**Greater Boston Jewish Community Study, 2015**

**Cohen Center for Modern Jewish Studies, Brandeis University**

**Documentation of Public Use Data Set**

*For more information contact Janet Krasner Aronson,* [*jaronson@brandeis.edu*](mailto:jaronson@brandeis.edu)

*Version January 9, 2017*

**Introduction**

This document describes the Public Use Data Set for the 2015 Greater Boston Jewish Community Study. It explains the constructed variables in the dataset and the procedures for statistical weighting.

Variables that begin with the prefix “x\_” were constructed from original survey responses, some of which are not included in the public data set (this is noted in the descriptions where applicable). All other variables are documented in the codebook (Appendix D of the report) and are taken directly from the survey.

You can find the report and technical appendices at: <http://www.brandeis.edu/ssri/communitystudies/bostonreport.html>

**About the Public Use Data Set**

The Public Use Data Set contains all raw data provided by respondents except for data that might be used to identify individual respondents. Primarily, these removed variables were open-ended responses and household ZIP codes.

Descriptions of some variable naming conventions follow:

1. All variables with a “resp” in the name refer to the respondent (e.g., respage is the age of the respondent and x\_isrresp is the constructed variable noting if the respondent is Israeli).
2. Variables with a “hhad” prefix refer to non-respondent adults in the household; these variables range from 2-9 because the respondent is considered as the first adult (e.g., hhadage3 is the age of the third adult in the household).
3. Variables with a “hhch” prefix refer to the children in the household; these range from 1-9 (e.g., hhchgrd1 is the grade-level of the first child in the household).
4. Variables with an “ac” prefix refer to adult children; these are children of the respondent who are 18 years of age or older (e.g., acempnow2 is the employment status of the second adult in the household, if that adult is the child of the respondent).

*Potential Issues with Data Interpretation*

CMJS as a rule maintained the integrity of the data as collected. As such, two potential issues warrant caution. First, due to programming conventions, both refusals and questions skipped through survey logic are coded the same way, as system missing. Second, there are cases where respondents answered a question and its follow-up before seemingly backtracking and changing the original response, which would otherwise render the follow-up response invalid. Such overwritten responses remain in the data set.

The variables charany and charanyamt provide examples of both types of potential issues. Regarding system missing, those for whom charnay=0 were not asked charanyamt. Regarding probable backtracking, there are three cases with responses for charanyamt for whom charany=0.

It is up to the analyst’s own interpretation as to how to account for these characteristics of the data set.

**Definitions of Constructed Variables (x\_ prefix)**

Variables were constructed from raw data for two purposes:

1. Standardized coding of open-ended data
2. Analytical variables created through complex combinations of multiple raw variables. The syntax used to create complex variables will be included in the report appendix.

x\_region: Respondents supplied the ZIP code of their primary residence in the Greater Boston area. To protect their identities, this variable groups ZIP codes into 11 regions that make up the Greater Boston area (detailed explanations of these regions can be found in Chapter 2 of the report). The regional categories are adapted from those used in the 2005 study.

x\_newregion: These are new regional categories as defined by Combined Jewish Philanthropies after the publication of the report. The detailed explanations of these regions appear at the end of this document.

x\_hhchrelrsd1-9: In the original survey, after identifying the religion in which the first child is being raised, respondents were asked if all children are being raised in the same religion (see variable hhchrelsame). These variables fill in the responses for children 2-10 who have the same religion as child 1.

x\_pkc\*; x\_jec\*: Separate, parallel questions were asked of households with one child and households with multiple children. Variables with the prefix x\_pkc and x\_jec contain the combined single- and multiple-child household data regarding children in pre-kindergarten and K-12, respectively. Note that students in K-12 include 18 and 19 year olds who are currently enrolled in high school (see variables resphs and hhadhs2-9). Note that one method of calculating household-level participation in types of Jewish education requires accounting for eligible children for whom x\_pkc and x\_jec data are missing.

*Due to a programming error on the instrument, single-child households whose children were in K-8 were not asked if the child participated in a youth group, and so are coded as missing from x\_jecygnow and x\_jecygk8.*

*Due to a programming error on the instrument, households with all children enrolled in part-time school were not asked if the children were enrolled in non-Jewish private schools, and so are coded as missing from x\_jecprivnow, x\_jecprivk8, and x\_jecpriv912.*

x\_rlsyntype1-5: A recoding of the up-to-five synagogues respondents say they belong to (the names were removed from the data set to protect identities). Traditional, brick-and-mortar synagogues are separated by denomination, while independent *minyanim*, *havurot*, and Chabad are listed separately, regardless of denomination.

x\_hhsize: This variable combines the number of adults (hhadct) and children (hhchct) in the household.

x\_respjewtype, x\_hhadjewtype2-9: These variables denote the “type of Jew” corresponding to the respondent or the household adults. Jews by Religion (JBR) are those who say their religion is Jewish and have a Jewish background (i.e., a Jewish parent, was raised Jewish, or converted). Jews of No Religion (JNR) are atheists and agnostics who consider themselves Jewish aside from religion and have a Jewish background, or say they are Jewish and atheist/agnostic, and have a Jewish background. Jews of Multiple Religions (JMR) either say they have two religions, one of which is Judaism, and have a Jewish background, or have another religion but consider themselves Jewish aside from religion, and have a Jewish background. Jewish Background (JB) say they have a religion other than Judaism, do not consider themselves Jewish aside from religion, but have a Jewish background. Jewish Affinity (JA) consider themselves Jewish either by religion or not by religion, but do not have a Jewish background. Non-Jews (NJ) have a religion other than Judaism, do not consider themselves Jewish in any way, and do not have a Jewish background. Messianic Jews with a Jewish background were categorized as JB, and those without a Jewish background were categorized as NJ.

x\_hhchjew1-9: Children are counted as Jews if they are being raised Jewish by religion, culturally Jewish, or Jewish and another religion (corresponding to the variables x\_hhchrelrsd1-9).

x\_spjewtype: This is the “type of Jew” of the respondent’s live-in spouse, fiancé/e, or significant other.

x\_spjewish: This notes if the respondent’s live-in spouse, fiancé/e, or significant other is Jewish (i.e., JBR, JNR, or JMR).

x\_martypehh: This notes if the household contains an inmarried or intermarried couple, or no couple (whether or not the respondent is part of the married couple).

x\_martyperesp: This notes if the respondent is inmarried, intermarried, or unmarried.

x\_jengage: This is the pattern of engagement of the respondent (see Chapter 3 of the report).

x\_isrhh: This notes if there is an Israeli in the household. Israelis were either born or raised in Israel, consider themselves Israeli in any way, or have Israeli citizenship.

x\_isrresp: This notes if the respondent is Israeli. Israelis were either born or raised in Israel, consider themselves Israeli in any way, or have Israeli citizenship.

x\_rushh: This notes if there is a Russian speaker in the household. Russian speakers were either born or raised in Russia/the Former Soviet Union, were raised in a Russian-speaking household, or currently speak Russian in their household.

x\_rusresp: This notes if the respondent is a Russian speaker. Russian speakers were either born or raised in Russia/the Former Soviet Union, were raised in a Russian-speaking household, or currently speak Russian in their household.

x\_glbtqhh: This notes if there is a GLBTQ individual in the household. In addition to a survey question about the presence of gay, lesbian, or bisexual individuals in the household, open-ended responses that indicated identification as neither male nor female were used to place people into this category.

x\_glbtqresp: This notes if the respondent is GLBTQ. In addition to a survey question about whether the respondent was gay, lesbian, or bisexual, open-ended responses that indicated an identification as neither male nor female were used to place people into this category.

**Constructed variables for analysis**

x\_hhjewish: Jewish households contain at least one JBR, JNR, or JMR adult. There are 47 households that screened into the survey but do not contain any Jewish adults or children. These households include at least one person of Jewish Background or Jewish Affinity.

x\_respjewish: Jewish respondents are JBR, JNR, or JMR.

**Weighting**

Four weight variables are available for this dataset. Two are at the household level (x\_wthh and x\_wthhf), and two are at the respondent level (x\_wtresp and x\_wtrespf). Household-level weights should be used to calculate characteristics of the household, population counts, and anything involving children. Respondent-level weights should be used to calculate characteristics of respondents (e.g., behaviors and attitudes).

The weight variables are also segmented by whether they refer to the primary sample (x\_wthh and x\_wtresp) or the full sample (x\_wthhf and x\_wtrespf). Primary weights are the only ones appropriate for generating counts or characteristics of the overall population. The full-sample weights are appropriate only for generating characteristics of subpopulations.

Note: The public-use dataset includes all screener data, but households that screened out of the survey have all their weights set at 0. Weighting instructions are designed for use with Stata.

**Constructed variables for weighting**

x\_stratagrp: The strata identified is used for weighting.

x\_wthh: The primary-sample household weight.

x\_wthhf: The full-sample household weight.

x\_wtresp: The primary-sample respondent/individual weight.

x\_wtrespf: The full-sample respondent/individual weight.

Primary Weights

*Use ONLY the primary sample for generating population counts*

For household estimates—estimations on the number of households—use wthh to estimate the percentage of households. For estimations on the number of people—i.e., counts—use the household weights with totals of count variables—e.g., hhadct, hhchct.

svyset \_n [pweight= x\_wthh], strata(x\_stratagrp) vce(linearized) singleunit(missing)

Use the respondent weights for questions asked only of respondents. Use wtresp to estimate percentage of adults (including respondent and non-respondent adults):

svyset \_n [pweight=x\_wtresp], strata(x\_stratagrp) vce(linearized) singleunit(missing)

Primary weights should be used for generating characteristics of the population as a whole, including counts and subgroups.

*Example: Household Characteristics*

The number or proportion of households that are synagogue members:

svyset \_n [pweight= x\_wthh], strata(x\_stratagrp) vce(linearized) singleunit(missing)

svy, subpop(x\_hhjewish): tab rlsynany, count

svy, subpop(x\_hhjewish): tab rlsynany

*Example: Counts of People*

svyset \_n [pweight= x\_wthh], strata(x\_stratagrp) vce(linearized) singleunit(missing)

svy, subpop(x\_hhjewish): total hhadct

*Example*: *Respondent or Individual Characteristics*

The proportion of close Jewish friends:

svyset \_n [pweight= x\_wtresp], strata(x\_stratagrp) vce(linearized) singleunit(missing)

svy, subpop(x\_respjewish): tab jlfriend

Full Weights

*DO NOT use the full weights to generate population counts.* Full weights should be used for characteristics of subgroups (e.g., percentage of synagogue-member households that belong to other Jewish organizations)

svyset \_n [pweight= x\_wthhf], strata(x\_stratagrp) vce(linearized) singleunit(missing)

svyset \_n [pweight= x\_wtrespf], strata(x\_stratagrp) vce(linearized) singleunit(missing)

*Example: Household Subpopulations*

The proportion of synagogue-member households that also belong to Jewish organizations:

svyset \_n [pweight= x\_wthhf], strata(x\_stratagrp) vce(linearized) singleunit(missing)

svy, subpop(x\_hhjewish): tab actmem rlsynany, col

*Example: Respondent or Individual Characteristics*

The proportion of close Jewish friends held by those born locally:

svyset \_n [pweight= x\_wtrespf], strata(x\_stratagrp) vce(linearized) singleunit(missing)

svy, subpop(x\_respjewish): tab jlfriend locborn, col

**Combined Jewish Philanthropies’ New Regional Definitions (x\_newregion)**

Downtown Boston: Back Bay, Bay Village, Beacon Hill, Chelsea, Chinatown, Downtown, East Boston, Everett, Fenway, Leather District, North End, Revere, South Boston, South Boston Waterfront, South End, West End, Winthrop

Brookline, Brighton, Allston, and Surrounding: Allston, Brighton, Brookline, Longwood, Mission Hill

Cambridge, Somerville, Medford, Charlestown: Cambridge, Charlestown, Medford, Somerville

Jamaica Plain, W. Roxbury, Roslindale, and Surrounding: Dorchester, Hyde Park, Jamaica Plain, Mattapan, Roslindale, Roxbury, West Roxbury

Newton, Needham, Wellesley, Weston: Needham, Newton, Wellesley, Weston

Lexington, Waltham, Watertown, Arlington, Belmont: Arlington, Belmont, Lexington, Waltham, Watertown,

South Area: Abington, Avon, Braintree, Bridgewater, Brockton, Canton, Cohasset, Dedham, East Bridgewater, Easton, Foxborough, Halifax, Hanover, Hanson, Harbor Islands, Hingham, Holbrook, Hull, Kingston, Mansfield, Marshfield, Milton, Norwell, Norwood, Pembroke, Quincy, Randolph, Rockland, Scituate, Sharon, Stoughton, Walpole, West Bridgewater, Westwood, Weymouth, Whitman, Wrentham

Metrowest: Ashland, Bellingham, Dover, Framingham, Franklin, Holliston, Hopkinton, Hudson, Marlborough, Maynard, Medfield, Medway, Milford, Millis, Natick, Norfolk, Sherborn, Southborough, Stow, Sudbury, Wayland

North and Northwest Suburbs: Acton, Bedford, Boxborough, Burlington, Carlisle, Concord, Lincoln, Malden, Melrose, North Reading, Reading, Stoneham, Wakefield, Wilmington, Winchester, Woburn, Sherborn, Southborough, Stow, Sudbury, Wayland

North Shore: Beverly, Boxford, Danvers, Essex, Georgetown, Gloucester, Hamilton, Ipswich, Lynn, Lynnfield, Manchester, Marblehead, Middleton, Nahant, Newbury, Newburyport, Peabody, Rockport, Rowley, Salem, Saugus, Swampscott, Topsfield, Wenham