POVERTY TRACKER

Technical Notes

Last updated: May 17, 2023

ROBIN HOOD

CENTER ON POVERTY & SOCIAL POLICY at Columbia University

COLUMBIA | CPRC Columbia Population Research Center

This document contains technical notes on imputation and weighting for the Poverty Tracker Study. If you have any questions about the methodology or the study, please get in touch with us at povertytracker@columbia.edu.

Poverty Tracker Study Investigators: Qin Gao, Irwin Garfinkel, Ronald Mincy, Kathryn Neckerman, Julien Teitler, Jane Waldfogel, Christopher Wimer

Suggested citations:

For the Poverty Tracker Data

Poverty Tracker Research Group at Columbia University (2023). *Poverty Tracker Data*. Center on Poverty and Social Policy at Columbia University and Columbia Population Research Center. https://www.povertycenter.columbia.edu/poverty-tracker-data.

For the Poverty Tracker Technical Notes

Yajun Jia, Lauren Kennedy, Sophie Collyer, Matthew Maury, Schuyler Ross, Christopher Wimer (2023). *Poverty Tracker Data User Guide*. Center on Poverty and Social Policy at Columbia University and Columbia Population Research Center. https://www.povertycenter.columbia.edu/poverty-tracker-data.

Acknowledgments

We are grateful for the support and partnership of Robin Hood. We thank Lily Bushman-Copp, Chantal Bannerman, Chloé Cargill, Lolita Colon, Katherine Friedman, Janira Gayle, Anastasia Koutavas, Xiaofang Liu, Kahlen Washington, Nicholos Wilkinson, Ho Yan Wong, and Christopher Yera for their contributions and assistance in the preparation of this guide.

Disclaimer

Note that all documentation files are working documents and are subject to change as the study evolves.

Technical Note A: Imputation for the First, Second, and Third Poverty Tracker Cohorts

Overview

The Poverty Tracker uses a two-step multiple imputation approach to produce imputations for key variables. Multiple imputation is carried out primarily using the expectation-maximization with bootstrapping algorithm provided by the ***Amelia*** R package.[1] The multiple imputation methodology entails making standard assumptions about the underlying data generating process: (1) the complete data (after appropriate transformations of constrained variables) can be described by a multivariate Gaussian distribution; and (2) the data are *missing at random* (commonly abbreviated as MAR). In short, the MAR assumption says that missingness may depend on observed values, but it is independent of the missing values themselves. See Honaker (2011) for more details on these assumptions and the imputation method.[2]

The following notes describe the imputation process for all of the imputed variables in the Poverty Tracker data from the baseline annual surveys. This process is carried forward to produce imputations for all of the annual surveys. The imputed variables fall into seven groups:

(1) demographics,
(2) phone type (which is necessary for producing design weights),
(3) health,
(4) material hardship,
(5) income and expenses,
(6) housing, and
(7) services use frequency.

The approaches below are used when producing imputations for our first, second, and third Poverty Tracker cohorts, unless otherwise specified. The imputation methodology is being updated for the fourth and fifth cohort to account for the oversample of Chinese-origin New Yorkers. Documentation for the fourth cohort publicly available when it is finalized.

Variables with an asterix (*) are not includes in the public use dataset.

1. Demographics

The study imputes demographic information collected when respondents join the panel.
On the baseline annual survey, we impute:

imp_qfemale	Gender
imp_qeducat	Education level

[1] Honaker J, King G, Blackwell M (2011). “Amelia II: A Program for Missing Data.” *Journal of Statistical Software*, **45**(7), 1–47.
[2] See Honaker J, King G, Blackwell M (2011). “Amelia II: A Program for Missing Data.” *Journal of Statistical Software*, **45**(7), 1–47.

imp_qrace	Race/ethnicity
imp_qimmigrant	Immigration status
imp_qage	Age
imp_qbor	Borough

Note that we only impute these demographic characteristics based on baseline annual survey responses as these questions are not repeated in subsequent surveys.

2. Phone type

The study also imputes phone type information collected at the baseline survey. These variables are used when we produce design weights.

On the baseline annual survey, we impute:

imp_qcellphone*	Phone type respondent contacted on complete the intake survey (cellphone/landline)
imp_qi10*	Has a working landline inside the house
imp_qi11*	Anyone in household has a working cellphone
imp_qi12n*	Number of landline telephone numbers inside the house
imp_qi13n*	Number of cellphone numbers at which the respondent or spouse/partner can be reached (excluding work phones)

Note that for all of these phone-type variables, we only produce one set of imputations based on responses to the baseline annual survey as these questions are not repeated in subsequent surveys.

Note that we only impute these phone-type variables based on baseline annual survey responses as these questions are not repeated in subsequent surveys.

3. Health

The study also imputes health information collected at each annual survey.

On the baseline annual survey, we impute:

imp_qhealth	Self-reported health
imp_qhealthlim	Working limiting health condition
imp_qshealth	Spouse/partner health
imp_qshealthlim	Spouse/partner working limit health condition
imp_qsadblue	Felt sad or blue in past 30 days
imp_qtense	Felt tense in the past 30 days
imp_qdistress	Felt distress in the past 30 days

The corresponding variables on subsequent annual surveys are also imputed. Following the Poverty Tracker naming convention (see Appendix A), these variables are named according to their survey wave. For example, health status on the 12-month annual survey is *imp_q4qhealth*, and on the 24-month survey, it's *imp_q8qhealth*, and so on.

4. Material Hardship

The study imputes hardship variables at each annual survey. The hardship variables measure housing, bill, medical, financial, and food hardships in the past 12 months.

On the baseline annual survey, we impute:

imp_qhous1	Fail to pay rent or mortgage because of financial problems
imp_qhous2	Move in with other people because of financial problems
imp_qhous3	Stay at a place not meant for regular housing
imp_qbill1	Fail to pay phone, gas, oil or electricity bill
imp_qbill2	Cut off of phone, gas, oil or electricity service
imp_qmedic1	Not able to meet medical needs
imp_qfinanc1	Run out of money between paychecks or before the end of the month
imp_qfood1	Food eaten situation in the household
imp_qfood2	Worry about running out of food
imp_qfood3	Fail to buy enough food

The corresponding variables on subsequent annual surveys are also imputed. Following the Poverty Tracker naming convention (see Appendix A), these variables are named according to their survey wave. For example, health status on the 12-month annual survey is *imp_q4hous1*, and on the 24-month survey, it's *imp_q8hous1*, and so on. These different hardship material variables are combined to create our composite material hardship measures, as defined in the Core Measures section of the **Poverty Tracker Data User Guide**.

5. Income and Expenses

The income- and expense-related variables are imputed conditioning on available demographic information (e.g., age, race, gender, education level, immigration status, etc.). Since the imputations are performed jointly, we also use information from the observations of the other survey variables being imputed.

Earnings of respondent and spouse/partner

For survey questions related to the earnings of the respondent and their spouse or partner, respondents can provide:

1. A continuous numerical value (a dollar amount) or
2. a categorical value (an earnings bracket, provided if the respondent reported that they did not know or refused to give a numerical value).

For respondents with missing continuous values but observed categorical values, we impute a continuous value for them according to the distribution of the continuous values for respondents in the same earnings bracket. If neither a dollar amount nor a bracket is provided, then we directly impute a positive dollar amount for respondents who reported working in the past 12 months.

Some respondents with missing values on the earnings question also do not report the number of months that they and/or their spouse or partner worked during the previous year. For these cases we also impute the number of months worked (between 0 and 12) and subsequently we only require the imputations for earnings if the number of months worked is greater than zero.

On the baseline annual survey, we impute:

imp_qmoswork	Respondent months worked
imp_qmosworksp	Spouse/partner months worked
imp_qearn	Respondent annual earnings
imp_qearnsp	Spouse/partner annual earnings

The corresponding variables on subsequent annual surveys are also imputed. Following the Poverty Tracker naming convention (see Appendix A), these variables are named according to their survey wave. For example, respondent earnings in the 12-month annual survey is *imp_q4earn*, and on the 24-month survey, it's *imp_q8earn*, and so on.

Income from other family members in household; expenses related to childcare, work, and medical out-of-pocket spending

Like the questions about earnings for the respondent and their spouse or partner, the questions about income from other family members in the household and about expenses can be answered either by providing a numerical dollar value or a categorical value indicating an income bracket. If neither is provided then we directly impute a dollar amount. If the dollar value is missing but the respondent provides a categorical value then we impute a continuous value according to the distribution of the continuous values in the same income bracket.

On the baseline annual survey, we impute:

imp_qincothhh	Income from other family members in the same household
imp_qmoop	Medical out-of-pocket spending
imp_qchwoop	Capped out-of-pocket spending on childcare and work expenses[3]
imp_qwoop	Out-of-pocket spending on work-related expenses
Imp_qchoop	Out-of-pocket spending on childcare

The corresponding variables on subsequent annual surveys are also imputed. Following the Poverty Tracker naming convention (see Appendix A), these variables are named according to

[3] Following the Supplemental Poverty Measure methodology, we cap the combined value of work and childcare expenses so that they do not exceed the minimum reported earnings of the respondent or their spouse/partner (if applicable). See Fox, L. (2020). The supplemental poverty measure: 2019. *Current population reports.* Access at: https://www.census.gov/content/dam/Census/library/publications/2020/demo/p60-272.pdf.

their survey wave. On the 12-month annual survey, it is *imp_q4incothhh*, and on the 24-month annual survey, it's *imp_q8incothhh*, and so on.

Income from other sources

For each survey question related to other sources of income excluding respondent and/or their spouse or partner's earnings and income from other family members in the, we use a two-stage imputation process. First, we impute a binary value indicating whether or not the respondent received this type of income. Conditional on receiving this type of income, we then impute a positive dollar value for the amount received.

On the <u>baseline</u> annual survey, we impute:

imp_qretyes	Receive retirement income (including Social Security or survivor's benefits)
imp_qdisyes	Receive disability income
imp_qwelfyes	Receive income from welfare payments
imp_quiyes	Receive income from unemployment payments
imp_qsnapyes	Receive income from SNAP food assistance program
Imp_qwicyes	Received income from WIC (Women, Infant and Children Nutrition Program
imp_qregyes	Receive income from regular financial assistance from someone outside the household
imp_qothyes	Receive income from other sources
imp_qlunch	Children receive free or reduced priced school lunches
imp_qincret	Income from retirement funds
imp_qincdis	Income from paid disability
imp_qincwelf	Income from welfare payments
imp_qincui	Income from unemployment payments
imp_qincsnap	Income from SNAP food assistance program
imp_qincreg	Income from regular financial assistance from someone outside the household
imp_qincoth	Income from other sources

The corresponding variables on subsequent annual surveys are also imputed. Following the Poverty Tracker naming convention (see Appendix A), these variables are named according to their survey wave. For example, receipt of disability income on the 12-month annual survey is *imp_q4disyes*, and on the 24-month survey, it's *imp_q8disyes*, and so on.

6. Housing

We also impute housing status variables collected at each annual survey.

On the baseline annual survey, we impute:

imp_qmortgage	Value of mortgage
imp_qgovhous	Lives in public housing or receives government rental assistance
imp_qrcontrol	Lives in a rent controlled or stabilized apartment
imp_qrent	Monthly rent
imp_qbedrooms	Number of bedrooms

The corresponding variables on subsequent annual surveys are also imputed. Following the Poverty Tracker naming convention (see Appendix A), these variables are named according to their survey wave. On the 12-month annual survey, it is *imp_q4rent*, and on the 24-month annual survey, it's *imp_q8rent*, and so on.

7. Service use frequency

The study also imputes service use frequency information collected at the baseline survey for the first and second panels.

On the baseline annual surveys for these panels, we impute:

imp_qservfreq	Frequency of using social service

The corresponding variables on subsequent annual surveys are also imputed. Following the Poverty Tracker naming convention, these variables are named according to their survey wave. For example, respondent frequency of using social service in the 12-month annual survey is *imp_q4servfreq*, and on the 24-month survey, it's *imp_q8servfreq*, and so on.

Technical Notes B: Weighting the First, Second, and Third Poverty Tracker Cohorts

Overview

This note details the study's approach to survey weighting, by which the Poverty Tracker sample aims to be representative of New Yorkers aged 18 and over. Survey weights adjust statistical parameters (estimates) so that inferences made from the weighted data could be applied to the overall population from which the sample was drawn (in this case, adults in New York City). Sample weights are constructed by weighting the sample to a three-year American Community Survey (ACS) dataset.[4] The ACS is a nationally and regionally representative survey conducted by the U. S. Census Bureau.[5] The weighting approach employed in the Poverty Tracker study adjusts for oversampling, random over- or under-representation, non-response, and attrition. The approach follows many nationally and locally representative studies.

The remainder of this note describes:

(1) the Poverty Tracker cohorts and their sampling frames;
(2) the survey weights construction for the Poverty Tracker's **baseline** annual surveys, and;
(3) the survey weights construction weights for **subsequent survey waves post-baseline**.

This note focuses on weights constructed for the first, second, and third Poverty Tracker cohorts. The study's weighting methodology is being updated for the fourth and fifth cohorts to account for the oversample of Chinese-origin New Yorkers. That documentation will be publicly available when finalized.

I. Poverty Tracker Cohorts

The study recruits the bulk of respondents in the Poverty Tracker cohorts using a Random Digit Dial (RDD).[6] In addition, some cohorts have been supplemented with additional subsamples, as described in the **Poverty Tracker Data User Guide**. See Table TN1 for the sample size of the various Poverty Tracker RDD samples and other subsamples.

Table TN1. Sample Size, Poverty Tracker Cohorts

Cohort	Sample Size of Subsamples			Total Sample Size
	RDD	Social Service Users	Chinese Oversample	
1	2,002 total 500 cellphone 1,502 landline	226	*n/a*	2,228
2	3,403 total	505	*n/a*	3,908

[4] Table TN2 lists years of data from the ACS to which we weight each panel
[5] Learn more about the ACS at: https://www.census.gov/programs-surveys/acs
[6] The RDD is conducted by a survey research firm that manages this component of recruitment; a Columbia-based team manages all subsequent surveying. See the Poverty Tracker User Guide for additional information.

	1,904 cellphone 1,499 landline			
3	853 total 551 cellphone 302 landline	*n/a*	*n/a*	853
4	1,491 total 897 cellphone 594 landline	*n/a*	421[7]	1,912

*Details regarding the weighting methods for the fourth cohort are discussed in a separate technical note which is not yet publicly available.

All cohorts include both landline and cell phone samples in the sampling frame. In addition, Cohort 1 included an oversample[8] of landline numbers from high-poverty (greater than 20% poor) zip codes, but the study did not carry this design forward when recruiting our subsequent cohorts.

II. Constructing the Poverty Tracker Baseline Weights

Even with careful planning with the sampling and data collection, the sample you end up reaching might not match the population that you were aiming to recruit. This could be due to factors such as unequal selection probability for subgroups, high non-response from certain groups, or a sample frame that did not perform as expected. Estimates produced with this initial/unweighted sample could be biased and are not representative of the target population. Researchers often use survey weighting, which includes adjusting the design effects and post-stratify to known population proportions, to mitigate the design effects and any sample imbalances from the population.

The first part in Poverty Tracker weighting adjusts for any design effects, which account for the different probabilities of being sampled that respondents may have. By doing so, the weights aim to address any over- or under-represented issues. The second part in Poverty Tracker weighting is post-stratification, which accounts for differences in the propensity to respond. In this case, with information about the population, weights could be applied to “correct” the sample. Specifically, the raking procedure is used, which unlike other calibration methods, requires just marginal population counts.[9]

Below, the sections provide details on how the Poverty Tracker construct weights.

[7] Out of 421 cases, 196 cases were recruited using a Random Digit Dial (RDD) targeting geographic areas with high density of residents with Chinese origin, and the remaining 225 cases were recruited using community sampling (WeChat community groups).

[8] Oversampling on specific populations characteristics (e.g., poverty) is a statistically appropriate and efficient way to increase the sample sizes of populations of interest in surveys. The Poverty Tracker study is focused on dynamics of poverty and hardship, so we oversampled low-income neighborhoods and low-income individuals who use social services. The alternative would have been to draw a much larger sample to yield equivalent statistical power. The oversample of Robin Hood agencies had the added benefit of providing information about the population directly supported by Robin Hood-funded programs.

[9] Lumley, T. (2010) Complex Surveys: A Guide to Analysis Using R. John Wiley and Sons, Hoboken, Washington. http://dx.doi.org/10.1002/9780470580066

Constructing the weights for data collected on the baseline annual survey consists of four steps:

Step 1: Adjust for probability of selection and weight the RDD sample to match the New York City adult population in the year of the cohort's recruitment (represented by the New York City sample in the ACS).
Note: For the first and second cohorts, we need to make additional adjustments for the service-user sample, detailed in steps 2 and 3. For the third cohort, we do not need further adjustments other than step 1.
Step 2: Adjust for the probability of selection into the service-use sample (Cohorts 1 and 2 only).
Step 3: Use the weighted RDD sample to estimate the population distribution of social service users (Cohorts 1 and 2 only).
Step 4: Combine the RDD and service-user samples and weight to match social service users and demographics (Cohorts 1 and 2 only).

Below, we provide more details about how we took each step above.

Step 1: Adjust for probability of selection for the RDD sample, and weight the RDD sample to match the NYC adult population (represented in the ACS).

Using the RDD sample only, we adjust for selection bias and nonresponse to match the ACS data through the following methods:

1a. Adjustment for the number of adults in the household and family.

The first form of design effect in PT weighting is to adjust for variations due to household and family sizes. The larger the household, the smaller the selection probability is for each individual in the household. However, the larger the family, the larger the response probability from the family. It is therefore necessary to weight up individuals in larger families, while weighting down larger families. Gelman and Little (1998)[10] argued that inverse probability weights for household sizes tend to overcorrect in telephone surveys and they recommended using square roots for this weighting adjustment. Thus, the study uses the Gelman and Little (1998) approach and takes the square root of the ratio of the number of adults in the household to the number of adults in the family. Note that after each step of the weighting, it generates a set of weights (for all respondents). Eventually, we multiply all weights coming from different steps together, which will be the base weight that comes into the post-stratification.

1b. Adjustment for phone availability.

The next form of design effects that the Poverty Tracker weights adjust for are the variations in probability of selection due to phone service interruptions and the availability of landlines/cellphones. The more phones that are available in the household, the larger the

[10] Gelman, A. and T. C. Little (1998). *Improving on probability weighting for household size.* Public Opinion Quarterly} 62(3), 398--404.

selection probability is for the household. However, those who experience interrupted phone service have a smaller probability of selection. It is therefore necessary to weight up respondents with interrupted service while weighting down respondents with multiple phones in the household. This stage of the weighting process assigns respondents in these two categories weights of ½ and 2, respectively.

In addition, because the landline and cellular RDD frames overlap, there are cases of dual-service, that is, respondents from the landline sample who also have a cell phone or respondents from the cell phone sample who also have landline service in the household. The study uses frame integration weights (Lohr, 2009)[11] to combine the landline and cellular components of the sample, with the dual-service respondents from the two frames integrated in proportion to their effective sample sizes. This integration assumes that the dual-service households from each of the two groups are random samples from the population of dual-service households.

To compute the effective sample sizes, we first calculate a design effect for both the landline and cellular groups. For the cellular sample, this requires taking the weights for the respondents who also have landlines and compute the coefficient of variation cv_C. For the landline sample the calculation of cv_L is analogous. We take the design effects to be $1+cv_C^2$ and $1+cv_L^2$, respectively. The effective sample sizes (ESS) for the dual-service cases are then computed as the raw sample sizes divided by the design effects.

Finally, the frame integration weights for the dual-service cell phone cases (cell phone respondents who also have a landline) are $fiw_C = \frac{ESS_C}{ESS_C + ESS_L}$ which is the ratio of the effective number of dual-service cases among the cell phone respondents to the total effective number of dual-service cases in the landline and cellphone groups combined. For the dual-service landline respondents the frame integration weights are computed analogously as $fiw_L = \frac{ESS_L}{ESS_C + ESS_L}$. Single-service cases (i.e., landline-only or cell-only) are given a frame integration weight of 1.

*1c. Adjustment for oversampling households in high-poverty neighborhoods (**only in Cohort 1**).*

The study adjusts for the oversampling of poor households by first obtaining the total number of households by zip code from the 2000 Decennial Census for New York City. Second, the zip codes are divided into three strata by poverty rates (below 10%, 10% to 20%, above 20%) and then the number of households in each stratum are counted. These marginal frequencies are taken to be the population information and are used as the benchmark for matching the weighted sample. Finally, Poverty Tracker respondents are stratified by poverty level and their matching their zip codes to the ACS zip codes[12] to adjust

[11] Lohr, S. (2009). Multiple-frame surveys. In D. Pfeffermann and C. R. Rao (Eds.), *Handbook of Statistics, Vol. 29A: Sample Surveys: Design, Methods and Applications*, Chapter 4. Elsevier/North Holland.
[12] Only 1,976 of the 2,002 cases have zip code information we can match to the ACS data. We match their strata number with those in the ACS. Zip codes 11247 and 11249 are known to be oversampled and belong to the third

for the oversampling on zip codes. Then the sample is post-stratified by matching the sample margins to the population (ACS) margins and post-stratification weights are trimmed at the 99% quantile.

1d. Post-stratification adjustment to corresponding weighted ACS NYC totals.

The accuracy of the estimates is improved if we adjust the composition of the survey sample to the known population composition. As such, the weighted estimates more accurately represent the characteristics of the population.

The most commonly used adjustment method for making a sample representative of the population is known as *poststratification*, or *raking*. Basic selection bias and nonresponse-adjusted sampling weights (see more details for the adjustments in previous steps 1a to 1c) are used as an input for the raking process. We use the R ***Survey*** package[13] to rake the weights to known margins in the population. We use three years of data on the New York City population from the American Community Survey (ACS) to represent the New York City population. Table TN2 lists years of data from the ACS to which we rake to in each cohort.

Table TN2. ACS Datasets used in Post-Stratification, by Cohort

	ACS Years (NYC Sample)
Cohort 1	**2011-2013**
Cohort 2	**2014-2016**
Cohort 3	**2016-2018**
Cohort 4	**2019-2021**

Before the post-stratification, we first obtain individual weights by multiplying the family weights obtained in the previous step by the number of adults in the household.[14]

Next, the known marginal distributions of post-stratification variables in the sample and the population are matched via a raking procedure. The population data used for post-stratification comes from the American Community Survey's (ACS) New York City sample. The population distributions of the post-stratification variables are approximated using weighted ACS data.[15] The post-stratification variables include gender, age, education, race, the number of children in the family, the number of seniors in the family, the number of working aged (18-64) people in the family, an income-to-needs measure for the family[16], and

strata. For the remaining cases, we randomly assign them into the first two strata with probability proportional to strata size.

[13] Lumley, T. (2014). Survey: analysis of complex survey samples. *R package*, version 3.30.

[14] For this calculation the number of adults in the household is capped at 4 due to sparseness at larger values.

[15] That is, we use the weights provided with the ACS and obtain a weighted frequency for each post-stratification variable.

[16] From the World Bank: Poverty gap is the mean shortfall from the poverty line (counting the non-poor as having zero shortfall), expressed as a percentage of the poverty line. This measure reflects the depth of poverty as well as its incidence.

interactions between many of the demographics and the income-to-needs measure to account for dependencies between these factors. After the raking procedure, resulting weights are trimmed at the 97.5% percentile.

Step 2: Adjust for the probability of selection for the service-use sample.

For Cohorts 1 and 2, the next step in the weighting procedure adjusts for the probability of being interviewed among the social-service user sample using the frequency of service use. Similar to the RDD household adjustment, the more frequent the respondents use social service, the highly probability of being selected during recruitment visits to service agencies. Thus, it is necessary to weight up the less-frequent service users while weighting down the more-frequent service users. This stage of the weighting process assigns respondents inverse probability weights for service use (for both RDD and service-use samples) to adjust for the selection probability variations due to the frequent of service use.

Step 3: Use the weighted RDD sample to estimate the population distribution of social service users.

Data on use of social services is not available in the ACS data, therefore it is not possible to simply match respondents from the social service user subsample to corresponding New Yorkers in the ACS in the post-stratification. Therefore, before combining the RDD and the service-use samples, it is necessary to use the weighted RDD sample (from Step 1) to estimate the population distribution of social service users in the ACS data. Since the weighted RDD sample is assumed to be representative of the ACS data, therefore, the frequency of service use in the RDD sample can be used to estimate the frequency of service use in the ACS population data.

Step 4: Combine the RDD and service-user samples and weight to match social service user and demographics

Finally, we combine the service-user and RDD samples to construct weights for all respondents.[17] We again correct for differences due to over-sampling low-income individuals in the service-use sample by post-stratifying the family weights to the equivalent ACS information. This is essentially the same adjustment that is made for the RDD weights (step 4 adjustment in the RDD section) but here we perform the adjustment on the combined Agency and RDD samples.

We first produce poverty-unit weights that are then multiplied by the number of adults in the family to obtain the person (individual) weights. These person-level weights are then post-stratified using the ACS-NYC data to adjust for deviations of the two samples from the corresponding ACS-NYC weighted totals. The frequency of social service use is also included as a post-stratification variable.[18]

[17] The weights for each sample are also separately normalized to each have a mean of 1.
[18] There will be unbalanced coverage of agency service visitors because frequent service users will be over-represented in the Agency sample. For the purpose of representing the general population of NYC adults, it is necessary to down-weight individuals in the sample who frequently use social service agencies. In order to post-

III. Constructing weights for subsequent survey waves

After constructing the baseline weights, longitudinal weights are constructed for subsequent survey waves. These weights use the baseline weights as a basis, but make two adjustments. The first adjustment corrects for non-response between the baseline and wave of interest. The second adjustment is a raking adjustment back to the baseline population demographics. The method for the first adjustment is inverse propensity scoring. This method uses a logistic regression to predict non-response given baseline characteristics.[19] From this model, the probability of responding to a follow-up survey is predicted for each individual, which we then break into 20 quantiles to reduce noise. For each quantile, the inverse of the probability of responding is calculated, which we then use to adjust the baseline weights for nonresponse. Weights are then raked and trimmed, and replicate weights are created with the same technique as at baseline.

IV. Replicate Weights

Altogether, the Poverty Tracker weights adjust for unequal selection, under-coverage and nonresponse, but there is no simple formula for estimating the error around estimates produced with the weights (such as the variance around the weighted poverty rate). For this reason, there are also replicate weights for each set of sampling weights are available for data users that were produced using the bootstrapping method implemented in R's ***Survey*** package.[20]

In the PT data, there are 50 separate replicate weights at the person and poverty-unit levels that allow users to generate empirically derived standard error estimates. In theory, the standard error of an estimate measures the variation of a statistic across multiple samples of a given population. A true standard error of estimates from a single sample can never be known with certainty. Replicate weights allow a single sample to simulate multiple samples, therefore generating more informed standard error estimates that mimic the "true" standard errors while retaining all information about the complex sample design. For example, the replicate weights can be used to produce margins of error around the estimated poverty rates calculated using the weighted data. More details about how to implement Poverty Tracker weights and replicate weights in the analysis can be found in **Guide for Using PT Weights**.

stratify on frequency of service use we need a measure of the distribution of social service use in the population. Unfortunately, we do not have any gold standard for the distribution of service use in the general population. Instead, we estimate it from the responses in the now weighted phone sample. We match the frequency of social service use from the combined sample (RDD and Agency) to the frequency estimated only using the RDD sample.

[19] These characteristics include race and ethnicity, education level, how the participants were originally contacted, service frequency, marital and cohabiting status, experiences of severe material hardship, number of working age adults in the family, number of elderly adults in the family, whether born in another country, experiences of health problems, income level (measured using the log of their income to needs), poverty status (measured under the official poverty measure), residence in public housing, and mental health status.

[20] Lumley, T. (2014). *survey*: analysis of complex survey samples. R package version 3.30.

Appendix A. Poverty Tracker variable naming convention

Naming convention
All Poverty Tracker variables have a **prefix** that indicates the survey where the data stored in the variable was collected**.** Nearly all variables begin with a "***q***," for questionnaire, followed by a number corresponding to a particular survey. Table 3 maps the variable prefixes to the corresponding survey wave.

Table 3. Survey waves and variable prefixes for Poverty Tracker data

Survey	Baseline	3-month	6-month	9-month	12-month	15-month	18-month	21-month
Prefix	q	q1	q2	q3	q4	q5	q6	q7
Survey	24-month	27-month	30-month	33-month	36-month	39-month	42-month	45-month
Prefix	q8	q9	q10	q11	q12	q13	q14	q15
Survey	48-month	60-month	72-month					
Prefix	q16	q17	q18					

Note: This table provides a guide to understanding the variable naming structure in the Poverty Tracker data. A "q" denotes each questionnaire, followed by the associated survey number.

You can also use the prefix to identify if the variable has been imputed (that is, if the missing values have been imputed) because the variable will have the prefix "***imp_***" preceding the prefixes in Table 3.

The variables also have a **suffix** that indicates one of three things: (1) If the variable directly matches a survey question and has not been imputed,[21] or (2) If the variable matches a survey question but was imputed, if (3) if the variable does not match a survey question, but was constructed based on responses to survey questions. You can also use the suffix to identify if it's a top coded variable because you'll see the suffix "***_tc***" at the end of the variable name.

See the **Poverty Tracker User Guide** to learn more.

[21] Variable names in the surveys may not line up perfectly with those in the public datasets. Instead you should reference the codebooks to confirm the meaning of each variable in the dataset. Please reference the variable crosswalk document for information on how survey questions are mapped to final variable names.