General Overview of the Project

- Goals
- Core measures and additional topics covered
- Cohorts, timeline, and evolution
- Survey schedule
- Pooling Poverty Tracker cohort data

Imputation and Weighting

- Imputation methods
- Weighting methods
- Types of weights and how to use them

Accessing Poverty Tracker Data and Documentation
Overview of the Poverty Tracker

Original goals of the project

Get a better measure of poverty in New York City.

Provide a more comprehensive understanding of disadvantage beyond poverty.

Understand the dynamics of poverty and disadvantage how these experiences respond to policy interventions and other life events and circumstances.
Poverty Tracker Topics Covered: Core Measures

Three Core Measures of Disadvantage

*All Collected on the Annual Poverty Tracker Surveys*

The Poverty Tracker collects all data necessary to measure poverty under the *Supplemental Poverty Measure.*
Poverty Tracker Topics Covered: Core Measures

Three Core Measures of Disadvantage

All Collected on the Annual Poverty Tracker Surveys

Measure material hardship in five domains:

Poverty Tracker Topics Covered: Core Measures

Five Domains of Material Hardship
Poverty Tracker Topics Covered: Core Measures

Three Core Measures of Disadvantage

*All Collected on the Annual Poverty Tracker Surveys*

- Income Poverty
- Material Hardships
- Health Problems

Health problems defined as having:

- Poor Health
- Work-Limiting Health Condition

**NYC Households Experiencing Some Disadvantage**

- The reality is worse than the official statistics (53%)
- 37% Experience at least one disadvantage: severe health problem, poverty or severe material hardship
- 23% Suffer at least one severe health problem
- 23% Under the supplemental poverty line
- 21% Under the official poverty line
Poverty Tracker Topics Covered: “Shocks”

In addition to our core measures, the Poverty Tracker includes a series of questions about changes and shocks that a respondent might have experienced, including:

<table>
<thead>
<tr>
<th>Starting or losing a job</th>
<th>Having a child</th>
<th>Having an accident or injury</th>
<th>Being a victim of a crime</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginning to receive benefits or losing benefits</td>
<td>Having a major increase or decrease in income</td>
<td>Being arrested or being stopped by police</td>
<td>Lost/broke expensive belongings or having a major unanticipated expense</td>
</tr>
<tr>
<td>Changes in household composition and moving</td>
<td>Starting or ending a relationship</td>
<td>Changes in childcare</td>
<td>Having to take time off work to talk with your child’s teacher or guidance counselor</td>
</tr>
</tbody>
</table>

These questions are included on every survey for the first three cohorts in reference to the previous three-months, thus tracking various changes in the lives of New Yorkers.
Likelihood of Falling into Poverty for New York City Adults by Life Event


Poverty Tracker Topics Covered

- Service Use
- City Services
- Assets & Debts
- Housing
- Student Loans
- Consumption
- Employment
- Sleep
- Immigration
- Commuter Benefits
- Childcare
- Sick Leave
- Caretaking
- Education History
- Child health
- Caregiving
- Discrimination
- COVID-19
- Affording food
- Aging
- Child development
- Discrimination
- COVID-19
- Affording food
- Aging
- Minimum Wage
- Policing

*Not exhaustive, please see Poverty Tracker Data User Guide*
Cohorts, Timeline, and Evolution

The study surveys a representative sample of adult New Yorkers multiple times a year for a period of up to six years. (Originally planned to just survey for two years but later extended to six.)

Since initial launch, the study recruited five distinct samples of New Yorkers: Cohorts 1 through 5 (more details to come!)

The majority of these samples were recruited through a Random Digit Dial (RDD) methodology. Periodically includes supplementary samples, including:

- Respondents recruited from social service agencies (in the first two cohorts)
- New Yorkers of Chinese-origin (in the fourth and fifth cohorts)

---

Cohorts, Timeline, and Evolution

|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

**Cohort 1**

(2012 - 2014)

- **Cohort 1** recruited in 2012
- Respondents interviewed for up to two years, with the final surveys launching in 2014
- Baseline sample size: **2,228**
- Subsamples:
  1) RDD methodology: **2,002** respondents (Incl. oversampling of New Yorkers in high-poverty zip codes)
  2) Randomly selected social service participants: **226** respondents (Oversample of social service users that may be disproportionately low-income)
- All of the **Cohort 1** data is publicly available on the CPSP website.
Cohorts, Timeline, and Evolution

<table>
<thead>
<tr>
<th>Year</th>
<th>Cohort 2</th>
<th>Cohort 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2021</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2022</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2023</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2024</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2025</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2026</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2027</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2028</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Cohort 1** (2012 - 2014)
- Baseline launch year: 2012
- Final survey launch year: 2014
- Sample size at baseline: 2,228

**Cohort 2** (2015 - 2023)
- **Cohort 2** recruited in 2015
- **Design change**: Respondents Interviewed for up to six years (instead of two), with the final surveys launched in 2023 (pandemic caused some delays)
- Baseline sample size: **3,908**
- Subsamples:
  1) RDD methodology: **3,403** respondents
     (Recruited from pool of 10,000 individuals who completed the NYC DOHMH’s 2015 Community Health Survey)
  2) Randomly selected social service participants: **505** respondents
- **The first four years** of the Cohort 2 data are **publicly available** on the CPSP website

**Cohort 3** (2017 - 2025)
- **Cohort 3** recruited in 2017
- **Design change**: Moved to rotating cohort design
- Final surveys scheduled for 2025
- Baseline sample size: **853**
- Data available upon request
## Cohorts, Timeline, and Evolution

<table>
<thead>
<tr>
<th>Year Range</th>
<th>Baseline Launch Year</th>
<th>Final Survey Launch Year</th>
<th>Sample Size at Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012-2014</td>
<td>2012</td>
<td>2014</td>
<td>2,228</td>
</tr>
<tr>
<td>2015-2023</td>
<td>2015</td>
<td>2023</td>
<td>3,908</td>
</tr>
<tr>
<td>2017-2025</td>
<td>2017</td>
<td>2025</td>
<td>853</td>
</tr>
<tr>
<td>2020-2027</td>
<td>2020</td>
<td>2027</td>
<td>1,912</td>
</tr>
<tr>
<td>2022-2028</td>
<td>2022</td>
<td>2028</td>
<td>1,548</td>
</tr>
</tbody>
</table>

- **Cohort 4** recruited in **2020**
- **Design change**: Introduced Chinese oversample
- Final surveys scheduled for **2027**
- Baseline sample size: **1,912**
- Additional design changes made in 2022 (see Data User Guide)
- Data available upon request

### Cohorts

- **Cohort 1** (2012-2014)
  - Baseline launch year: 2012
  - Final survey launch year: 2014
  - Sample size at baseline: 2,228

- **Cohort 2** (2015-2023)
  - Baseline launch year: 2015
  - Final survey launch year: 2023
  - Sample size at baseline: 3,908

- **Cohort 3** (2017-2025)
  - Baseline launch year: 2017
  - Final survey launch year: 2025
  - Sample size at baseline: 853

- **Cohort 4** (2020-2027)
  - Baseline launch year: 2020
  - Final survey launch year: 2027
  - Sample size at baseline: 1,912

- **Cohort 5** (2022-2028)
  - Baseline launch year: 2022
  - Final survey launch year: 2028
  - Sample size at baseline: 1,548
Poverty Tracker Survey Schedule

**Cohort 1**

<table>
<thead>
<tr>
<th>Year 1</th>
<th>Annual surveys (Baseline)</th>
<th>Interim survey (3 month)</th>
<th>Interim survey (6 month)</th>
<th>Interim survey (9 month)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 2</td>
<td>Annual surveys (12 month)</td>
<td>Interim survey (15 month)</td>
<td>Interim survey (18 month)</td>
<td>Interim survey (21 month)</td>
</tr>
<tr>
<td>Year 3</td>
<td>Annual surveys (24 month)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Interim Surveys cover a variety of topics

Annual surveys repeatedly collects core measures (poverty, material hardship, and health problems) as well as housing, mental health, among other topics.

**Cohort 2**

<table>
<thead>
<tr>
<th>Year 1</th>
<th>Annual surveys (Baseline)</th>
<th>Interim survey (3 month)</th>
<th>Interim survey (6 month)</th>
<th>Interim survey (9 month)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 2</td>
<td>Annual surveys (12 month)</td>
<td>Interim survey (15 month)</td>
<td>Interim survey (18 month)</td>
<td>Interim survey (21 month)</td>
</tr>
<tr>
<td>Year 3</td>
<td>Annual surveys (24 month)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 4</td>
<td>Annual surveys (27 month)</td>
<td>Interim survey (30 month)</td>
<td>Interim survey (33 month)</td>
<td></td>
</tr>
<tr>
<td>Year 5</td>
<td>Annual surveys (36 month)</td>
<td>Interim survey (39 month)</td>
<td>Interim survey (42 month)</td>
<td>Interim survey (45 month)</td>
</tr>
</tbody>
</table>

Interim Surveys cover a variety of topics

Annual surveys repeatedly collects core measures (poverty, material hardship, and health problems) as well as housing, mental health, among other topics.
### Poverty Tracker Survey Schedule and Topics

<table>
<thead>
<tr>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Annual Baseline</strong>&lt;br&gt;Housing</td>
<td>Housing</td>
<td>Housing</td>
<td>Housing</td>
</tr>
<tr>
<td>Childcare</td>
<td>Childcare</td>
<td>Childcare</td>
<td>Childcare</td>
</tr>
<tr>
<td>Service Use</td>
<td>Service Use</td>
<td>Service Use</td>
<td>Service Use</td>
</tr>
<tr>
<td>Adult Health &amp; Well-Being</td>
<td>Adult Health &amp; Well-Being</td>
<td>Adult Health &amp; Well-Being</td>
<td>Adult Health &amp; Well-Being</td>
</tr>
<tr>
<td>Material Hardship</td>
<td>Material Hardship</td>
<td>Material Hardship</td>
<td>Material Hardship</td>
</tr>
<tr>
<td>Employment</td>
<td>Employment</td>
<td>Employment</td>
<td>Employment</td>
</tr>
<tr>
<td>3-Month</td>
<td>15-Month</td>
<td>27-Month</td>
<td>39-Month</td>
</tr>
<tr>
<td>Adult health</td>
<td>Housing, Consumption</td>
<td>Housing, Assets &amp; Debts</td>
<td>Adult &amp; Child Health &amp; Well-being, Caretaking</td>
</tr>
<tr>
<td>6-Month</td>
<td>18-Month</td>
<td>30-Month</td>
<td>42-Month</td>
</tr>
<tr>
<td>Service Use, City Services</td>
<td>Employment, Volunteering</td>
<td>Service Use, City Services, Employment</td>
<td>Housing, Employment, Commuter Benefits</td>
</tr>
<tr>
<td>9-Month</td>
<td>21-Month</td>
<td>33-Month</td>
<td>45-Month</td>
</tr>
<tr>
<td>Assets &amp; Debts</td>
<td>Adult &amp; Child Health, Immigration, Sick Leave</td>
<td>Fairness &amp; Opportunity, Discrimination</td>
<td>Housing, Sleep, Education History, Student Loans</td>
</tr>
</tbody>
</table>

*Not exhaustive, please see Poverty Tracker Data User Guide*

### Pooling Poverty Tracker Data from Multiple Cohorts

#### Pooling Cohort 1 and 2

<table>
<thead>
<tr>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Annual surveys (Baseline)</strong></td>
<td><strong>Annual surveys (12 month)</strong></td>
<td><strong>Annual surveys (24 month)</strong></td>
</tr>
<tr>
<td>Interim survey (3 month)</td>
<td>Interim survey (15 month)</td>
<td>Interim survey (21 month)</td>
</tr>
<tr>
<td>Interim survey (6 month)</td>
<td>Interim survey (18 month)</td>
<td>Interim survey (21 month)</td>
</tr>
<tr>
<td>Interim survey (9 month)</td>
<td>Interim survey (21 month)</td>
<td>Annual surveys (24 month)</td>
</tr>
</tbody>
</table>

**This is just one way to pool the data! Please reach out if you're interested in learning about other approaches!**

---

*Powered by The Robinhood Poverty Tracker*
Imputation & Weighting

Documents

• Technical Notes
• Guide for Using Weights
Dealing with Missingness

**Missing Data**

- **Common in survey research** - respondents might be skeptical about data protection, uncomfortable disclosing personal information, or experience survey fatigue...
- **Going back** to conduct the survey may not be the best solution because you’d have to spend twice as much time and money.
- Most commonly, folks **impute** due to issues of power associated with reduced sample size due to missingness.

**Imputation**

- Substituting missing information with an estimated value depending on other available information/data.
- Imputation preserves all cases so that we would be able to analyze “complete” data once all values have been imputed.

How the Poverty Tracker Deals with Missingness

**Missing at Random (MAR)**

- It’s an underlying assumption (missing at random but not completely at random) for Poverty Tracker imputations where we impute missing values using other information/characteristics in the data.
- MAR means there might be systematic differences between the missing and observed values, but these can be entirely explained by other observed variables.

**Multiple Imputation (MI)**

- Simple imputation assume perfect estimation of imputed values and ignore between-imputation variability.
- In contrast, MI generates replacement values (“imputations”) for missing data and repeating this procedure many times, resulting in many data sets with replaced missing information.
- Incorporates uncertainty into the standard errors of imputed values by accounting for variability between imputed solutions.

Acock, 2005; Graham, 2009; Hibel, Farkas, & Morgan, 2010; Schafer, 1999
Variables Imputed in Poverty Tracker Data

Demographics [Baseline]
- Demographic variables (e.g. education, immigration status)

Core Measures [Baseline & Post-Baseline]
- Health (e.g. limiting health condition, mental health)
- Material hardship (e.g. housing, bill, medical, financial, food hardships)
- Income and expenses
  - Earnings (e.g. respondent/spouse months worked, earnings)
  - Income from other sources (e.g. retirement, disability, welfare, unemployment, SNAP, WIC)
    - Two-step imputation
      - Binary value indicating whether or not received type of income
      - Amount of income
  - Income from other family members
  - Expenses (e.g. childcare, work, medical-related expenses)
- Housing (e.g. mortgage, rent)

Variables Imputed in Poverty Tracker Data

Impute Variables for Weighting
- Phone type (necessary for producing design weights; e.g. cellphone/landline)
- Service use frequency

Impute for Other Measures
- Assets and debts
Weighting

What is a Survey Weight?

• A value assigned to each case in the data, normally used to adjust statistical parameters/estimates so that inferences from the data are more representative of the target population.

• In this case, survey weight allows us to use Poverty Tracker sample to make statements about the city’s population.

• *Example*: statements about what share of adult New Yorkers or families are in poverty instead of statements just about the share of the sample.

Weights in the Poverty Tracker

Survey weights are used to ensure that the PT sample is representative of adults and poverty units (families) in New York City.

The weighting approach adjusts for **oversampling**, **random over- or under-representation**, **non-response**, and **attrition**.

*For example*, the value of the weight indicates how much each case will count in a statistical procedure. A weight of 2000 in Poverty Tracker data indicates that this weighted case is representative of 2000 New Yorkers.

To achieve this, Poverty Tracker weights are produced by…

• **Step 1**: Adjust for sampling design
• **Step 2**: Adjust for non-response - post-stratify to “population”
• **Step 3**: Adjust post-baseline surveys for sample attrition
Step 1: Adjust for Sampling Design

*How Design Weights Adjust for Sampling Design?*

- Design weights compensate for over- or under-sampling of specific cases or for disproportionate stratification.
- *Example:* oversample Chinese New Yorkers 4 times more than White New Yorkers…

*Design Weights in Poverty Tracker*

- Household size
- Phone availability
- Oversample households in high-poverty neighborhoods (cohort 1)
- Service use frequency

Gelman, & Little, 1998; Lohr, 2009

---

Step 2: Adjust for Non-response - Post-stratify to “Population”

*Post-stratification Weights Adjust for Non-response*

- Post-stratification weights compensate for the fact that persons with certain characteristics are not as likely to respond to the survey (e.g. more females than males). Corrects for departures from known population totals.
- Requires the use of auxiliary information about the population and may take a number of different variables into account.
- Commonly adjusted respondent characteristics: age, gender, education, race/ethnicity, etc.

*Post-stratification Weights in Poverty Tracker*

- “Population” data: data are weighted to three-year American Community Survey (ACS) dataset.
- Post-stratification variables include gender, age, education, race/ethnicity, number of children, number of seniors, number of working-aged people, income-to-needs measure, and interactions between many of the demographics and the income-to-needs measure to account for dependencies between these factors.
### Longitudinal Weights in Poverty Tracker

Longitudinal weights are computed from two components:
- The weight from the baseline survey
- The weight calculated to adjust for attrition between the waves
  - Logistic regression with response to the wave as outcome variable (0=no, 1=yes)
  - Predict probability of responding
  - **attrition** weight = Inverse the probability of responding

### Using Poverty Tracker Weights

<table>
<thead>
<tr>
<th>Wave</th>
<th>Person-level Weights</th>
<th>Poverty-unit Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>qweight_p</td>
<td>qweight_pu</td>
</tr>
<tr>
<td>3-Month</td>
<td>q1weight_p</td>
<td>q1weight_pu</td>
</tr>
<tr>
<td>6-Month</td>
<td>q2weight_p</td>
<td>q2weight_pu</td>
</tr>
</tbody>
</table>

Person-level weights: generalize to characteristics of adults in NYC…

“**What percent of adults experienced X in 2015?**” Apply **qweight_p**

Poverty-unit weights: generalize to characteristics of poverty units/families in NYC…

“**What percent of families (poverty units) experienced X in 2015?**” Apply **qweight_pu**
Replicate Weights

Replicate weights allow a single sample to simulate multiple samples, thus generating more informed standard error estimates that mimic the theoretical basis of standard errors while retaining all information about the complex sample design. These standard errors can then be used to obtain more precise confidence intervals and significance test. [Guide for Using Poverty Tracker Weights]

<table>
<thead>
<tr>
<th>Wave</th>
<th>Person-level Replicate Weights</th>
<th>Poverty-unit Replicate Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>qweight_p_rep1</td>
<td>qweight_pu_rep1</td>
</tr>
<tr>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td></td>
<td>qweight_p_rep50</td>
<td>qweight_pu_rep50</td>
</tr>
<tr>
<td>3-Month</td>
<td>q1weight_p_rep1</td>
<td>q1weight_pu_rep1</td>
</tr>
<tr>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td></td>
<td>q1weight_p_rep50</td>
<td>q1weight_pu_rep50</td>
</tr>
</tbody>
</table>

… and so on for each wave

Accessing Poverty Tracker Data & Other Resources
Public Data

Much of the data we’ve collected over the years is publicly available on the CPSP website.

As of today, you can download the full two years of data from our first cohort (2012 - 2014) and the first four years of data from our second cohort (2015 - 2019). Data collection with the second cohort is ongoing.

Similar resources are available for both cohorts:
- **Data files** in both Stata 14 and CSV format. (Stata file includes labels.)
- Complete copies of all **survey questionnaires** administered to respondents.
- **Codebooks** for all survey waves that list the variables, full questions, answer options, and labels.

---

Accessing Public Data: The Process

**STEP #1**


New Users
Register for public-use data access.

**STEP #2**
Accessing Public Data: The Process

**STEP #3**

![Registration Form](image)

**STEP #4**

![Password](image)

**STEP #5**

![Poverty Tracker Data Portal](image)

Panel 1 (2012-2014)
- Codebooks
- Stata 14 DTA
- CSV
- Surveys
- All Files Zip

Panel 2 (2015-2017)*
- Codebooks
- Stata 14 DTA
- CSV
- Surveys
- All Files Zip

Last updated: May 24th, 2021

*Panel 2 was followed through 2019. The files available publicly include survey waves through 2017.
Accessing Public Data: The Process

**STEP #6**

Returning Users

Log in to access the public-use data.

---

**Restricted-Use Data**

- We collect more data than we publish online. Some variables are deemed sensitive and only made available upon request.

- **Sensitive variables include:**
  - Non-top-coded income measures
  - Geographic indicators (e.g., zip code and census tract)
  - Open-ended text responses (e.g., “What is the reason you did not seek help with your health issue?”)
  - Information on sexuality, immigration history, salary and wages
Accessing Restricted-Use Data: The Process

1. **Let us know!** Put your request in the public data registration form or shoot us an email at povertytracker@columbia.edu. Include a brief overview of your research project and variables-of-interest.

2. **Set up a meeting** with a member of the data team to discuss details, timeline, and next steps.

3. **Obtain approval** from Poverty Tracker project directors.

4. **Submit a Data Usage Agreement** (DUA) packet.

---

So what’s a “DUA”?

- Extended Research Abstract
- Application for Obtaining Restricted Use Data (Cover Page)
- Restricted Use Data License Application
- Supplemental Agreement with Research Staff
- Restricted Use Data Protection Plan
- Copies of CVs for all Research Staff
- Proof of IRB Approval from Researcher’s Institution (Proposal + Data Protection Plan)

*Don’t worry!! We’re here to help!*
New Public Resources

• 2-Page Overview of the Poverty Tracker
• Comprehensive Guide to the Poverty Tracker
• Supplemental Guide to Using the Weights
• Technical Appendix on Imputation and Weighting
• Searchable Question Library Spreadsheet
  - Includes all questions/variables in the public datasets
  - Organized by survey wave and topic
Poverty Tracker Data Demonstration

- Downloading the Poverty Tracker Data
- Pointers for using the Poverty Tracker data
- Producing summary statistics across survey waves

Example analysis:

What shocks are correlated with entrances and exits from poverty and hardship?

---

Downloading the Poverty Tracker Data

→ Go to the website: povertycenter.columbia.edu/poverty-tracker-data
Pointers for using the Poverty Tracker data

Variable Naming Convention: Prefixes

Prefix indicating the survey on which the data stored in the variable was collected. Begin with a “q” for questionnaire, followed by a number corresponding to a particular survey.

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

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.

Pointers for using the Poverty Tracker data

Variable Naming Convention: Suffixes

The variables also have a suffix that indicates one of three things:

1. if the variable matches a survey question and has not been imputed;
2. if the variable matches a survey question but was imputed; or
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 is a top-coded variable because you will see the suffix “_tc” at the end of the variable name.
Pointers for using the Poverty Tracker data

**Variable Suffix Example 1:** If the variable matches a survey question and *has not been* imputed

The 3-month and 21-month quarterly surveys ask respondents, *How would you rate your fatigue on average?* On both surveys, this is question B9 (the 9th question in Module B).

In the public dataset, the variables representing these questions are:

- *q1b9* for the 3-month survey
- *q7b9* for the 21-month survey

The prefix is *q1* for the responses from the 3-month survey and *q7* for those from the 21-month survey.

Because this variable matches a survey question, the suffix (*"b9"*) indicates the question number.

*Warning:* Suffix does not always match question number - please consult codebooks and questionnaires!

---

Pointers for using the Poverty Tracker data

**Variable Suffix Example 2:** if the variable matches a survey question but *was* imputed

Imputes variables concerning income (all forms), hardship, health, assets, debts, or demographics.

Imputed variable names are not associated with their survey question number. They are descriptive.

In the public dataset, the variables representing disability income are:

- *imp_qincdis_tc* for the baseline annual survey
- *imp_q4incdis_tc* for the 12-month annual survey
- *imp_q8incdis_tc* for the 24-month annual survey … and so on for each annual survey

You can see the complete list of imputed variables in Poverty Tracker Technical Notes.
Pointers for using the Poverty Tracker data

**Variable Suffix Example 3:** if the variable does not match a survey question, but was constructed based on responses to survey questions.

The Poverty Tracker data includes several variables *constructed* using responses to the individual survey questions (e.g. poverty status)

All of the constructed variables also have descriptive names.

For example the variables representing poverty status (measured using the SPM) are:

- \textit{qspmpov} for the responses from the baseline survey
- \textit{q4spmpov} for the responses from the 12-month survey
- \textit{q8spmpov} for the responses from the 24-month survey... and so on for each annual survey

Questions about this research?

**Contact:**

General Inquiries, povertytracker@columbia.edu

Yajun Jia, Survey Methodologist, yj2648@columbia.edu

Schuyler Ross, Data Manager, sr3820@columbia.edu

Sophie Collyer, Research Director, smc2246@columbia.edu

Acknowledgments

We thank Robin Hood for funding the Poverty Tracker, which provided data for this study. We are grateful to all members of the Poverty Tracker team based at Columbia, past and present!