0.00%
Total	100.00%	100.00%	100.00%		
Hispanic					
Yes	3.91%	9.40%	9.40%	-5.49%	0.00%
No	96.09%	90.60%	90.60%	5.49%	0.00%
Total	100.00%	100.00%	100.00%		
Respondent Age					
<20	0.16%	29.10%	29.10%	-28.94%	0.00%
20 - 24	0.47%	6.40%	6.40%	-5.93%	0.00%
25 - 54	40.80%	42.80%	42.80%	-2.00%	0.00%
55 – 64	26.68%	9.00%	9.00%	17.68%	0.00%
65 +	31.89%	12.70%	12.70%	19.19%	0.00%
Total	100.00%	100.00%	100.00%		
Strata					
1	19.62%	24.60%	24.60%	-4.98%	0.00%
2	18.92%	22.80%	22.80%	-3.88%	0.00%
3	28.69%	20.60%	20.60%	8.09%	0.00%
4	22.38%	19.40%	19.40%	2.98%	0.00%
5	10.40%	12.60%	12.60%	-2.20%	0.00%
Total	100.00%	100.00%	100.00%		
County					
Lake	64.72%	65.50%	65.50%	-0.78%	0.00%
La Porte	14.30%	19.70%	19.70%	-5.40%	0.00%
Porter	20.97%	14.80%	14.80%	6.17%	0.00%
Total	100.00%	100.00%	100.00%		

FINAL HOUSEHOLD WEIGHT

The final analytic weight is simply the product of sampling weight, adjustment for unknown eligibility, adjustment for non-response, adjustment for multiple phone numbers, dual sampling frame weight, and raking adjustment. This weight was normalized to represent the number of survey respondents in the study area.

EXPANSION

The expansion process simply takes the weighted total (10,477 households for Chicago and 3,838 households for NIRPC) and multiplies each household by a factor that, when applied, will produce the household universe of 2,940,704 and 277,396 households for Chicago and NIRPC study area respectively. To derive the expansion factor, simple division was used: $\text{Expansion Factor} = N(\text{Universe}) / N(\text{Surveyed})$. The expansion factor was 280.681 for Chicago and 72.276 for NIRPC study area.