# Weight Construction and Usage

# in Wave One of the Three-City Study

Andrew Cherlin
Paula Fomby
Robert Moffitt

Johns Hopkins University

March 2002

This document has two purposes. First, it briefly reviews the construction of the weights for Wave One of the Three-City Study. Second, it summarizes the decisions of the Co-Principal Investigators in three areas of weight usage: (1) weight variance and trimming; (2) weight normalization; and (3) adjustment of variances of estimated quantities for the complex survey design.

**Construction of the Weights**

A full description of the construction of the weights can be found in Jones (2000). The sampling design and its scientific rationale can be found in Winston et al. (1999). Here we provide a relatively nontechnical summary.

The unit of observation in the Three-City Study is a child age 0-4 or 10-14 years. Each unit has a constructed weight. Each is a product of four separate weights:

$$W = (W1) \times (W2) \times (W3) \times (W4)$$

where

W1 = inverse of the probability that a dwelling unit (DU) was selected for screening

W2 = inverse of the probability that a screened eligible household was selected for the sample and offered participation in the study

W3 = inverse of the probability of an affirmative response

W4 = inverse of the probability that a child in the appropriate age range was selected for our study

The weight W1 is based on the probability that a DU was selected for screening. The firm conducting the survey, the Research Triangle Institute (RTI), constructed eight sets of block groups from all the blocks in the city in the 1990 Census. Each set ranked all the block groups in the city in descending order of the poverty rate of children in a particular race-ethnic group. In Boston, three such sets were compiled and ranked --one each for Non-Hispanic Whites, Non-Hispanic Blacks, and Hispanics. Three analogous sets were compiled for Chicago, and two sets for San Antonio--Non-Hispanic Blacks and Hispanics. Low-income non-Hispanic Whites were excluded from the San Antonio sample because of their relatively high geographic dispersion in that city.

Each block group in each set was called a primary unit (PU). In each of the eight sets, a minimum percentage poor was selected as a cutoff point, and PUs with percentages poor less than the cutoff were excluded from further consideration (the cutoffs were chosen to balance the desire to cover as great a percentage of the city's poor child population as possible without increasing survey costs inordinately). In Boston, the cutoff was a poverty rate of 15.3 percent in the PU; in Chicago, 8.2 percent; and in San Antonio, 17.1 percent (Winston et al. 1999, table A-2). A random set of the remaining PUs was selected with probability proportional to size. Then the selected PUs were divided into segments--which are areas of a size typically regarded as convenient for surveying and generally consist of 90-120 dwelling units--and a set of segments was chosen randomly from the PUs. All selected segments were then counted and listed (i.e., interviewers visited the segments, counted the housing units, and wrote down the addresses of all occupied dwelling units). A random sample of DUs (identified by street addresses), was then selected from within each segment. As noted earlier, W1 is the inverse of the probability that a DU was selected for screening, and that probability was calculated as the product of the probability that the PU in which the DU is located was selected from the set of block groups with a poverty rate above the cutoff, the probability of being in a segment selected within a PU, and the probability of being selected within a segment.

The weight W2 is based on the differential selection rates of families arising from the scientific goals of the study, which were to focus primarily on children living in single-mother families with income less than the federal poverty line but to include some married couple families and some families (both single-mother and married) with incomes greater than the poverty line. A screening interviewer visited each sampled DU to determine whether four eligibility conditions were met: whether there was an eligible child in the household (a child aged 0-4 or 10-14), whether the household had income less than twice the poverty level for its size, whether the primary caregiver of the child was female, and whether the household head was non-Hispanic White, non-Hispanic Black, or Hispanic. If these conditions were met, the household's status on four specific yes/no variables was determined: (1) whether household income was less than the poverty line or between 100 percent and 200 percent of the poverty line; (2) into which of the three race-ethnic groups the household head self-identified; (3) whether the household head was married or unmarried; and (4) if the household received both Food Stamps and Medicaid or if it received only one or neither. There are twenty-four possible combinations resulting from combining these four variables. For each of the 24 cells (e.g., children living in a household with income below 100 percent of the poverty line, Hispanic, unmarried, and receiving both Food Stamps and Medicaid), a different sampling rate was set. These 24 sampling rates were also set separately for each of the eight sets of race-ethnic-specific block group lists mentioned previously. The rates were design to result in a sample consisting of equal numbers of families of each race-ethnic group, and consisting primarily of single-mother families with incomes below the poverty line. The targets also aimed to achieve a sample which was half on TANF and half not on TANF.[1] Combing the 24 cells in each of the 8 block group lists, the design includes 176 cells , each with a different sampling probability.[2] W2 is the inverse of the sampling probability for each family that was selected for the study and offered participation in it.

As for W3, a response probability was calculated as the product of the probability that a selected DU was screened (i.e., that the screener spoke to someone in the household and that person complied with the screener) multiplied by the probability that, if offered participation in the study, the DU accepted participation. The product of these probabilities was, on average, 74 percent. This figure represents the response rate for the survey component of the study. The first probability was disaggregated by which of the eight race-ethnicity-city sets of block groups it came from; and the second probability was disaggregated by several of the characteristics

---

[1] These targets were obtained by using estimates of the fraction of families on Medicaid and Food Stamps, or on neither, who were on TANF. Using Medicaid and Food Stamp receipt for the design variable rather than TANF receipt directly was based on the presumption that asking TANF receipt at the front door might result in underreporting.

[2] If there are 24 cells and 8 lists, there would be 192 individual cells. However, 16 of these cells were for San Antonio white families, who were not sample. Hence the number of cells was 176.

gathered in the screening interview (e.g., married or unmarried). The inverse of the response probability for each nonresponse class yielded W3.

If more than one child in either the 0-4 or 10-14-year age range was in the household, one was chosen at random to be the focal child for the study. Therefore, the weight W4 was constructed as the inverse of the number of age-eligible children in the family, which is the probability that the focal child was selected. The product of W1 * W2 * W3 * W4, which we will call the child weight, is the appropriate weight to use for analysis. Since this sample was conceived as a sample of children age 0 to 4 or 10 to 14, we include the child weight on the public use file. It should be emphasized that the weights W are focal child weights, not adult caregiver weights.

**Weight Usage Decisions by the Co-PIs**

Weight Variance and Trimming. The weights in the Three-City Study have a high variance, as shown in Table 1 which reports the mean, standard deviation, and percentile points of W1, W2, the components of W3, and their product. W1 has significant variation, ranging from 0.8 to 94.3, with a standard deviation of 17.7 relative to its 8.9 mean. W2 ranges from 1.0 to 20, again a wide range. Values of W2 are derived from the inverses of the 176 design cell probabilities, which range from a low of .053 (e.g., for Hispanic married mothers in San Antonio with income below poverty who are not receiving Medicaid and Food Stamps) to 1.0 (e.g., for Black single mothers in Chicago with income below poverty receiving Medicaid and Food Stamps). On average, married mothers, those with income between 100 percent and 200 percent of poverty, and those not receiving Medicaid and Food Stamps have low selection probabilities and high weights; and single mothers with income below poverty receiving Medicaid and Food Stamps have high selection probabilities and low weights. The aim of the design was a sample consisting primarily of TANF-eligible families but with small samples of comparison groups (e.g., married heads of household, higher income households, and single mothers not on welfare) included. The large range of selection probabilities reflects this desire, and also reflects the desire to have equal numbers of each of the three race-ethnicity groups in the sample; to achieve this, is was necessarily to undersample such groups that were numerous in a city (e.g., Hispanics in San Antonio). This further increases the variation of the weights. The nonresponse weights in Table 1 that constitute W3 vary very little from dwelling unit to dwelling unit, indicating that response rates for the screening and main interviews did not vary much among the design cells.

The final column shows the "DU" weight, the product of the three weights.[3] The DU weight shown in the table, however, has been transformed in two ways. First, RTI trimmed (i.e., slightly reduced) a few of the high weights to reduce variance, in order to keep design effects for the DU weight below 2.0 and the design effect for the child weight below 3.0.[4] Second, the DU weight has been normalized to one (i.e., the DU weights of all the cases in the sample average 1.0). As the table shows, the dispersion of this weight is quite large, ranging

---

[3] The weight W1 was not called the "DU weight" by RTI. All three weights--W1, W2, and W3--are, in a sense, "DU" weights since they all involve probabilities of selecting a DU at different stages of the sampling process. But the term "DU weight" is reserved for the final probability of selecting a DU, which is the product of the three weights.

[4] The design effect is calculated as the ratio of the sum of the squared weights and the square of the sum of the weights times sample size. It is a measure of the variance of normalized weights, and a design effect over 2 implies approximately that the variance of the weights is more than twice its mean.

from a low of .03 to a high of 13.42. There are several weights in the "13" range, all of which are for married Hispanics in San Antonio.

The co-PIs considered additional trimming of the weights in order to reduce the weight variation in the sample. A decision was made to consider trimming only W1 and W3, and not W2, since trimming on W2 would bias estimates of population proportions by race-ethnicity, income, welfare receipt, and family structure, and these distributions were too central to the study to afford compromise. However, investigation revealed that variation in W2 accounts for 90% of the total variance of the weights.[5] The "within design cell" residual weights have a range of only 3, which means that the weights could not be reduced by more than 3 even by complete trimming on the non-W2 weights. A 10% trim on these residual weights was conducted and had essentially no effect on the variance of the weights and did not eliminate any of the very highest weights in the data. As a consequence, the project decision was to leave the weights as is.

<u>Further normalization for city size</u>. There are two versions of the child weight, normalized to 1.0, on the Three-City public use file. The "raw" child weight has no further transformations. It weights children in proportion to the population size of the city in which they reside (technically speaking, in proportion to the sampled population in each city, which is only a subset of the total city population but also varies greatly across the three cities). In other words, it weights a child from Chicago more heavily than a comparable child from San Antonio. However, the co-PIs decided that a weight based on a population- or sample-weighted average of the three cities was of little interest either scientifically or for policy makers. A second

[5] The variance of the weights is 2.18, the square of 1.48 shown in Table 1; a regression of the weights on the 176 dummies for the W2 categories leaves residuals with a variance of only .209. So the R-squared is [(2.18-.209)/2.18], approximately .90.

consideration was that welfare policies differ across cities and not within them, and there are effectively only three observations on welfare policies, which should be weighted equally.

Consequently, the co-PIs constructed a weight which was further normalized in such a way as to give the observations in each of the three cities equal weight (i.e., one-third each, while preserving the property that the average weight across all cases is 1.0). This "city-normalized" child weight, as we will call it, eliminated two sources of variation in the raw child weights. First, it eliminated variation arising from the different population sizes of the three cities. Second, it eliminated variation arising from the different unweighted numbers of observations in the three cities: 926 in Boston, 762 in Chicago, and 714 in San Antonio. All analyses by the co-PIs use this city-normalized child weight. Note that any analyses which are stratified by city are unaffected by whether the analyst uses the raw or city-normalized child weights It should also be noted that analyses which use subsamples of the total sample must renormalize the weights again if this equal-city- weighting principle is to be maintained.[6] The public use file contains both the raw child weight and the city-normalized child weight.

Variance Adjustment. The co-PIs investigated the need to adjust the variances of statistics using the Three-City data to take account of the complex survey design. Illustrative ordinary least squares regressions were estimated using the software program STATA which incorporated design elements by defining strata and clusters appropriate to the design. A complicating factor is that, in three of the eight race-ethnic-city sets of PUs, all PUs with poverty levels above the cutoff were chosen, rather than a sample of all such PUs. Upon the advice of

[6] A SAS macro which renormalizes the weights for any subsample is available upon request from the co-PIs.

RTI, strata were defined by first constructing a geographic ordering of the PUs in the five sampled lists and a geographic ordering of the DUs in the three 100-percent lists, by grouping adjacent pairs of PUs (in the first case) and DUs (in the second case) into stratum, and by defining the PUs or DUs within each stratum as clusters.

The Appendix shows STATA output from four regressions that illustrate the results. The first two have "minor11" as the dependent variable, which is the number of minors in the household. The second two have "ineed11" as the dependent variable, which is the income of the household relative to its poverty line (i.e., the income to needs ratio). The independent variables are a dummy for White and two dummies for Boston and Chicago, with San Antonio the omitted group. In each pair of regressions, the first defines strata and cluster as described above, whereas the second ignores strata and cluster entirely, for comparison purposes, i.e., it ignores the complex nature of the design.

The results show that the standard errors are affected only slightly by the adjustment. Given these results, the project decision was to ignore any standard error adjustment in future analyses. Of the three variables needed to do this adjustment, only two (SEGID and SCRID) are on the public use file. The other one (PU) is available upon request from the co-PIs.

## References

Jones, B.L. 2000. "Welfare, Children, and Families: Sample Selection and Weight Development." Research Triangle Institute.

Winston, Pamela, Ronald J. Angel, Linda M. Burton, P. Lindsay Chase-Lansdale, Andrew J. Cherlin, Robert A. Moffitt, and William Julius Wilson 1999. *Welfare, Children, and Families: A Three-City Study, Overview and Design.* Johns Hopkins University. Available upon request or at www.jhu.edu/~welfare.

Table 1

Weight Variation in the Three-City Study

| | W1 | W2 | Nonresponse Weight (W3) | | DU Weight[a] |
|---|---|---|---|---|---|
| | | | Screening | Interview | |
| Mean | 8.9 | 2.2 | 1.1 | 1.2 | 1.00 |
| Standard Deviation | 17.7 | 3.2 | 0.1 | 0.2 | 1.48 |
| Percentile Points | | | | | |
| 100 | 94.3 | 20.0 | 1.1 | 2.2 | 13.42 |
| 90 | 16.1 | 3.9 | 1.1 | 1.3 | 2.53 |
| 75 | 3.5 | 2.1 | 1.1 | 1.3 | 1.29 |
| 50 | 2.4 | 1.0 | 1.1 | 1.2 | 0.41 |
| 25 | 2.2 | 1.0 | 1.0 | 1.1 | 0.17 |
| 10 | 1.2 | 1.0 | 1.0 | 0.9 | 0.12 |
| 0 | 0.8 | 1.0 | 1.0 | 0.5 | 0.03 |

Notes:

[a] This weight is the product of W1, W2, and W3, and is called the “DU weight.” It differs from the child weight described in the text by the omission of W4. It is trimmed and normalized to one.

# Appendix

```
svyreg minor11 white boston chicago [pweight=normchwt]

Survey linear regression

pweight:  normchwt                                  Number of obs    =      2401
Strata:     newst                                         Number of strata =
 728
PSU:       newcl                                         Number of PSUs   =
1456
                                                    Population size  = 2401.1361
                                                    F(   3,    726)  =     13.31
                                                    Prob > F         =    0.0000
                                                    R-squared        =    0.0467

------------------------------------------------------------------------------
 minor11 |      Coef.    Std. Err.       t     P>|t|      [95% Conf. Interval]
---------+--------------------------------------------------------------------
   white |  -.3225021    .1917331     -1.682   0.093     -.6989178    .0539136
  boston |   .5402056    .1383595      3.904   0.000      .2685745    .8118368
 chicago |  -.2271986    .1417695     -1.603   0.109     -.5055244    .0511271
   _cons |   2.993269    .0982369     30.470   0.000      2.800407     3.18613
------------------------------------------------------------------------------

. svyset, clear

. svyreg minor11 white boston chicago [pweight=normchwt]

Survey linear regression

pweight:  normchwt                                  Number of obs    =      2401
Strata:   <one>                                     Number of strata =         1
PSU:      <observations>                            Number of PSUs   =      2401
                                                    Population size  = 2401.1361
                                                    F(   3,   2398)  =      9.68
                                                    Prob > F         =    0.0000
                                                    R-squared        =    0.0467
------------------------------------------------------------------------------
 minor11 |      Coef.    Std. Err.       t     P>|t|      [95% Conf. Interval]
---------+--------------------------------------------------------------------
   white |  -.3225021    .1904366     -1.693   0.090     -.6959393     .050935
  boston |   .5402056    .1779697      3.035   0.002      .1912154    .8891958
 chicago |  -.2271986    .1410533     -1.611   0.107     -.5037975    .0494003
   _cons |   2.993269    .1174488     25.486   0.000      2.762957     3.22358
------------------------------------------------------------------------------


.
svyreg ineed11 white boston chicago [pweight=normchwt]

Survey linear regression

pweight:  normchwt                                  Number of obs    =      2399
Strata:   snewst                                    Number of strata =       726
PSU:      snewcl                                    Number of PSUs   =      1452
```

```
                                                  Population size  = 2394.4117
                                                  F(   3,    724)  =      2.25
                                                  Prob > F         =    0.0810
                                                  R-squared        =    0.0058

------------------------------------------------------------------------------
 ineed11 |      Coef.   Std. Err.       t     P>|t|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
   white |   .0333805     .074934      0.445   0.656     -.1137327    .1804938
  boston |  -.0582959    .0603108     -0.967   0.334     -.1767002    .0601084
 chicago |   .0405888    .0586917      0.692   0.489     -.0746368    .1558145
   _cons |   .7336379    .0524946     13.975   0.000      .6305786    .8366972
------------------------------------------------------------------------------

. svyset, clear

. svyreg ineed11 white boston chicago [pweight=normchwt]

Survey linear regression

pweight:  normchwt                                Number of obs    =      2399
Strata:   <one>                                   Number of strata =         1
PSU:      <observations>                          Number of PSUs   =      2399
                                                  Population size  = 2394.4117
                                                  F(   3,   2396)  =      1.82
                                                  Prob > F         =    0.1422
                                                  R-squared        =    0.0058
------------------------------------------------------------------------------
 ineed11 |      Coef.   Std. Err.       t     P>|t|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
   white |   .0333805    .0751822      0.444   0.657     -.1140482    .1808093
  boston |  -.0582959    .0564417     -1.033   0.302     -.1689755    .0523837
 chicago |   .0405888    .0564837      0.719   0.472     -.0701731    .1513507
   _cons |   .7336379    .0464126     15.807   0.000       .642625    .8246509
------------------------------------------------------------------------------
. save strat
```