Appendix I: Accuracy of the Data and Changes to the NYCgov Model
Appendix I

Accuracy of the Data and Changes to the NYCgov Model

The principal data set for NYC Opportunity poverty estimates is the American Community Survey (ACS) Public Use Micro Sample (PUMS). The ACS is designed to sample three percent of U.S. households each year and the PUMS is a subset of the full ACS sample. It provides information collected from roughly 26,000 households in New York City annually. Because the ACS is a survey, it is subject to two types of error: nonsampling error and sampling error.

Nonsampling error is the error within survey data that is not specifically associated with the statistical sampling procedures of the sample data. Nonsampling error may occur because of erroneous responses by survey respondents, for example. Another source of nonsampling error may come from mistakes in the processing of the data by the Census Bureau, such as when data are edited or recoded.

Nonsampling error may affect the data in two ways: either randomly, which increases the variability of the data, or systematically, which introduces bias into the results. To minimize bias in the survey, the Census Bureau conducts extensive research of sampling techniques, questionnaire design, and data collection and processing procedures. For instance, after identifying a systematic underreporting of Supplemental Nutritional Assistance Program (SNAP) receipt and benefit dollar values in the ACS, the Census Bureau researched methods to increase the reported participation rate. The Census Bureau concluded through this research that changing the SNAP question to include the words “SNAP benefit card,” as well as not asking about the SNAP benefit value, would significantly increase the number of households responding to whether they have received SNAP.¹

Sampling Error: Sampling error occurs in the ACS, as in other sample survey data, because inferences about the entire population, such as the poverty rate for New York City, are derived from a sample of individuals and housing units. Another sample drawn from the same population would provide a different estimate of the poverty rate. The sampling error is estimated by the standard error, which can be thought of as a measure of the deviation of an estimate drawn from one sample from the average estimate of all possible samples.

For this report, NYC Opportunity employed the replicate weight method recommended by the Census Bureau to compute direct standard errors for our estimated poverty rates. The standard errors provide a measure of sampling error and some types of nonsampling error. Using the standard errors, we tested the statistical significance of differences and changes in the report’s poverty rates at the 10 percent level of significance. In the report’s tables, we highlight, in bold, statistically significant differences between poverty rates.

An additional source of error in the data results from NYC Opportunity’s need to impute information on items such as the value of SNAP benefits, housing status, childcare expenditures, and medical out-of-pocket expenditures from other survey data into the ACS sample. We do not, however, account for the imputation error in this report.

Changes to the Poverty Model in This Report

The statistics shown here are slightly different from what is reported in previous issues of our poverty reports due to adjustments to our methodology. The changes are described in the appendices on the Unit of Analysis (addition of health insurance units), Work Related Expenses (change in imputation of childcare costs), Medical Out-of-Pocket Expenditures (change in imputation of medical costs), and Housing Status (change in imputation of housing status and HVS match for 2013). These changes were applied to our 2015 estimate and also to prior years, when appropriate.