

## Appendix WW

### Details of Imputation, Calculation of the Survey Weights, and Nonresponse Bias Analysis

#### Imputation

Imputation will be used to adjust for item nonresponse, i.e., missing data for particular items among those who respond to a given wave. By using imputation to “plug holes” due to item nonresponse, we mitigate issues analysts would encounter in trying to analyze data with “swiss cheese” patterns of missingness. As with weighting, a carefully designed imputation procedure will reduce bias due to nonresponse (in this case, item nonresponse).

For imputation, a cyclical n-partition hot deck (an approach analogous to the Gibbs sampler but using the hot deck to generate the imputations) will be used. (See Judkins 1997; Judkins et al. 2007; Judkins, Piesse, and Krenzke 2008; and Krenzke and Judkins 2008.) The cyclical n-partition hot deck relies primarily on the hot deck method of imputation, beginning with a simple hot deck to initialize the process, and iterating with successive rounds of hot deck imputation until convergence of the imputation model is reached. This approach is designed to preserve multivariate distributions; in implementing the approach, care will be taken to ensure that imputations maintain skip patterns and adhere to constraints. After imputation, the same analytic edits that we ran on the raw data will be run again on the imputed data.

## **Calculation of the Survey Weights**

The weighting adjustments will be fairly standard. The approach entails giving a zero weight to the nonresponding case and redistributing the base weight of the nonrespondent to responding but otherwise similar cases. This process is done within nonresponse adjustment “cells.” The approach we propose to use to form the cells for nonresponse adjustment uses a class of procedures known as “doubly robust” adjustments. In contrast to traditional approaches for forming nonresponse adjustment cells, these procedures place greater emphasis on the modeling of critical outcomes in the development of cells and somewhat reduced emphasis on the modeling of nonresponse propensity. In a survey with many outcomes, the challenge is determining the key outcomes to use in this modeling exercise. For WIC ITFPS-2, we propose to develop a binary indicator at each wave for whether the mother is following recommended feeding practices for the age of the infant. We will then model this in terms of data from prior waves to obtain a set of cells that vary in maternal conformance to recommended feeding practices. We will then cross these cells with the cells defined more traditionally to predict nonresponse propensity.

The key to effective nonresponse adjustments is the availability of good auxiliary variables to be used in the adjustment. The adjustment for those who initially consent but do not respond to the initial (prenatal or 1-month) interview is most limited in this regard. We should be able to use local administrative data on food package/voucher receipt to adjust for

nonresponse to the initial interview. This will require us to give a list of recruited participants at the site to a local clerk and for the clerk to then keep a record of package/voucher usage over the next year or so – enough time for all the pregnant enrollees to have given birth. Of course, depending on the sophistication of the local office, we can also do an electronic merge of their voucher records with our sample. Also, we will build models of attrition at each wave based on the data collected to date. Various modeling methods could be considered, and these methods have been found to work approximately equally well (Folsom and Witt, 1994; Rizzo, Kalton, and Brick, 1996; Judkins, et al 2005); the real question is which variables to allow into the modeling and how to deal with missing data in the early wave data. The variables under consideration in this modeling process will include variables available for attritors from earlier waves.

In many surveys, one step (generally the final step) in the sequence of weighting adjustments is to calibrate the weights (e.g., using poststratification or raking adjustments) to control totals from trusted sources, such as census totals or estimates from administrative record systems or larger surveys. In this case, no such trusted source exists, so this calibration step will not be possible.

Variability in the weights is a concern because highly variable weights reduce the precision of survey estimates, and there is the potential for the cases with large outlying weights to have undue influence on estimates. If there are isolated incidences of sites with weights much larger

than the mean, we will consider trimming the weights to avoid the situation where the results from such an office dominate the national estimates on a weighted basis.

Table B2.3 shows a planned set of weights. In addition to the cross-sectional weight for each wave, we will create longitudinal weights for analyses that require analysis of linked data across waves. We note that simple change estimates (as in the percent still breastfeeding) do not require linked data. These change estimates will be prepared by forming point estimates for each wave with the wave-specific cross-sectional weight and then subtracting the two estimates at the macro level to get estimates of net change. In creating the weights, we will consider the use of variables available from earlier waves to adjust for nonrespondents who completed some waves (but not enough waves to constitute “response”).

Table B2.3. Weights to be prepared and delivered

Weight name	Core only or combined ?	Positive for respondents at which waves?	Additional notes
PrenatalWgt	Core only	Prenatal	
Month1CoreWgt	Core only	1-mo	Only prenatally recruited infants and infants recruited postnatally within the window for the 1-mo interview
Month1CombWgt	Combined	1-mo	Only prenatally recruited infants and infants recruited postnatally within the window for the 1-mo interview
Month3CoreWgt	Core only	3-mo	
Month5CoreWgt	Core only	5-mo	
Month7CoreWgt	Core only	7-mo	
Month7CombWgt	Combined	7-mo	
Month9CoreWgt	Core only	9-mo	
Month11CoreWgt	Core only	11-mo	
Month13CoreWgt	Core only	13-mo	
Month13CombWgt	Combined	13-mo	
Month15CoreWgt	Core only	15-mo	
Month18CoreWgt	Core only	18-mo	
Month24CoreWgt	Core only	24-mo	
Month24CombWgt	Combined	24-mo	
HazardModelCoreWgt	Core only	Prenatal + 1-mo or 3-mo if recruited postnatal	Good for modeling hazard of weaning & introduction of various foods; good for modeling of BF initiation
HazardModelCombWgt	Combined	At least one postnatal interview	Good for modeling hazard of weaning & introduction of various foods. Earliest weight that uses entire sample. Larger sample size than HazardModelCoreWgt but can't be used in conjunction

---

w/prenatal data.

---

InfantCoreLongWgt	Core only	Responded every wave from birth through 13 mo	Good for growth curve modeling of calories or other variables that are measured each wave.
-------------------	-----------	---	--

Table B2.3. Weights to be prepared and delivered (Continued)

Weight name	Core only or combined ?	Positive for respondents at which waves?	Additional notes
ToddCoreLongWgt	Core only	Responded every wave from birth forward	Good for growth curve modeling of calories or other variables that are measured each wave. No plans to use in our analysis, but would be expected by many users on a RUF.
CritWaveLongWgt	Combined	1/3, 7, 13, 24	If prenatally recruited or recruited postnatally within the 1-mo interview window, responded at mos 1, 7, 13, and 24. If postnatally recruited after the 1-mo interview window, responded at mos 3, 7, 13, and 24. Good for growth curve modeling with procedures that cannot handle missed waves.

## **Nonresponse Bias Analysis**

To the extent that respondents are systematically different from the population as a whole with respect to characteristics used in an analysis, the potential for nonresponse bias exists. Statistical methods used to compensate for missing data (weighting and imputation) aim to reduce nonresponse bias. Since there is generally no way to directly measure the difference in key survey characteristics between respondents and the population as a whole, various methods have been developed that aim to assess the potential for nonresponse bias.

One approach we will use is to examine bivariate cross tabulations of data from one wave by response status at a followup wave to check for evidence of nonresponse bias at followup. Since there will be eight waves of followup on the core sample after the first interview for infants recruited after birth, and ten waves of followup on the core sample for those required prenatally, there will be many possible cross-tabs that could be run. By the 24-month interview, there will be thousands of measurements from prior waves that could be used to check for nonresponse bias at the 24-month interview. Obviously the scope of these tabulations could quickly become unmanageable. We will identify a few key variables from early waves to use as benchmarks for nonresponse bias analyses.

As discussed above, the weighting class adjustments for nonresponse aim to reduce nonresponse bias. Thus, while the subgroup response rate analysis described above may be useful in identifying the

*potential* for nonresponse bias due to varying response propensities among key subgroups, this nonresponse bias may be mitigated through the adjustments for nonresponse. To examine this, we will compare unadjusted estimates (i.e., computed using weights that do not include the adjustment for nonresponse to the particular wave) to adjusted estimates.

With a longitudinal study such as WIC IFTPS-2, another technique that can be used is to compare prior-wave estimates for key statistics for respondents to the given wave to the corresponding prior-wave estimates computed using the full set of prior-wave respondents.

Another method that could be considered is benchmarking estimates from WIC IFTPS-2 to estimates from other sources, provided such external estimates are available. Although benchmarking to external estimates is a method commonly included in a repertoire of nonresponse bias analysis techniques, it is recognized that this approach does not allow for isolation of bias due to nonresponse. Besides nonresponse bias, differences between the survey estimates and external estimates might be attributable to temporal differences, differences in survey populations or survey measures, or other sources of error such as coverage bias.