Fragile Families & Child Wellbeing Study: Methodology for Constructing Mother, Father, and Couple Weights for Core Telephone Public Survey Data Waves 1-4

Prepared by:

Barbara Lepidus Carlson

Mathematica Policy Research

April 2008

Table of Contents

Section	Page
1. Mother Weights	
A. Baseline Weights	3
B. Follow-up Weights	14
2. Father Weights	
A. Baseline Weights	20
B. Follow-up Weights	22
3. Couple Weights	
A. Baseline Weights	28
B. Follow-up Weights	28
4. Variance Estimation	36
Appendix	
A. Table A – Variables Used in the Response Propensity Score Models	41

SECTION 1 – MOTHER WEIGHTS

This memorandum documents the methodology for constructing the weights used for analysis of the survey data provided by mothers as part of the core data collection effort for the Fragile Families and Child Well-Being Study. The target survey population for this study can be summarized as: live births occurring in large cities, by mothers who plan to keep the child, can identify the still-living father, and speak English or Spanish.

The first part of the memorandum will cover the baseline weights, and the second will cover the weights for the follow-up surveys conducted at one year, three years, and five years after baseline. For each of these four time periods, there are three sets of weights: two for doing analysis at the national level and one for analysis at the city level. These are discussed in detail below.

The most important step was to create a set of baseline weights that make the sample of births represent all eligible births occurring in large metropolitan areas of the United States during the study period. All the follow-up weights would then sum up to this baseline total, which represents the population of children eligible for the study at baseline. For each stage of sampling and data collection, I discuss below which units are considered to be eligible respondents. Each stage of weighting follows three basic steps: (1) calculate the probability of selection, (2) adjust weights for nonresponse, and (3) poststratify weights to known totals. Issues related to variance estimation will be covered in a separate memorandum.

A. BASELINE WEIGHTS

Though the mother completed the baseline survey, the ultimate sampling unit for this study is actually a birth. The first several steps therefore involve calculating the probability of selection for the birth as the sampling unit, followed by nonresponse adjustments related to the mother as a data collection unit. This was a multistage sample, where the first stage of selection was the city, the second stage the hospital, and the third the birth. The final weight for the birth (which accounts for the probability of selection and nonresponse at various stages) is cumulative across these three stages. (See *Fragile Families: Sample and Design* by Reichman et al. for more details on the sampling methodology at each stage.[1])

Selecting Cities. There were 77 cities on the sampling frame. These included all U.S. cities with a population of 200,000 or more in 1994.[2] These cities were classified into 9 strata: 8 strata with 1 selection each, and 1 stratum with 8 selections. These 16 cities were selected with probability proportional to size (1994 population). Two (Santa Ana and Birmingham) were later considered to be “nonparticipating” cities, and were replaced (non-probabilistically) by other cities in the same strata: San Jose and Baltimore, respectively. These two cities were treated as certainty selections for the purpose of making national estimates. Four additional cities were not selected randomly but were included in the study because of interest to funders. Because it was not decided *a priori* that these four cities would be included with certainty, but that they would be included only after they were not randomly selected, we could not assign them a probability of selection of 1. So these non-probabilistically selected cities would be included only for city-specific, not national, analyses.

---

[1]*Children and Youth Services Review*, Vol. 23, nos. 4/5, pp. 303-26, 2001, available at [www.fragilefamilies.princeton.edu/surveys/Reichman_et_al_2001.pdf].

[2]In this memorandum, the term *national* refers to all 77 U.S. cities with 1994 populations of 200,000 or more.

To summarize, we selected 22 cities to be in the study, 2 of which were considered to be nonparticipants within the national sample. Among the 20 participating cities, 16 were considered part of the national sample (including San Jose and Baltimore). The first 2 participating cities started baseline data collection in 1998. The next 5 cities started in 1999. And the final 13 started data collection in 2000.

For the 16 cities involved in national estimates, we calculated the probability of selection as:

$p_c = \frac{pop94_c \cdot n_h}{\sum_{\text{cities c in } h} pop94_c}$, where n_h is the number of cities selected in stratum h and $pop94_c$ is the city's population in 1994. (San Jose and Baltimore were given a probability of selection equal to 1, and the probability of selection of other selections in their strata were adjusted so that San Jose or Baltimore was excluded from the n_h and the sum in the denominator.)

We then did a nonresponse adjustment for the two nonparticipating cities. Two weighting cells were formed: one comprising the 8 sampling strata from which 1 city was selected in each, and one comprising the sampling strata from which 8 cities were selected. Using the inverse of the probability of selection (calculated using the formula above), the nonresponse adjustment was the inverse of the weighted response (participation) rate within each cell.

$$cityadj_{cell} = \frac{\sum_{\text{selected cities in cell}} (1/p_c)}{\sum_{\text{participating cities in cell}} (1/p_c)}$$

Princeton University also asked us to create a series of weights that treated City X like a nonparticipating city, along with Santa Ana and Birmingham. This was because City X, being the first city in the study, did not have the same data collected as other cities. For analyses that involved the data not collected for City X, these weights would be appropriate. Using the same methodology as outlined above, we used the inverse of the participation rate within cells to construct a weight that considered City X as a nonparticipating city. So in addition to city-specific weights for each of the 20 cities in the study, there are two sets of national weights, one in which City X is considered a participating city, and one in which it is not. Below is the nonresponse adjustment factor excluding City X.

$$cityadj2_{cell} = \frac{\sum_{\text{selected cities in cell}} (1/p_c)}{\sum_{\text{participating cities in cell, excluding Austin}} (1/p_c)}$$

The final weight for cities that are part of the national sample can be calculated as:

$$citywt_c = 1/p_c \cdot cityadj_{cell} \text{ and}$$
$$citywt2_c = 1/p_c \cdot cityadj2_{cell}.$$

Selecting Hospitals. In the 5 smallest cities, all hospitals were selected for inclusion in the study. In the 13 largest cities (other than New York and Chicago), the hospitals with the largest number of births to nonmarried parents ("nonmarital births") were included, such that 75 percent of such births occurred in these hospitals. Because this was a non-probabilistic sample, we decided to treat the hospitals in the city as having been selected with certainty (as in the 5 smallest cities), but with undercoverage in the frame, to be taken care of with poststratification. In New York and Chicago, a random sample of hospitals with 1,000 or more nonmarital births per year was selected. We constructed a sampling weight accounting for the random selection, and then accounted for the births in hospitals with fewer than 1,000 nonmarital births using poststratification at the birth level. For cities other than New York and Chicago, the probability of selection of each hospital *s* was set to 1. For New York and Chicago, the probability of selection of each hospital *s* was set to:

$$p_s = \frac{\text{number of hospitals selected in city}}{\text{number of hospitals in city eligible for selection}}$$

There were 210 hospitals in the 20 participating cities, and there were 82 hospitals selected for the core study and 128 hospitals not selected according to the methods described above. Among the 82 selected hospitals, 72 participated in the study. There were 9 hospitals that did not agree to participate, and 1 considered to be ineligible. We adjusted the hospital sampling weight to account for nonresponse among eligible hospitals, using the city as the weighting cell.

$$hospadj_c = \frac{\sum_{\text{eligible sampled hospitals in city}} (1/p_s)}{\sum_{\text{participating sampled hospitals in city}} (1/p_s)}$$

The hospital weight can then be calculated as:

$$hospwt_s = 1/p_s \cdot hospadj_c.$$

This weight can be used as is for city-specific estimates about hospitals. For national estimates about hospitals, the weights that account for the probability of selection and participation status of the city would be calculated as:

$$hospwt_s(nat) = citywt_c \cdot hospwt_s$$
$$hospwt_s(nat2) = citywt2_c \cdot hospwt_s$$

Selecting Births. Within each of the selected hospitals, during the time period the study was ongoing, we selected all nonmarital and marital births until their respective quotas were achieved. For each hospital or city, there were quotas set for the number of nonmarital and marital births to be selected. The quota for marital births was set as 23 or 25 percent of the quota for all births. Because there was no quantifiable probability of selection for births, we used the following method to calculate a weight given the available information.

For each selected city, we had available some vital statistics (number of births) by hospital and marital status that were from 1996, two to four years old (depending on the city) at the time of baseline data collection. For the hospitals selected for inclusion in this study—but not for the non-selected hospitals—we also had available the total number of births from the year prior to data collection, but not by marital status. Using the number of baseline completes by hospital and marital status, we constructed a base weight using the procedures described below.

First, we used the marital and nonmarital *proportions* found in the older vital statistics for each hospital, and applied these hospital-specific proportions to the more current number of births for each selected hospital. We then created a preliminary weight that was equal to the estimated number of births divided by the number of completes (by hospital and marital status). This preliminary weight partially accounts for both selection probability and response rate. (We did not have available the total number of eligible births, the number of selected births, or the number of selected but ineligible births by hospital and marital status. These would have allowed us to calculate a probability of selection and then perform a nonresponse adjustment to the sampling weight.)

$$BASEWT1_{s,m} = \frac{\text{estimated annual births}_{\text{s,m}}}{\text{baseline completes}_{\text{s,m}}}$$

We then looked at the coverage rate of births by marital status; that is, what percent of births in a city were covered by the selected hospitals. This was based on the older vital statistics data that we had for all hospitals in a city.

$$COVGRATE_{c,m} = \frac{\text{births in selected city hospitals}_{\text{m}}}{\text{births in all city hospitals}_{\text{m}}}$$

$$BASEWT2_{s,m} = \frac{\text{BASEWT1}_{\text{s,m}}}{\text{COVGRATE}_{\text{c,m}}}$$

We then looked at the eligibility rate within city; that is, what percent of selected births in a city were eligible at baseline, and applied this to the coverage-adjusted weight.[3] The counts for various status codes were aggregated into categories as follows:

[3] Note that this was not done within marital status because the majority of the ineligible births had unknown marital status and the eligibility rate formula includes ineligible births in the denominator. For the first two cities, the eligibility rate was not available, and so was imputed as .85, which was the mean eligibility rate from the other 18 cities.

How Statuses Were Aggregated in Counts Used for Calculating Eligibility Rate	Status of Mother Baseline Survey
Eligible	Complete
	refused interview/screening complete
	mother too ill/screening complete
	mother left hospital after screening
Ineligible	does not speak English or Spanish
	baby to be adopted
	baby's father is deceased
	mother is a minor (in hospitals not agreeing to interview minor parents)
	baby was stillborn
Noncompletes with Undetermined Eligibility (Excluded from Eligibility Rate Calculation)	refused screening
	mother too ill for screening
	mother left hospital before screening
	mother about to leave hospital before screening
	hospital denied permission
	married—quota already met
	births not considered part of the core study

$$ELIGRATE_c = \frac{\text{count of eligible sampled births in city}}{\text{count of eligible + ineligible sampled births in city}}$$

$$BASEWT3_{s,m} = BASEWT2_{s,m} \cdot ELIGRATE_c$$

Finally, we adjusted for sampling in the presence of multiple births. If a woman had twins or triplets, one child was randomly selected to be part of the study.

$$BIRTHWT_i = BASEWT3_{s,m} \cdot NUMBABIES_i$$

To construct city-specific birth weights that account for the probability of selection and participation status of both the hospital and the birth/mother (*i*), we multiply the final hospital weight by the birth weight.

$$birthwt_i(city) = hospwt_s \cdot birthwt_m$$

To construct national birth weights that further account for the probability of selection and participation status of the city (with and without City X), we add one more factor to the product: the final city weight.

$$birthwt_i(nat) = citywt_c \cdot hospwt_s \cdot birthwt_m$$
$$birthwi_i(nat2) = citywt2_c \cdot hospwt_s \cdot birthwt_m$$

The final number of eligible responding baseline mother interviews (births) that were part of the core study was 4,789.

Challenges. There were several difficulties encountered in calculating the mother baseline (birth) weights. At certain stages of sampling, the lack of available information on selection probabilities and dispositions posed challenges in the weighting process.

At the city level, most of the cities were selected with probability proportional to size. Others were added purposively after the random selection, either to replace selected but nonparticipating cities or for other reasons. All but two of the non-randomly selected cities are included in the study for city-specific estimates only. (See earlier discussion of San Jose and Baltimore being included in national estimates, even though those cities were not selected probabilistically.)

At the hospital level, the various methods used to include hospitals in the sample were not probabilistic (unless all hospitals were selected, as in the smallest cities). For the largest cities (other than New York and Chicago), a cumulative quota-type selection was used, which meant that a probability of selection could not be used to calculate a weight. Instead, the hospitals that were excluded (because 75 percent of nonmarital births in the state were covered by the included hospitals, which had more nonmarital births) would be accounted for in the final poststratification adjustments. While New York and Chicago had a probabilistic way of selecting hospitals, this was done only among those hospitals with a minimum annual number of nonmarital births, so the same undercoverage issue exists for the smaller hospitals in these two cities. Extra hospitals were included in the study, but they are not considered to be part of the core study and were excluded from the weighting process. Finally, there was some uncertainty about whether certain hospitals were "released" into the sample.

At the birth level, the lack of information about the number eligible for selection, the number selected, and the number of births later coded as ineligible--by hospital and marital status--required us to combine the sampling weight (accounting for the probability of selection) and the nonresponse adjustment into one factor, using population counts estimated from various sources (some as many as four years old as of the baseline data collection period). The weights also treated the sampled births as though they were a random sample from births in a hospital across the entire year when in fact they were a quota sample collected over a specific time period of several weeks that varied from hospital to hospital. We then had to account for coverage rates for selected hospitals and eligibility rates among selected births. These coverage rates differed from city to city, and by marital status within city.

The previous weighting steps take into account unequal selection probabilities and response rates across cities, hospitals, and births. However, the weighted estimates produced with these weights could, because of certain sampling features, over- or underestimate the true number in the Fragile Families population. For example, we knew about the coverage issue at the hospital level and the lack of sampling probabilities at the birth level. Thus, the distributions of characteristics in the sample may fail to reflect those in the population, even after adjustments. When we looked at the weighted number of births after all these adjustments, it was clear that the weighting to this point—that is, the sampling weights--did not fully account for the coverage, sampling, and nonresponse issues.

Raking.[4] The accuracy of the estimates might be improved if we know the population's distribution and are able to adjust the individual survey weights to it. The adjusted weights would properly represent births in the population. The most commonly used adjustment method is known as poststratification, or *raking.* Because of the challenges outlined above, the raking adjustments to external totals of births by city became crucial for making up for the deficiencies in the weights described above. The basic nonrespondent-adjusted sampling weights are used as an input for the raking process. Raking adjusts these weights by aligning the total sum of the weights for selected variables, which are considered as risk factors in the study.

Unfortunately, raking posed its own difficulties. The external totals that would be required for such an adjustment would be the number of marital and nonmarital births occurring in each of the 20 cities in the study (during 1998, 1999, or 2000, depending on the city).[5] Unfortunately, such numbers do not readily exist. The Centers for Disease Control and Prevention (CDC) provides annual natality data on a Natality Detail public use file. These data contain important characteristics, such as mother's marital status, race/ethnicity, age, and education. In addition to demographic variables, the CDC natality file has variables that represent geographical location for both the birth's occurrence and the mother's residency. This information may be used to match an individual CDC birth record to a geographical area. Each record can be identified by city, county, and state codes for the mother's residence. Unfortunately, city codes are not available for the birth's place of occurrence. So while one can find annual birth counts by the city of mother's residence, by the county of mother's residence, and by the county in which the birth occurred, one cannot find in these data what we need for our weighting purposes here: the counts by the city of occurrence. After we looked at our survey data and found substantial numbers of mothers whose city of residence did not match the city in which she gave birth, we determined that using the counts by city of mother's residence as a substitute for counts by city of birth occurrence was not a viable option.

If the city and county boundaries are one and the same, we can use the counts by the county of birth occurrence. Similarly, if no hospitals in a county are outside the city boundaries (discerned from American Hospital Association data), then we can use the counts by the county of birth occurrence. For all other situations, we had to create a synthetic estimate of the number of births occurring in a city.

First, more about raking in general. It is a method of poststratifying the weights to ensure that the weighted counts of the sample are consistent with the known counts of the population within raking cell. In this study, the variables used for the raking process are given in Table 1.1. Even though the adjustment was done within individual raking cells, the raking process requires only known marginal population totals for a single variable, rather than totals for individual cells in which multiple variables are crossed. We therefore do not need to worry about empty or very small cells formed by crossing all the variables used in the poststratification process.

---

[4]Acknowledgement: The section of this memorandum describing the raking procedure was originally drafted by Mary Edith Bozylinsky (formerly of Mathematica) and Amang Sukasih and Donsig Jang of Mathematica.

[5]Note that this still does not precisely match the *survey* target population defined on the first page of this memorandum, because these birth counts would not exclude those who do not plan to keep the child, those who cannot identify the still-living father, and those who speak a language other than English or Spanish. By raking to the total number of births in these cities, we make any estimated *totals* from the survey match a slightly larger population of births. If we assume that the *distribution* of key outcome variables is similar between the survey target population and these slightly larger outside totals, then distributional estimates should be reasonable after raking.

TABLE 1.1

LIST OF RAKING VARIABLES

Variable Name	Description	Levels
Marital Status	Mother's Marital Status	2
Education	Mother's Education Level	5
Race/Ethnicity	Mother's Race/Ethnicity	4
Age	Mother's Age	7

Raking is an iterative process in which adjustments are made to scale the weights to the known marginal population totals for each raking variable. In each step, the weights are adjusted so that the weighted counts equal the population totals for each level of a particular raking variable. After each step, however, the weighted counts and the population counts may not be equal for the levels of other variables, so the adjustment process is repeated in an iterative manner until the differences between the weights in the previous iteration and the current iteration converge to a predetermined value. We implemented the raking algorithm using a SAS macro.

For this study, raking adjusts the weights attached to individual sampled births, so that the sums of these weights match population counts. The goal is to produce three sets of weights—individual city-level, national-level with all cities part of the national sample, and national-level excluding City X—so we need population counts for the 77 cities that were eligible for sampling. As mentioned in the previous section, city information is not available in CDC data for the birth's occurrence; thus, population birth counts are not available on the city level. Population birth counts are available for county level, however, and other useful information from the CDC natality file is city and county information for the mother's residency. We estimated the total number of births for the city level using the information available from the sample data and from the CDC natality file as follows:

(1) Even though city information is not available for the birth's occurrence, certain cities may use the county's information in place of the missing city information. Two types of cities qualify: (a) those that have boundaries identical to those of their county, and (b) others that contain all the hospitals for their county. For these cities, we may use county birth totals as city birth totals. Table 1.2 presents a list of the cities from the sample that are type (a) or (b).

TABLE 1.2

LIST OF CITIES OF TYPE (A) OR (B)

Type	Fragile Families City
(a) Cities that have boundaries identical to those of their county	New York Norfolk Baltimore Philadelphia Richmond
(b) Other cities that contain all the hospitals for their county	Corpus Christi San Antonio

(2) For the remaining cities, we estimated the number of births for each city using both mother's residence and birth occurrence information. For each of the counties containing the 77 cities eligible for the Fragile Families study ("the FF cities"), we partitioned the county-level birth count (sum of births occurring within the county) into three parts:

A = total number of births given by mothers living in the FF city,

B = total number of births given by mothers living in a different city, but in the same county as the FF city,

C = total number of births given by mothers living in a county not associated with the FF city.

The estimate of FF city-level birth count is computed as

(1) $$D = rA + sB + tC$$

where

D = estimate of total number of births in the FF city,

r = among births in county given by mothers living in FF city, proportion of births occurring in the FF city,

s = among births in county given by mothers living in the county (but not in FF city), proportion of births occurring in the FF city,

t = among births in county given by mothers living outside the county, proportion of births occurring in the FF city.

The value of r is assumed to be large ($0.9 < r < 1$), since it is reasonable to assume that, in general, mothers who live in a particular city give birth in a hospital within the same city, especially for large cities such as those in this study.

Intuitively, s should be larger than t. In this case, however, there is no compelling reason to treat non-FF cities differently regardless of their locality within or outside an FF county; therefore, the values of s and t are assumed to be equal. Under this assumption, the above equation can be simplified into

$$(2) \qquad D = rA + u(B + C)$$

where

u = among births in county given by mothers living outside the FF city, proportion of births occurring in the FF city.

Now, among the births occurring in a particular FF city, the proportion of births where the mother lives in the FF city is:

$$(3) \qquad P = \frac{rA}{rA + u(B + C)}$$

The value of P can be estimated from the study sample (comprising mothers giving birth in the FF cities), because the survey questionnaire asks for the mother's city of residence. Suppose p denotes such a proportion from the sample data. The values of p as obtained from the survey data are shown in Table 1.3 for the 13 cities not of type (a) or (b) above.

Hence, the estimate of u can be obtained by replacing P with p in (3) and solving it for u as follows:

$$(4) \qquad u = \frac{\left(\frac{1}{p} - 1\right) rA}{B + C}$$

For purposes of raking, we assumed $r = 0.9$, we estimated p from the survey data as above, and we obtained the values of A, B, and C from the CDC data (1998, 1999, or 2000, as appropriate for the particular city). Plugging the values back into equation (2), we get an estimate of D_{level}, the number of births occurring within a particular FF city for a particular level of a raking variable. (For cities of type (a) or (b) above, we set D equal to $A+B+C$.)

TABLE 1.3

ESTIMATE OF PROPORTION OF BIRTHS GIVEN BY MOTHERS LIVING IN THE FF CITY, AMONG ALL BIRTHS OCCURRING IN THE FF CITY (*p*)

FF City	Proportion of Mothers Living in the City of Interview *(p)*
Austin	0.90
Boston	0.79
Chicago	0.85
Detroit	0.86
Indianapolis	0.78
Jacksonville	0.89
Milwaukee	0.83
Nashville	0.56
Newark	0.64
Oakland	0.95
Pittsburgh	0.80
San Jose	0.72
Toledo	0.65

Source: FF survey data

A series of raking adjustments across various raking variables would then be calculated as:

$$rakeadj_{level} = \frac{D_{level}}{\sum_{s,m \in level} \text{raking-adjusted birthwt}_{s,m}}$$

Applying the series of raking adjustments to the nonresponse-adjusted birth weights, we get:

$$birthwtrake_{citycell}(city) = birthwt_{s,m}(city) \cdot \prod_{levels} rakeadj_{level}$$
$$birthwtrake_{natcell}(nat) = birthwt_{s,m}(nat) \cdot \prod_{levels} rakeadj_{level}$$
$$birthwtrake_{natcell}(nat2) = birthwt_{s,m}(nat2) \cdot \prod_{levels} rakeadj_{level}$$

Trimming. After raking the weights as described above, we trimmed them to remove any outliers that may have occurred due to a large adjustment factor or combination of factors. For each of the three types of baseline weights, we determined the mean *(M)* and the standard deviation *(S)* of each weight, by marital status *(i)*. We set the trim value for marital status *i* to $M_i + 4S_i$. That is, we considered any weights that were more than four standard deviations higher than the mean weight value to be outliers and trimmed them to that maximum value. After trimming, we re-raked the weights. To give an example of the level of trimming, for one of the national weights (including City X), we ended up trimming 25 out of 4,789 weights, or 0.5 percent.

City-Level Weights. The birth weights for individual FF city-level (**m1citywt**) were developed to provide users of the mother baseline survey data with final survey weights for analyses within individual cities. Using the methods explained in the previous section, we adjusted/raked these weights so that they are consistent with total population counts of births in large U.S. cities based on CDC data.

National Weights. The national-level weights are the final survey weights attached to individual births for analyses that pool records for the 16 national-sample cities within the sample. The analysis generalizes to births occurring in the 77 large cities defined as the FF population. The weights were developed based on national weights (computed in the earlier steps), which were in turn raked to total (population) birth counts in the 77 cities based on CDC data.

We computed two sets of national-level weights: one based on all 16 of the national-sample cities in the sample, with all 77 cities as the population being targeted (**m1natwt**), and the other based on only 15 cities (City X is excluded) in the sample, with all 77 cities as the population being targeted (**m1natwtx**).[6]

B. FOLLOW-UP WEIGHTS

The final baseline weight serves as the anchor for all the follow-up weights. Because there is no subsampling at the various follow-ups, we concern ourselves mainly with nonresponse adjustments and re-raking to the baseline totals.

There was much discussion about how to define the eligible population at each stage of follow-up. But the final decision was that a case was ineligible at follow-up only if the child associated with the sampled birth died. (We also treated as ineligible those cases that were released to the sample in error, because they were duplicates.) Because of the rarity of this type of situation, we consider all sample members to have known eligibility status; that is, even if we are unable to locate a sample member, we assume the case is still eligible unless we learn otherwise.

Other types of situations were those in which a survey was not intended to be conducted, according to study protocol, because the questions did not apply. These included cases in which the child was adopted, neither parent had custody of the child, or one of the parents died. While no survey was completed (or only a few questions of the survey were answered), we decided that we would still consider these cases part of the target population, and that these situations (custody and parental death) were a type of outcome. Because we obtained all the needed information from the case as part of the disposition code, these cases were considered to be completes for weighting purposes. Princeton University constructed separate indicators (**cm*W*samp** flags, where *W* is the wave) to indicate whether some or all of a questionnaire was completed, among those considered to be part of the eligible completes within the sample. All other final dispositions were considered to be eligible noncompletes, subclassified as located or

---

[6]National analyses using **m1natwtx** should include only samples in the 15 cities, while analyses using **m1natwt** should include all records in the 16 national-sample cities.

unlocated. Table 1.4 shows how the various disposition codes were classified for weighting purposes, and Table 1.5 shows how the classification variables were assigned values.

TABLE 1.4

CLASSIFICATION OF FINAL DISPOSITION CODES FOR MOTHER WEIGHTING PURPOSES

<table>
<tr><th colspan="2">Classification for Weighting Purposes</th><th>Final Disposition Code</th></tr>
<tr><td rowspan="2">Eligible Complete</td><td>With Survey Data</td><td>01-Complete callout
05-HC Comp-field
06-HC Comp-field-incarcerated
07-Comp-field call in
08-Comp-incarcerated call in</td></tr>
<tr><td>Without Survey Data[7]</td><td>40-Deceased
42-Child adopted
46-Neither Parent had custody</td></tr>
<tr><td colspan="2">Ineligible[8]</td><td>43-Child deceased
48-Duplicate</td></tr>
<tr><td rowspan="2">Eligible Noncomplete</td><td>Located</td><td>20-Hung up during intro
21-Refusal
30-Language barrier
31-Ill/other barrier
32-Incarcerated
39-Other eligible
49-Other ineligible
54-Moved out of state
65-Maximum calls
66-Case retired
67-Marr dad/no field eff
69-Out of area/no effort</td></tr>
<tr><td>Unlocated</td><td>59-Cannot locate</td></tr>
</table>

[7]There were some cases that started to respond to the survey, but whose survey responses early in the questionnaire indicated that the child was adopted or that neither parent had custody, at which point the survey was terminated. The final disposition codes for these cases were changed from “complete” to “child adopted” or “neither parent had custody.”

[8]If a previous round of the survey indicated ineligibility (child deceased or duplicate), then the current round was classified as ineligible, regardless of the current disposition code.

TABLE 1.5

VALUES ASSIGNED TO CLASSIFICATION VARIABLES FOR MOTHER WEIGHTS

	LOC_*Mi* Located Status	ELIG_*Mi* Eligibility Status	ELIGD_*Mi* Eligibility Determined	COMP_*Mi* Completion Status
Eligible Complete	1	1	1	1
Ineligible	1	3	1	2
Located Eligible Noncomplete	0	1	1	.
Unlocated Eligible Noncomplete	1	1	1	2

Note: Subscript *i* in the variable names is equal to 2 for the 1-year follow-up, equal to 3 for the 3-year follow-up, and equal to 4 for the 5-year follow-up.

We adjust the weights of the eligible completes to account for those of the eligible noncompletes in two stages. Each follow-up weight starts with the final poststratified baseline weight (national, national without City X, or city, as appropriate). First we adjust for unlocatability; that is, we adjust the initial weights for all the eligible located cases upward to account for those of the eligible unlocated cases. Then we adjust for nonresponse among the located; that is, these adjusted weights for the eligible located completes are further adjusted upward to account for those of the eligible located noncompletes. These adjustments were made within weighting cells. (The formation of these weighting cells is described below.) Each cell comprises sample members who have similar response propensities. Once the cells are formed, the two sets of adjustments are made separately for each of the two national weights and the city weight.

To do these adjustments, we must form weighting cells within which to make the adjustments. We first modeled the propensity separately for the two types of nonresponse among eligibles: (1) unlocated, and (2) noncomplete among located. We determined a set of covariates to use as candidates for these models from among baseline survey variables, which are available for both respondents and nonrespondents in these follow-up surveys. These were baseline variables that we thought would predict response propensity. We temporarily imputed values if values were missing. We developed two separate unweighted stepwise logistic regression models to predict the two types of nonresponse. (The stepwise parameters were that the significance level for entry into the model was .15 and the significance level for staying was .20.) One of the baseline covariates was a city indicator. We separately examined whether any particular cities were significant predictors of each type of nonresponse. If so, we included them as possible covariates for the stepwise logistic regression models. After convergence of the final model, we used the propensity scores to form deciles for the national weights, and quintiles within city for the city-level weights. We used these deciles and quintiles to form the weighting cells for the nonresponse adjustments. Note that the ineligible sample members were excluded from these nonresponse adjustments and simply retained their initial weight from baseline.

Once the two national weights and the city weight were adjusted for the two types of nonresponse, we brought the ineligible weights back in. Then, as described above for the baseline weights, we rake the nonresponse-adjusted weights to their baseline totals, trim any outlier weights, and re-rake the weights.

Checking. After each set of weights is produced, we check them along several fronts. If we include the ineligible cases (which have positive weights for the follow-up surveys), the sum of the follow-up weight should be equal to the sum of the comparable baseline weight (national, national without City X, and city). We then classify the case according to its city; that is, whether the city is City X, or part of the national sample, or not part of the national sample. We cross that classification with the eligibility status, locatability status, and completion status, and then check whether the weight is appropriately missing or has a positive value. The city weights should all have a positive value if the case is (1) eligible and located and complete, or (2) ineligible and noncomplete. For the national weights (including and excluding City X), the same rules apply, except for the following. For national weights including City X, those from the four cities that not part of the national sample will have zero weights. For national weights excluding City X, those from those four cities plus City X will have zero weights.

We also look at summary statistics of the ratio between the follow-up weight and its comparable baseline weight (for the city-specific weight, this is done separately by city) to see if there are any extreme values. We print out the cases with the five highest and five lowest values of this ratio, to look at the various weighting adjustment factors for valid explanations as to why these ratios were so high or so low. Table 1.6 shows the final weighting variables and the sums of the weights.[9]

[9] Cases without valid survey data (except for cases who were dead/child not living with either parent) are recoded to missing in the final version of the weights so the sums in the data files will not match these sums.

TABLE 1.6

MOTHER BASELINE AND FOLLOW-UP WEIGHTS SUMMARY

Final Survey Mother Weight		Weight Variable Name	Sum of the Weights
National Level Including City X (16 cities)	Baseline	m1natwt	1,131,308.36
	1-Year Follow-up	m2natwt	1,131,308.36
	3-Year Follow-up	m3natwt	1,131,308.36
	5-Year Follow-up	m4natwt	1,131,308.36
National Level Excluding City X (15 cities)	Baseline	m1natwtx	1,131,308.36
	1-Year Follow-up	m2natwtx	1,131,308.36
	3-Year Follow-up	m3natwtx	1,131,308.36
	5-Year Follow-up	m4natwtx	1,131,308.36
City Level (20 cities)	Baseline	m1citywt	347,237.90
	1-Year Follow-up	m2citywt	347,237.80
	3-Year Follow-up	m3citywt	347,237.80
	5-Year Follow-up	m4citywt	347,237.81

Table 1.7 shows the covariates used for each of the response propensity models, after the stepwise regression procedure, and any other differences across the various weights.

TABLE 1.7

INDEPENDENT VARIABLES IN FINAL PROPENSITY MODELS USED TO FORM CELLS FOR MOTHER WEIGHTS[10]

Locatability Adjustment			Response Adjustment for Located		
1-year	3-year	5-year	1-year	3-year	5-year
M1CITY*	M1CITY*	M1CITY*	M1CITY*	M1CITY*	M1CITY*
M1A11B	M1A11A	M1A11B	M1B27	M1B27	M1B27
M1B8	M1A11C	M1B3	M1D2A	M1D2E	M1D1D
M1D2D	M1A15	M1D1B	M1D2F	M1E3A	M1D2C
M1D2E	M1B8	M1D1E	M1D2G	M1E3E	M1E3E
M1D2F	M1D2D	M1D2B	M1E4C	M1I3 (grpd)	M1I1 (grpd)
M1E3E	M1D2E	M1D2C	M1F4	M1G1	M1H3A+M1H3
M1E4A	M1E4C	M1D2D	M1J3	M1J5	M1A4+M1B2
M1I11	M1F2	M1E3E	CM1AGE (grpd)	M1G4 (grpd)	M1A9 (grpd)
M1I3 (grpd)	M1F3	M1E4A		M1A13+M1A13A	COMP_M1
M1F3	M1F6	M1F6		COMP_M1	COMP_M3
M1J3	M1G1	M1A13+M1A13A			
M1J5	M1I1 (grpd)	M1A9 (grpd)			
CM1AGE (grpd)	M1A13+M1A13A	LOC_M1			
M1G2 (grpd)	LOC_M1	LOC_M3			
M1G3 (grpd)					
M1A13+M1A13A					
*Significant cities for this model: New York San Jose	*Significant cities for this model: Newark Milwaukee New York San Jose	*Significant cities for this model: Oakland Baltimore Philadelphia New York San Jose	*Significant cities for this model: Oakland Newark New York San Jose	*Significant cities for this model: New York San Jose	*Significant cities for this model: Philadelphia New York San Jose Boston

Note: Cities not found to be significant predictors of nonresponse for a particular model were grouped together into one value, while significant cities retained their original identifiers for the model.

[10]See Appendix for what these baseline variables represent. We imputed missing values of these variables for purposes of propensity modeling only.

SECTION 2 – FATHER WEIGHTS

See Section 1 for how we created a set of baseline weights that make the sample of births represent all eligible births occurring in large metropolitan areas of the United States during the study period. All the father weights (baseline and follow-up) would then sum up to this baseline total, which represents the population of children eligible for the study at baseline. For each stage of sampling and data collection, we discuss below which units are considered to be eligible respondents. Each stage of weighting follows three basic steps: (1) calculate the probability of selection, (2) adjust weights for nonresponse, and (3) poststratify weights to known totals. Issues related to variance estimation will be covered in a separate memorandum.

A. BASELINE WEIGHTS

All fathers associated with sampled births are part of the study, so their marginal probabilities of selection and corresponding sampling weights are equal to one. Therefore, for father baseline weights, we actually start with the final *mother* baseline weights (which represent the sampled births), and then adjust these weights for father nonresponse. The father baseline sampling weight is:

$$\text{father baseline sampling weight} = \text{final } \textit{mother} \text{ baseline weight} \cdot 1$$

where the final mother baseline weight is *m1natwt*, *m1natwtx*, or *m1citywt*, depending on the type of estimate being made.

Because the baseline data collection involved more than one contractor, final baseline disposition codes for mothers and fathers were not readily available at the hospital level for many of the cities, although aggregate numbers were available by city and by marital status. For the first two cities (City X and Oakland), no disposition code information was retained, other than number of completes by city and marital status.

We classified the father baseline interview as either complete (COMP_F0 = 1) or noncomplete (COMP_F0 = 2). All fathers were assumed to be eligible at baseline (ELIG_F0 = 1). We calculated the survey nonresponse adjustment factor for father baseline weights as follows:

$$\textit{respadj}_{cell} = \frac{\sum_{\text{all fathers in cell}} (\text{father baseline sampling weight})}{\sum_{\text{responding fathers in cell}} (\text{father baseline sampling weight})}$$

where the cell is defined by city and marital status, and also by race/ethnicity if the cells were large enough.

To construct the two types of national weights and city weights for fathers at baseline, we multiply the final *mother* baseline weight by the nonresponse adjustment factor calculated above.

For responding fathers:

father0wt(nat)=m1natwt · respadj
father0wt(nat2)=m1natwtx · respadj
father0wt(city)=m1citywt · respadj

For nonresponding fathers, these weights were set equal to 0.

Raking. See the Section I for more information about raking in general and some specifics for this study. The variables used for the raking process are given in Table 2.1. Even though the adjustment was done within individual raking cells, the raking process requires only known marginal population totals for a single variable, rather than totals for individual cells in which multiple variables are crossed. We therefore do not need to worry about empty or very small cells formed by crossing all the variables used in the poststratification process.

TABLE 2.1

LIST OF RAKING VARIABLES FOR FATHER WEIGHTS

Variable Name	Description	Levels
Marital Status	Mother's Marital Status	2
Education	Mother's Education Level	5
Race/Ethnicity	Mother's Race/Ethnicity	4
Age	Mother's Age	7

The raking adjustment would be calculated as:

$$rakeadj_{cell} = \frac{D_{cell}}{\sum_{cell} father0wt},$$

where *D* is the synthetic estimate (using CDC Natality data) of the number of births by marital status in the appropriate city (or nationally) and year. Applying the raking adjustment to the nonresponse-adjusted birth weights, we get:

$$father0wtrake_{citycell}(city) = rakeadj_{citycell} \cdot father0wt_{hosp,maritalstatus}(city)$$
$$father0wtrake_{natcell}(nat) = rakeadj_{natcell} \cdot father0wt_{hosp,maritalstatus}(nat)$$
$$father0wtrake_{natcell}(nat2) = rakeadj_{natcell} \cdot father0wt_{hosp,maritalstatus}(nat2)$$

Trimming. After raking the weights as described above, we trimmed them to remove any outliers that may have occurred due to a large adjustment factor or combination of factors. For each of the three types of baseline weights, we determined the mean *(M)* and the standard deviation *(S)* of each weight, by marital status *(i)*. We set the trim value for marital status *i* to $M_i + 4S_i$. That is, we considered any weights that were more than four standard deviations higher than the mean weight value to be outliers and trimmed them to that maximum value. After trimming, we re-raked the weights.

City-Level Weights. The father baseline city-level weight (**f1citywt**) was developed to provide users of the father baseline survey data with final survey weights for analyses within individual cities. Using the methods explained in the previous section, we adjusted/raked these weights so that they are consistent with total population counts of births in the large U.S. cities in this study, based on CDC data.

National-Level Weights. The national-level weights are the final father baseline survey weights attached to individual births for analyses that pool records for the 16 national-sample cities within the sample. The analysis generalizes to births occurring in the 77 large cities defined as the Fragile Families population. The weights were developed based on national weights (computed in the earlier steps), which were in turn raked to total (population) birth counts in the 77 cities based on CDC data.

We computed two sets of father baseline national-level weights: one based on all 16 of the national-sample cities in the sample, with all 77 cities as the population being targeted (**f1natwt**), and the other based on only 15 cities (City X is excluded) in the sample, with all 77 cities as the population being targeted (**f1natwtx**).[11]

Checking. After the three baseline weights were produced, we checked them along several fronts. The sum of the father baseline weight should have been equal to the sum of the corresponding mother baseline weight (national, national without City X, and city). We then classified the case according to its city; that is, whether the city was City X, or part of the national sample, or not part of the national sample. We crossed that classification with the completion status, and then checked whether the weight was appropriately missing or had a positive value. The city weights should have all had a positive value if the case was complete. For the national weights (including and excluding City X), the same rules applied, except for the following. For national weights including City X, those from the four cities that were not part of the national sample should have zero weights. For national weights excluding City X, those from those four cities plus City X should have zero weights.

We also looked at summary statistics of the ratio between the father baseline weight and its comparable mother baseline weight (for the city-specific weight, this is done separately by city) to see if there were any extreme values. We printed out the cases with the five highest and five lowest values of this ratio, to look at the weighting adjustment factors for valid explanations as to why these ratios were so high or so low.

B. FOLLOW-UP WEIGHTS

The final mother baseline weight also serves as the anchor for all the father follow-up weights at one year, three years, and five years after baseline. Because there is no subsampling

---

[11] National analyses using **f1natwtx** should include only samples in the 15 cities, while analyses using **f1natwt** should include all records in the 16 national-sample cities.

at the follow-ups, we concern ourselves mainly with nonresponse adjustments and re-raking to the mother baseline totals.

There was much discussion about how to define the eligible population at each stage of follow-up. The final decision was that a case was ineligible at follow-up only if the child associated with the sampled birth had died. Because of the rarity of this type of situation, we considered all sample members to have known eligibility status; that is, even if we could not locate a sample member, we assumed the case was still eligible unless we learned otherwise. (We also treated as ineligible those cases that were released to the sample in error, because they were duplicates.)

Other types of situations were those in which a survey was not intended to be conducted, according to study protocol, because the questions did not apply. These included cases in which the child was adopted, neither parent had custody of the child, or one of the parents had died. It also included cases where the father was unknown, the father denied paternity, the father was not told of the pregnancy, or the DNA test indicated this was not the father. While no survey was completed (or only a few questions of the survey were answered) for these situations, we decided that we would still consider these cases part of the target population, and that these situations (custody and parental death) were a type of outcome. Because we obtained all the needed information from the case as part of the disposition code, these cases were considered to be completes for weighting purposes. Princeton University constructed separate indicators (**cf*W*samp** flags, where *W* is the wave) to indicate whether some or all of a questionnaire was completed, among those considered to be part of the eligible completes within the sample. All other final dispositions were considered to be eligible noncompletes, subclassified as located or unlocated. Table 2.2 shows how the disposition codes were classified for weighting purposes, and Table 2.3 shows how the classification variables were assigned values.

We adjusted the weights of the eligible completes to account for those of the eligible noncompletes in two stages. Each follow-up weight started with the final poststratified mother baseline weight (national, national without City X, or city, as appropriate). First we adjusted for unlocatability; that is, we adjusted the initial weights for all the eligible located cases upward to account for those of the eligible unlocated cases. Then we adjusted for nonresponse among the located; that is, these adjusted weights for the eligible located completes were further adjusted upward to account for those of the eligible located noncompletes. These adjustments were made within weighting cells. (The formation of these weighting cells is described below.) Each cell comprised sample members who had similar response propensities. After the cells were formed, the two sets of adjustments were made separately for each of the two national weights and the city weight.

TABLE 2.2

CLASSIFICATION OF FINAL DISPOSITION CODES FOR FATHER WEIGHTING PURPOSES

<table>
<tr><th colspan="2">Classification for Weighting Purposes</th><th>Final Disposition Code</th></tr>
<tr><td rowspan="2">Eligible Complete</td><td>With Survey Data</td><td>01-Complete callout
05-HC Comp-field
06-HC Comp-field-incarcerated
07-Comp-field call in
08-Comp-incarcerated call in</td></tr>
<tr><td>Without Survey Data[12]</td><td>40-Deceased
41-DNA not dad
42-Child adopted
44-Father unknown
45-Father not told of pregnancy
46-Neither parent had custody
47-Father denies paternity</td></tr>
<tr><td colspan="2">Ineligible[13]</td><td>43-Child deceased
48-Duplicate</td></tr>
<tr><td rowspan="2">Eligible Noncomplete</td><td>Located</td><td>21-Refusal
28-Mother refused info on dad
30-Language barrier
31-Ill/other barrier
32-Incarcerated
33-Unavailable during field period
37-No prior interview – no field
39-Other eligible
49-Other ineligible
54-Moved out of state
65-Maximum calls
66-Case retired
67-Marr dad/no field eff
68-Mom no contact with dad
69-Out of area/no effort</td></tr>
<tr><td>Unlocated</td><td>59-Cannot locate</td></tr>
</table>

[12] There were some cases that started to respond to the survey, but whose survey responses early in the questionnaire indicated that the child was adopted or that neither parent had custody. At that point, the survey was terminated. The final disposition codes for these cases were changed from "complete" to "child adopted" or "neither parent had custody."

[13] If a previous round of the survey (mother's or father's) indicated ineligibility (child deceased or duplicate), or the current mother survey indicated ineligibility, then the current round was classified as ineligible, regardless of the current disposition code.

TABLE 2.3

VALUES ASSIGNED TO CLASSIFICATION VARIABLES

	LOC_*Fi* Located Status	ELIG_*Fi* Eligibility Status	ELIGD_*Fi* Eligibility Determined	COMP_*Fi* Completion Status
Eligible Complete	1	1	1	1
Ineligible	1	3	1	2
Located Eligible Noncomplete	0	1	1	.
Unlocated Eligible Noncomplete	1	1	1	2

Note: Subscript *i* in the variable names is equal to 2 for the one-year follow-up, equal to 3 for the three-year follow-up, and equal to 4 for the five-year follow-up.

To do these adjustments, we must form weighting cells within which to make the adjustments. We first modeled the propensity separately for the two types of nonresponse among eligibles: (1) unlocated, and (2) noncomplete among located. We determined a set of covariates to use as candidates for these models from among baseline survey variables, which are available for both respondents and nonrespondents in these follow-up surveys. These were baseline variables that we thought could predict response propensity. We temporarily imputed values if values were missing. We developed two separate unweighted stepwise logistic regression models to predict the two types of nonresponse. (The stepwise parameters were that the significance level for entry into the model was .15 and the significance level for staying was .20.) One of the baseline covariates was a city indicator. We separately examined whether any particular cities were significant predictors of each type of nonresponse. If so, we included them as possible covariates for the stepwise logistic regression models. After convergence of the final model, we used the propensity scores to form deciles for the national weights, and quintiles within city for the city-level weights. We used these deciles and quintiles to form the weighting cells for the nonresponse adjustments. Note that the ineligible sample members are excluded from these nonresponse adjustments and simply retain their initial weight from baseline.

After the two national weights and the city weight have been adjusted for the two types of nonresponse, we bring the ineligible weights back in. Then, as described above for the father baseline weights, we rake the nonresponse-adjusted weights to the mother baseline totals, trim any outlier weights, and re-rake the weights.

Checking. After each set of follow-up weights was produced, we checked them along several fronts. If we included the ineligible cases (which have positive weights), the sum of the follow-up weight should have been equal to the sum of the corresponding mother baseline weight (national, national without City X, and city). We then classified the case according to its city; that is, whether the city was City X, or part of the national sample, or not part of the national sample. We crossed that classification with the eligibility status, locatability status, and completion status, and then checked whether the weight was appropriately missing or had a positive value. The city weights should have all had a positive value if the case was (1) eligible and located and complete, or (2) ineligible and noncomplete. For the national weights (including and excluding City X), the same rules applied, except for the following. For national weights including City X, those from the four cities that were not part of the national sample should have zero weights. For national weights excluding City X, those from those four cities plus City X should have zero weights.

We also looked at summary statistics of the ratio between the follow-up weight and its comparable mother baseline weight (for the city-specific weight, this is done separately by city) to see if there were any extreme values. We printed out the cases with the five highest and five lowest values of this ratio, to look at the weighting adjustment factors for valid explanations as to why these ratios were so high or so low. Table 2.4 shows the final weighting variables and the sums of the weights.[14]

Table 2.5 shows the covariates used for each of the response propensity models, after the stepwise regression procedure, and any other differences across the weights.

TABLE 2.4

FATHER BASELINE AND FOLLOW-UP WEIGHTS SUMMARY

Final Survey Father Weight		Weight Variable Name	Sum of the Weights
National Level Including City X (16 cities)	Baseline Father	f1natwt	1,131,308.36
	1-Year Follow-up	f2natwt	1,131,308.36
	3-Year Follow-up	f3natwt	1,131,308.36
	5-Year Follow-up	f4natwt	1,131,308.36
National Level Excluding City X (15 cities)	Baseline Father	f1natwtx	1,131,308.36
	1-Year Follow-up	f2natwtx	1,131,308.36
	3-Year Follow-up	f3natwtx	1,131,308.36
	5-Year Follow-up	f4natwtx	1,131,308.36
City Level (20 cities)	Baseline Father	f1citywt	347,237.80
	1-Year Follow-up	f3citywt	347,237.81
	3-Year Follow-up	f3citywt	347,237.80
	5-Year Follow-up	f4citywt	347,237.80

[14] Cases without valid survey data (except for cases who were dead/child not living with either parent) are recoded to missing in the final version of the weights so the sums in the data files will not match these sums.

TABLE 2.5

INDEPENDENT VARIABLES IN FINAL PROPENSITY MODELS USED TO FORM CELLS FOR FATHER WEIGHTS[15]

Locatability Adjustment			Response Adjustment for Located		
1-year	3-year	5-year	1-year	3-year	5-year
M1CITY*	M1CITY*	M1CITY*	M1CITY*	M1CITY*	M1CITY*
M1A13+M1A13A	M1A13+M1A13A	M1B8	M1A9 (grpd)	M1A11A	M1A11D
M1B3	M1A15	M1D1D	M1B3	M1A13+M1A13A	M1E4B
M1D2B	M1B3	M1D2A	M1D1A	M1D2E	M1F5
M1D2F	M1B28	M1I1 (grpd)	M1D1C	M1E3C	M1I2A (grpd)
M1E3E	M1D2F	LOC_M3	M1D1D	M1E4A	COMP_M2
M1H3+M1H3A	M1G2 (grpd)	LOC_M4	M1D2B	M1G6	COMP_M4
M1I11	M1G6	LOC_F3	M1F7 (grpd)	M1I1 (grpd)	COMP_F0
M1J4	M1I3 (grpd)		M1G3 (grpd)	M1I11	COMP_F2
	LOC_M2		M1G4 (grpd)	COMP_M3	COMP_F3
	LOC_M3		M1G6	COMP_F0	
	LOC_F2		M1I1 (grpd)	COMP_F2	
			M1I3 (grpd)		
			M1I11		
			M1J3		
			COMP_M2		
			COMP_F0		
*Significant cities for this model:	*Significant cities for this model:	*Significant cities for this model:	*Significant cities for this model:	*Significant cities for this model:	*Significant cities for this model:
Oakland	Oakland	Oakland	Newark	Baltimore	Baltimore
City X	Detroit	Detroit	Corpus Christi	Corpus Christi	Newark
Baltimore	Philadelphia	Indianapolis	New York	Milwaukee	Philadelphia
Detroit	Richmond	Milwaukee	Chicago	New York	New York
Newark	San Jose	Boston		San Jose	San Jose
Philadelphia		Jacksonville		Boston	Boston
San Jose				Jacksonville	
				Pittsburgh	

Note: Cities that were not significant predictors of nonresponse for a particular model were grouped together.

[15] See the appendix for what these baseline variables represent. We imputed missing values of these variables for propensity modeling only.

SECTION 3 – COUPLE WEIGHTS

This memorandum documents the methodology for constructing the weights used for analysis of the survey data provided by mothers and fathers as part of the core data collection effort for the Fragile Families and Child Well-Being Study. These weights would be used when the analyst wants to include data from both the mother and the father for a particular analysis. The target survey population for this study can be summarized as: live births occurring in large cities, by mothers who plan to keep the child, can identify the still-living father, and speak English or Spanish.

The memorandum covers the weights for the follow-up surveys conducted at one year, three years, and five years after baseline. For each of these three time periods, there are three sets of weights: one for analysis at the city level and two for analysis at the national level. We discuss these in detail below. Please see Section 1 for how we created a set of baseline weights that make the sample of births represent all eligible births occurring in large metropolitan areas of the United States during the study period. All the couple weights would then sum up to this baseline total, which represents the population of children eligible for the study at baseline. For each stage of sampling and data collection, we discuss below which units are considered to be eligible respondents. Because we are starting with the final mother and father weights, each stage of weighting the couples follows two basic steps: (1) adjust weights for nonresponse, and (2) poststratify weights to known totals. Issues related to variance estimation will be covered in a separate memorandum.

A. BASELINE WEIGHTS

For analyzing couples at baseline, the father baseline weight (**f1natwt, f1natwtx,** or **f1citywt** depending on the sample/measure of interest) can be used. Please see Section 2 for how we created the father baseline weights.

B. FOLLOW-UP WEIGHTS

The final *mother baseline* weights, which represent the sampled births, serve as the anchor for all the couple follow-up weights at one year, three years, and five years after baseline. Because there is no subsampling at the follow-ups, we concern ourselves mainly with nonresponse adjustments and re-raking to the mother baseline totals.

There was much discussion about how to define the eligible population at each stage of follow-up. The final decision was that a case was ineligible at follow-up only if the child associated with the sampled birth had died. Because of the rarity of this type of situation, we considered all sample members to have known eligibility status; that is, even if we could not locate a sample member, we assumed the case was still eligible unless we learned otherwise. (We also treated as ineligible those cases that were released to the sample in error, because they were duplicates.)

Other types of situations were those in which a survey was not intended to be conducted, according to study protocol, because the questions did not apply. These included cases in which the child was adopted, neither parent had custody of the child, or one of the parents had died. It also included cases where the father was unknown, the father denied paternity, the father was not told of the pregnancy, or the DNA test indicated this was not the father. While no survey was completed (or only a few questions of the survey were answered) for these situations, we decided that we would still consider these cases part of the target population, and that these situations (custody and parental death) were a type of outcome. Because we obtained all the needed information from the case as part of the disposition code, these cases were considered to be completes for weighting purposes. Princeton University constructed separate indicators (**cm*W*samp/cf*W*samp** flags, where *W* is the wave) to indicate whether some or all of a questionnaire was completed, among those considered to be part of the eligible completes within the sample. All other final dispositions were considered to be eligible noncompletes, subclassified as located or unlocated. Table 3.1 shows how the mother and father disposition codes were classified for weighting purposes, Table 3.2 shows how we combined these classifications for the couple weight, and Table 3.3 shows how the classification variables were assigned values.

TABLE 3.1

CLASSIFICATION OF FINAL MOTHER AND FATHER DISPOSITION CODES FOR COUPLE WEIGHTING PURPOSES

Classification for Weighting Purposes		Final Disposition Code
Eligible Complete	With Survey Data	01-Complete callout 05-HC Comp-field 06-HC Comp-field-incarcerated 07-Comp-field call in 08-Comp-incarcerated call in
	Without Survey Data[16]	40-Deceased 41-DNA not dad 42-Child adopted 44-Father unknown 45-Father not told of pregnancy 46-Neither parent had custody 47-Father denies paternity
Ineligible[17]		43-Child deceased 48-Duplicate
Eligible Noncomplete	Located	20-Hung up during intro 21-Refusal 28-Mother refused info on dad 30-Language barrier 31-Ill/other barrier 32-Incarcerated 33-Unavailable during field period 37-No prior interview – no field 39-Other eligible 49-Other ineligible 54-Moved out of state 65-Maximum calls 66-Case retired 67-Marr dad/no field eff 68-Mom no contact with dad 69-Out of area/no effort
	Unlocated	59-Cannot locate

[16] There were some cases that started to respond to the survey, but whose survey responses early in the questionnaire indicated that the child was adopted or that neither parent had custody. At that point, the survey was terminated. The final disposition codes for these cases were changed from “complete” to “child adopted” or “neither parent had custody.”

[17] If a previous round of the survey indicated ineligibility (child deceased or duplicate), or the spouse’s current survey indicated ineligibility, then the current round was classified as ineligible, regardless of the current disposition code.

TABLE 3.2

COMBINING MOTHER AND FATHER CLASSIFICATIONS

Mother Classification	Father Classification	Couple Classification
Eligible Complete	Eligible Complete	Eligible Complete
Eligible Complete	Ineligible	Ineligible
Eligible Complete	Eligible Noncomplete Located	Eligible Noncomplete Located
Eligible Complete	Eligible Noncomplete Unlocated	Eligible Noncomplete Located
Ineligible	Eligible Complete	Ineligible
Ineligible	Ineligible	Ineligible
Ineligible	Eligible Noncomplete Located	Ineligible
Ineligible	Eligible Noncomplete Unlocated	Ineligible
Eligible Noncomplete Located	Eligible Complete	Eligible Noncomplete Located
Eligible Noncomplete Located	Ineligible	Ineligible
Eligible Noncomplete Located	Eligible Noncomplete Located	Eligible Noncomplete Located
Eligible Noncomplete Located	Eligible Noncomplete Unlocated	Eligible Noncomplete Located
Eligible Noncomplete Unlocated	Eligible Complete	Eligible Noncomplete Located
Eligible Noncomplete Unlocated	Ineligible	Ineligible
Eligible Noncomplete Unlocated	Eligible Noncomplete Located	Eligible Noncomplete Located
Eligible Noncomplete Unlocated	Eligible Noncomplete Unlocated	Eligible Noncomplete Unlocated

If either member of the couple was classified as “ineligible,” then the couple was classified as “ineligible.” If both members of the couple were classified as “eligible completes,” the couple was classified as an “eligible complete.” Otherwise, the couple was classified as “eligible noncomplete.” To be classified as “eligible noncomplete unlocated,” both members of the couple had to be unlocatable.

We adjusted the weights of the eligible completes to account for those of the eligible noncompletes in two stages. Each follow-up weight started with the final poststratified mother baseline weight (national, national without City X, or city, as appropriate). First we adjusted for unlocatability; that is, we adjusted the initial weights for all the eligible located cases upward to account for those of the eligible unlocated cases. Then we adjusted for nonresponse among the located; that is, these adjusted weights for the eligible located completes were further adjusted upward to account for those of the eligible located noncompletes. These adjustments were made within weighting cells. (The formation of these weighting cells is described below.) Each cell comprised sample members who had similar response propensities. After the cells were formed, the two sets of adjustments were made separately for each of the two national weights and the city weight.

TABLE 3.3

VALUES ASSIGNED TO CLASSIFICATION VARIABLES FOR COUPLE WEIGHTS

	LOC_*Ci* Located Status	ELIG_*Ci* Eligibility Status	ELIGD_*Ci* Eligibility Determined	COMP_*Ci* Completion Status
Eligible Complete	1	1	1	1
Ineligible	1	3	1	2
Located Eligible Noncomplete	0	1	1	.
Unlocated Eligible Noncomplete	1	1	1	2

Note: Subscript *i* in the variable names is equal to 2 for the one-year follow-up, equal to 3 for the three-year follow-up, and equal to 4 for the five-year follow-up.

To do these adjustments, we must form weighting cells within which to make the adjustments. We first modeled the propensity separately for the two types of nonresponse among eligibles: (1) unlocated, and (2) noncomplete among located. We determined a set of covariates to use as candidates for these models from among baseline survey variables, which are available for both respondents and nonrespondents in these follow-up surveys. These were baseline variables that we thought would predict response propensity. We temporarily imputed values if values were missing. We developed two separate unweighted stepwise logistic regression models to predict the two types of nonresponse. (The stepwise parameters were that the significance level for entry into the model was .15 and the significance level for staying was .20.) One of the baseline covariates was a city indicator. We separately examined whether any particular cities were significant predictors of each type of nonresponse. If so, we included them as possible covariates for the stepwise logistic regression models. After convergence of the final model, we used the propensity scores to form deciles for the national weights, and quintiles within city for the city-level weights. We used these deciles and quintiles to form the weighting cells for the nonresponse adjustments. Note that the ineligible sample members were excluded from these nonresponse adjustments and simply retained their initial weight from baseline. After the two national weights and the city weight had been adjusted for the two types of nonresponse, we brought the ineligible weights back in.

Raking. See the Section I for more information about raking in general and some specifics for this study. In this study, the variables used for the raking process are given in Table 3.4. Even though the adjustment was done within individual raking cells, the raking process requires only known marginal population totals for a single variable, rather than totals for individual cells in which multiple variables are crossed. We therefore do not need to worry about empty or very small cells formed by crossing all the variables used in the poststratification process.

TABLE 3.4

LIST OF RAKING VARIABLES FOR COUPLE WEIGHTS

Variable Name	Description	Levels
Marital Status	Mother's Marital Status	2
Education	Mother's Education Level	5
Race/Ethnicity	Mother's Race/Ethnicity	4
Age	Mother's Age	7

The raking adjustment would be calculated as the synthetic estimate (using CDC Natality data) of the number of births by marital status in the appropriate city (or nationally) and year, divided by the sum of the nonresponse-adjusted weights within raking cell. This adjustment is then applied to the nonresponse-adjusted weight for each completed case.

Trimming. After raking the weights as described above, we trimmed them to remove any outliers that may have occurred due to a large adjustment factor or combination of factors. For each of the three types of weights, we determined the mean *(M)* and the standard deviation *(S)* of each weight, by marital status *(i)*. We set the trim value for marital status *i* to $M_i + 4S_i$. That is, we considered any weights that were more than four standard deviations higher than the mean weight value to be outliers and trimmed them to that maximum value. After trimming, we re-raked the weights.

City-Level Weights. The couple city-level weight at time *t* (**c1citywt**) was developed to provide users of the couple follow-up survey data with final survey weights for analyses within individual cities. Using the methods explained in the previous section, we adjusted/raked these weights so that they are consistent with total population counts of births in the large U.S. cities in this study, based on CDC data.

National-Level Weights. The national-level weights are the final couple follow-up survey weights attached to individual births for analyses that pool records for the 16 national-sample cities within the sample. The analysis generalizes to births occurring in the 77 large cities defined as the Fragile Families population. The weights were developed based on national weights (computed in the earlier steps), which were in turn raked to total (population) birth counts in the 77 cities based on CDC data.

We computed two sets of couple follow-up national-level weights at each time point: one based on all 16 of the national-sample cities in the sample, with all 77 cities as the population being targeted (**c1natwt**), and the other based on only 15 cities (City X is excluded) in the sample, with all 77 cities as the population being targeted (**c1natwtx**).[18]

Checking. After each set of follow-up weights was produced, we checked them along several fronts. If we included the ineligible cases (which have positive weights), the sum of the follow-up weight should have been equal to the sum of the corresponding mother baseline

[18] National analyses using **c1natwtx** should include only samples in the 15 cities, while analyses using **c1natwt** should include all records in the 16 national-sample cities.

weight (national, national without City X, and city). We then classified the case according to its city; that is, whether the city was City X, or part of the national sample, or not part of the national sample. We crossed that classification with the eligibility status, locatability status, and completion status, and then checked whether the weight was appropriately missing or had a positive value. The city weights should have all had a positive value if the case was (1) eligible and located and complete, or (2) ineligible and noncomplete. For the national weights (including and excluding City X), the same rules applied, except for the following. For national weights including City X, those from the four cities that were not part of the national sample should have zero weights. For national weights excluding City X, those from those four cities plus City X should have zero weights.

We also looked at summary statistics of the ratio between the couple follow-up weight and its comparable mother baseline weight (for the city-specific weight, this is done separately by city) to see if there were any extreme values. We printed out the five highest lowest values of this ratio to look at the weighting adjustment factors for valid explanations as to why these ratios were so high/low. Table 3.5 shows the final weighting variables and the sums of the weights.[19]

Table 3.6 shows the covariates used for each of the response propensity models, after the stepwise regression procedure, and any other differences across the weights.

TABLE 3.5
COUPLE BASELINE AND FOLLOW-UP WEIGHTS SUMMARY

Final Survey Couple Weight		Weight Variable Name	Sum of the Weights
National Level Including City X (16 cities)	Baseline Couple	c1natwt (same as f1natwt)	1,131,308.36
	1-Year Follow-up	c2natwt	1,131,308.36
	3-Year Follow-up	c3natwt	1,131,308.36
	5-Year Follow-up	c4natwt	1,131,308.36
National Level Excluding City X (15 cities)	Baseline Couple	c1natwtx (same as f1natwtx)	1,131,308.36
	1-Year Follow-up	c2natwtx	1,131,308.36
	3-Year Follow-up	c3natwtx	1,131,308.36
	5-Year Follow-up	c4natwtx	1,131,308.36
City Level (20 cities)	Baseline Couple	c1citywt (same as f1citywt)	347,237.90
	1-Year Follow-up	c2citywt	347,237.80
	3-Year Follow-up	c3citywt	347,237.80
	5-Year Follow-up	c4citywt	347,237.80

Reminder: For analyzing couple data at baseline, father weights can be used as all mothers responded to the baseline survey and the weight only needs to adjust for father non-response.

[19] Cases without valid survey data (except for cases who were dead/child not living with either parent) are recoded to missing in the final version of the weights so the sums in the data files will not match these sums.

TABLE 3.6

INDEPENDENT VARIABLES IN FINAL PROPENSITY MODELS USED TO FORM CELLS[20]

Locatability Adjustment			Response Adjustment for Located		
1-Year	3-Year	5-Year	1-Year	3-Year	5-Year
CM1AGE	CM1AGE	M1A9 (grpd)	M1CITY*	M1CITY*	M1CITY*
M1A11B	M1A11B	M1A11B	M1A9 (grpd)	M1A11B	M1A9 (grpd)
M1A13+M1A13A	M1A13+M1A13A	M1A11C	M1A13+M1A13A	M1D2A	M1A11D
M1B8	M1A15	M1D1B	M1A15	M1D2E	M1D1B
M1D1A	M1D1B	M1D1E	M1B3	M1E3C	M1D1D
M1D1F	M1D1F	M1D2C	M1D1C	M1E4A	M1D2B
M1D2A	M1D2C	M1D2F	M1D2B	M1G1	M1D2F
M1D2E	M1D2D	M1E4A	M1E4C	M1G3 (grpd)	M1E3E
M1E3E	M1D2E	M1F6	M1F7 (grpd)	M1I1 (grpd)	M1F5
M1E4A	M1D2F	M1G3	M1G4 (grpd)	M1J5	M1G2 (grpd)
M1F3	M1E4C	M1G4 (grpd)	M1I2A (grpd)	COMP_F0	M1H3+M1H3A
M1I11	M1F6	M1G6	M1I3 (grpd)	COMP_C2	M1I1 (grpd)
M1J5	M1G6	M1H3+M1H3A	M1I11		COMP_F0
COMP_F0	M1H3+M1H3A	M1J4	COMP_F0		COMP_C2
	M1I1 (grpd)	LOC_C2			COMP_C3
	M1I2A (grpd)	LOC_C3			
	LOC_C2				
			*Significant cities for this model:	*Significant cities for this model:	*Significant cities for this model:
			Oakland	Oakland	Oakland
			Detroit	Baltimore	Baltimore
			Newark	Milwaukee	Philadelphia
			Philadelphia	New York	Indianapolis
			Corpus Christi	San Jose	New York
			New York	Jacksonville	San Jose
			Chicago	Pittsburgh	Boston
			Toledo		

Note: Cities that were not significant predictors of nonresponse for a particular model were grouped together.

[20] See the appendix for what these baseline variables represent. We imputed missing values of these variables for propensity modeling only.

SECTION 4 - VARIANCE ESTIMATION[21]

This section describes the methods used for constructing replicate weights for the public data (which do not contain the strata and PSU variables, except through a restricted use contract) and instructions on how to use these weights in Sudaan and Stata.

The Fragile Families study used a multistage design[22] and estimates can be made at the national or city level. There were 20 cities in the study, of which 16 are used for national estimates, and an average of 3.6 sampled and participating hospitals per city.

Because of the complex sample design, it is important to use specialized techniques when calculating the variance of estimates arising from the data. One technique is the Taylor Series approach, which is available in SUDAAN statistical software and in specialized survey procedures in SAS and Stata. These specialized survey procedures should be used when estimating variance using the strata and PSU variables. Another set of techniques involves replication procedures and the creation of a set of replicate weights. These techniques should be used when using the public data files that do not contain the geographic identifiers needed to construct the strata and PSU variables.

Preparation of Replicate Weights for Public Use Estimation

There are several techniques for creating replicate weights, all of which essentially randomly exclude certain sample members, reweight to account for the excluded sample members, calculate a new estimate based on the subsample, and then calculate the variance across a series of these subsamples. Methods of estimating standard errors that rely on Taylor Series approximations require specification of design parameters such as sampling strata and PSUs. There was concern that even knowing the PSU (city) could potentially compromise the anonymity of some sample members. Therefore, we created a set of replicate weights so that the Taylor Series approximations would not be necessary to calculate variances when using the public data. Jackknife estimation, a replication technique, will allow for the estimation of sampling errors while masking the cities and hospitals with which sample members were associated. Jackknife estimation requires the creation of replicate weights.

We employed a replication method—the random groups approach, discussed in Chapter 2 of Wolter (1985).[23] The excluded random groups were formed by taking random subsamples of nearly equal size such that, for each level of estimate (national or city), each PSU had the same representation in each random group. This was accomplished by selecting the sample for each random group as a stratified random subsample of the whole sample, using the PSUs as explicit strata. These subsamples were selected so that no case could appear in more than one excluded random group. For national estimates, the PSU was the city; for city-level estimates the PSU was the hospital. Cities that are not considered part of the sample for national estimates were excluded from the national level random groups; however, city-level random groups were formed for these cities. We formed 33 random groups for national estimates and 10 for each

[21] Adapted from a memorandum written by John Hall to Princeton University

[22] A description of the sample design can be found in *Children and Youth Services Review*, vol. 23, nos. 4/5, 2001, pp. 303-26, available at [www.fragilefamilies.Princeton.edu/surveys/Reichman_et_al_2001.pdf].

[23] Wolter, KM (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag.

city-level estimate. The number of random groups was based on having approximately 100 cases in each random group for national estimates (each random group being about 1/33 of the sample size) and anywhere from 9 to 42 cases in each random group for city-specific estimates (each random group being about 1/10 of the sample size for the city).

For estimation, we created a number of replicate weights equal to the number of random groups. For each replicate α, the weights of the α^{th} random group were set to zero, while the weights for those not in that random group were adjusted by (1) weighting them up to the full sample by hospital and marital status, and (2) raking and trimming these weights using the same program used for raking and trimming the base weights for the entire sample.

The resulting weights are displayed below.

TABLE 4.1

WEIGHTING VARIABLE NAMES FOR FRAGILE FAMILIES CORE TELEPHONE SURVEY

Wave	**National level weights**[i]		**City level weights**	
	Basic weight	**Replicate weights**	**Basic weight**	**Replicate weights**
Baseline	m1natwt f1natwt c1natwt	m1natwt_rep1-m1natwt_rep33 f1natwt_rep1-f1natwt_rep33 c1natwt_rep1-c1natwt_rep33	m1citywt f1citywt c1citywt	m1citywt_rep1- m1citywt_rep10 f1citywt_rep1- f1citywt_rep10 c1citywt_rep1- c1citywt_rep10
One-Year	m2natwt f2natwt c2natwt	m2natwt_rep1-m2natwt_rep33 f2natwt_rep1-f2natwt_rep33 c2natwt_rep1-c2natwt_rep33	m2citywt f2citywt c2citywt	m2citywt_rep1- m2citywt_rep10 f2citywt_rep1- f2citywt_rep10 c2citywt_rep1- c2citywt_rep10
Three-year	m3natwt f3natwt c3natwt	m3natwt_rep1-m3natwt_rep33 f3natwt_rep1-f3natwt_rep33 c3natwt_rep1-c3natwt_rep33	m3citywt f3citywt c3citywt	m3citywt_rep1- m3citywt_rep10 f3citywt_rep1- f3citywt_rep10 c3citywt_rep1- c3citywt_rep10
Five-year	m4natwt f4natwt c4natwt	m4natwt_rep1-m4natwt_rep33 f4natwt_rep1-f4natwt_rep33 c4natwt_rep1-c4natwt_rep33	m4citywt f4citywt c4citywt	m4citywt_rep1- m4citywt_rep10 f4citywt_rep1- f4citywt_rep10 c4citywt_rep1- c4citywt_rep10

[i] There is also an alternate set of national sample weights with the suffix "x" (e.g, **m1natwtx**). Applying these weights makes the data from 15 of the cities in the sample representative of births occurring in large U.S. cities. This weight achieves the same goal as the primary national sample weight (described above) but drops one of the cities that has a high rate of questions "not asked" (denoted by -5 in the data), particularly at the one-year follow-up, because of changes to the survey instruments between fielding the first two cities and the remaining 18 cities. For example, this weight could be used if you wanted to analyze responses in the following variables… m2b11, m2b11a, m2b11b, m2b21, m2c10, m2d4, m2d4a, etc. To identify variables when the alternative national weights may be necessary, see the annotated questionnaires for notations "18-cities only" or see data for questions with high percent of "-5" responses.

Analysis Using Replicate Weights

The variances for random groups and jackknife estimates are almost identical (see Wolter 1985, chapters 2 and 4), so software that computes standard errors using the jackknife method will produce appropriate estimates of sampling error for the FF data.

In SUDAAN, the statement "design=jackknife" invokes the jackknife estimation procedure. The analysis file contains a basic weight; this weight must be specified in the weight statement. The replicate weights should be listed as part of the "jackwgts" statement. The sample code below produces estimates using "Replicate Weight Jackknife" procedures in SUDAAN (RTI 2004, Chapter 3).[24] As mentioned above, the analysis file contains a basic weight; this weight must be specified in the weight statement (e.g., "WEIGHT m1natwt;"). The subgroup and levels statements are required for PROC CROSSTAB and are explained in the SUDAAN documentation. The replicate weights are listed as part of the "jackwgts" statement.

Sample SUDAAN code looks like this…

```
proc crosstab data=filename filetype=sas design=jackknife deft4;
weight fill in weight;   (e.g. m1natwt)
tables [some categorical variable]*[another categorical variable];
subgroup [some categorical variable];
levels [number of categories of categorical variable];
jackwgts fill in replicate weights (e.g. m1natwt_rep*) / adjjack = 1;
output colper serow secol deffrow deffcol/filename=out.filename_wA filetype=sas replace;
    run;
```

Sample Stata code would look like…

```
svyset [pw=BASICWEIGHT], jkrw(REPLICATES, multiplier(1)) vce(jack) mse
```

… where *BASICWEIGHT* and *REPLICATES* are replaced with the relevant (wave/sample) weights for your analyses. Then users would put the "svy:" command before the command they wanted to run (e.g. logistic or regress). See the document "*Fragile Families & Child Wellbeing Study: A Brief Guide to Using the Mother, Father, and Couple Weights for Core Telephone Surveys Waves 1-4*" for more examples using Stata.

Analysis using Taylor Series Approximation with Stratum and PSU (for users with geographic identifiers)

As alternative to using the replicate weights, those using restricted files with information on design parameters (stratum and PSU) could use the basic weight appropriate for the data and conduct analyses that rely on Taylor Series approximations to estimate sampling error. It is not clear that estimates of standard errors by this method will be much better (at least for cross-sectional estimates) than those using the replicate weights prepared for the public use files. While the estimates using replicate weights do not allow for treatment of PSUs as certainty

[24] Research Triangle Institute (2004). *SUDAAN Language Manual, Release 9.0* Research Triangle Park, NC: Research Triangle Institute.

selections, the fact that there are more random groups than there are cities may provide estimates of sampling error that are as stable as those using Taylor Series approximations.

In conducting the Taylor Series method in SUDAAN, it would be appropriate to specify selection with replacement (design = wr). In both SUDAAN and Stata, this is the default.

For city-specific analysis of births, the PSU would be the hospital (m1hospid), and the stratum would be the city (m1city -- which only comes into play if more than one city were to be included in the analysis). For convenience, these variables are recreated for you as **citystratum** and **citypsu**.

For national estimates: for cities selected with certainty, the stratum is the city (m1city), and the PSU is the hospital (m1hospid); for non-certainty cities, the stratum is a variation on the one used for selecting cities (STRAT_2) and the PSU is the city. Note that you must use STRAT_2 instead of STRAT, because STRAT_2 treats San Jose and Baltimore as if they had been a priori certainty selections. Because the STRAT_2 variable contains 5 strata with singleton PSUs, we have created an alternative stratum variable that collapses these five cities into one pseudo-stratum.[25] The variables are stratum = **natstratum** and the primary sampling unit (PSU) = **natpsu**. Users can consider alternate strategies.

Sample SUDAAN code using the Taylor Series methodology would look something like this for national estimates:

```
PROC CROSSTAB DATA=filename DESIGN=WR DEFF;
NEST [appropriate stratum and psu variables];    e.g. natstratum natpsu
WEIGHT [appropriate weight for wave/sample];  e.g. m2natwt
SUBPOPN [appropriate sample flag for wave/sample];  e.g. m2natsm
SUBGROUP [some categorical variable];
LEVELS [number of categories of categorical variable];
TABLES [some categorical variable]*[another categorical variable];
RTITLE '[title]';
```

The syntax of the svyset statement in Stata you should use is …

svyset [pw=*BASICWEIGHT*], strata(*STRATID)* psu*(PSUID)* clear

… where *BASICWEIGHT* is replaced with the relevant (wave) weights for your analyses and *STRATID* and *PSUID* are replaced with the appropriate strata/PSU for the sample you are analyzing (national or city). You also need to include the subpop command with the appropriate sample flag…

svy, subpop(*SAMPLEFLAG*):

…before the command you are running.

---

25 In the **natstratum** variable we collapsed the following five cities into one stratum as one strategy data users can consider for dealing with singleton PSUs. Data users who have access to city identifiers can consider alternate strategies as well.

To limit to your analytic sample, you would use the SUBPOPN statement in SUDAAN or the subpop option in Stata to specify the sample to be included. Note that using the subpopulation specification is preferred over subsetting the file to the domain of interest. By including the full file and specifying the subpopulation, your statistical package will make full use of the design when determining the degrees of freedom.

APPENDIX

SELECTED BASELINE VARIABLE NAMES AND DESCRIPTIONS

		Variable used in propensity models for		
Variable Name	Variable Label	Mother	Father	Couple
cm1age	Constructed—mother's age (years)	X		X
m1city	City of interview	X	X	X
m1a4	Is respondent married to baby's father	X		
m1a9	Do you feel ready to go home or would you rather stay longer?	X	X	X
m1a11a	Will the baby live with mother?	X	X	
m1a11b	Will the baby live with father?	X		X
m1a11c	Will baby live with other relatives?	X		X
m1a11d	Will baby live with non-relatives?		X	X
m1a13	Did you visit a doctor/other health care professional to check on the pregnancy?	X	X	X
m1a13a	In which month of pregnancy did you first see doctor/other health care provider?	X	X	X
m1a15	How are you paying for the baby's birth?	X	X	X
m1b2	Int chk: Is respondent married? (primary marriage variable)	X		
m1b3	Which statement best describes your current relationship with baby's father?	X	X	X
m1b8	Are you and bf living together now?	X	X	X
m1b27	When you found out you were pregnant, did you think about having an abortion?	X		
m1b28	Did bf suggest that you have an abortion?		X	
m1d1a	The main advantage of marriage is financial security?		X	X
m1d1b	There are more advantages to being single than to being married?	X		X
m1d1c	A single m can bring up a child as well as a married couple?		X	X
m1d1d	It is better for a couple to get married than just live together?	X	X	X
m1d1e	It is better for children if their parents are married?	X		X
m1d1f	Living together is just the same as being married?			X
m1d2a	How imp for successful marriage, have same friends?	X	X	X
m1d2b	How imp for successful marriage, husband has	X	X	X

		Variable used in propensity models for		
Variable Name	Variable Label	Mother	Father	Couple
	steady job?			
m1d2c	How imp for successful marriage, wife has steady job?	X		X
m1d2d	How imp for successful marriage, both same race/ethnicity?	X		X
m1d2e	How imp for successful marriage, have good sex?	X	X	X
m1d2f	How imp for successful marriage, share same religion?	X	X	X
m1e3a	During your preg, did you receive financial support from anyone besides birth father?	X		
m1e3c	During preg, did you receive a place to live?		X	X
m1e3e	During preg, did you receive child care?	X	X	X
m1e4a	Next yr, would someone in family loan you $200?	X	X	X
m1e4b	Next yr, would someone in family give place to live?		X	
m1e4c	Next yr, would someone help you with babysitting/child care?	X		X
m1f2	Is the home/apartment where you currently reside owned/rented?	X		
m1f3	Do you live in a public housing project?	X		X
m1f4	Is the fed/state/local govt helping to pay for your rent?	X		
m1f5	How safe are the streets around your home at night?		X	X
m1f6	About how often do you attend religious services?	X		X
m1f7	What is your religious preference?		X	X
m1g1	How is your health?	X		X
m1g2	During the preg, how often did you drink alcohol?	X	X	X
m1g3	During the preg, how often did you use drugs?	X	X	X
m1g4	During the preg, how many cigarettes did you smoke?	X	X	X
m1g6	Have you ever sought help or been treated for drug/alc problems?		X	X
m1h3	What is your race?	X	X	X
m1h3a	Are you of Hispanic or Latino origin or descent?	X	X	X
m1i1	What is the highest grade/years of school that you have completed?	X	X	X
m1i2a	When you last worked, how many hours per week did you work?		X	X

		Variable used in propensity models for		
Variable Name	Variable Label	Mother	Father	Couple
m1i3	What is the highest grade/years of school that bf has completed?	X	X	X
m1i11	Where does bf live most of the time?	X	X	X
m1j3	What was your total household income before taxes in the past 12 months?	X	X	
m1j4	At the end of the month how much money left over do you usually have?		X	X
m1j5	Do you own a car?	X		X