Methodology for Constructing Child Weights for Wave 3-5 Fragile Families and Child Wellbeing Study

Yajuan Si and Andrew Gelman

1 Overview

We construct Year 9, Year 5 and Year 3 child weights for the home visit samples, as collaborative studies for the Fragile Families (FF) and Child Wellbeing Study. For each year, there will be two sets of national weights and one set of city weights. The national child weights are based on 16 cities to represent national samples; the city child weights are constructed to represent the 20 city samples. The two national weights differ from each other by whether including City X, which conducts as a pilot study with different questionnaire from the remaining cities. The summary of weighting variable names is shown in Table 1.

Table 1: Weighting variable names for Fragile Families Wave 3-5 home visit survey.

	Basic weight	Replicate weights
National Level	k5natwt	k5natwt_rep1-k5natwt_rep26
	k4natwt	k4natwt_rep1-k4natwt_rep26
	k3natwt	k3natwt_rep1-k3natwt_rep26
National Level (without City X)	k5natwtx	k5natwtx_rep1-k5natwtx_rep23
	k4natwtx	k4natwtx_rep1-k4natwtx_rep23
	k3natwtx	k3natwtx_rep1-k3natwtx_rep23
City Level	k5citywt	k5citywt_rep1-k5citywt_rep72
	k4citywt	k4citywt_rep1-k4citywt_rep72
	k3citywt	k3citywt_rep1-k3citywt_rep72

2 Nine-year follow-up wave

2.1 Overview of the Nine-Year follow-up data collection

The fifth wave FF data collection around focal children's ninth birthdays, was conducted from August 2007 through April 2010. The Nine-Year wave of data collection integrated interviews

with 1) core biological parents, 2) primary caregivers (and in certain circumstances, a non-parental caregiver), 3) "focal" children, and 4) teachers. Home Visits were also conducted and included cognitive tests, in-home observations, a primary caregiver self-administered questionnaire, and saliva sample collection for genetic analysis. Interviewers completed "In-Home Observations" of the home environment following the Home Visit.

This wave of data collection was fielded to allow researchers to answer the following questions: How do children develop over time, and how do family resources influence children's health and development? How do the resources of unmarried parents evolve over time, relative to those of married parents? How do children's genetic endowments interact with their environments to influence their outcomes? How do school environments influence children's social and academic outcomes?

These survey components were typically administered in the following order: In most cases, the primary caregiver (PCG) survey was completed by Computer-Assisted Telephone Interviewing followed by the core biological parent interviews. Home Visits were typically scheduled during the primary caregiver and core biological parent phone interviews. During the Home Visit, a 20-minute interview was administered to the focal child (using Computer-Assisted Personal Interview technology), the primary caregiver completed a self-administered questionnaire, height (focal child only) and weight (focal child and biological mother) measurements were taken, a speech sample was taken from the primary caregiver, and cognitive assessments were conducted with the focal child. Saliva samples were also collected from biological mothers and focal children. Interviewers also collected consent and contact information in order to mail hard-copy interviews to focal children's teachers.

2.2 Child weighting

There is a child interview (n=3377—the indicator variable is *ck5kint*), and various components of the child assessments (n=3392—the indicator variable is *o5oint*) like PPVT, Woodcock Johnson or Digit Span, height/weight, saliva sample, or home visit observations. For the child survey, there could be numerous scenarios for completing a part of the Year 9 child interview or assessments and home visit. In Year 9, we will construct weights for every child who participated in any assessment or was interviewed (n=3393—the indicator variable is *Y9childsamp*).

Table 2 presents the classification for Year 9 biological mother samples and child samples with home visits or not. The Year 9 mother weights are assigned to those who responded in the survey (with survey data, mother died, and child adopted or neither parent has legal custody). For children from these 3596 families, 3295 of them participated the home visit while 301 did not. The 301 non-participants neither were not sampled for the home visit or refused to participate by the mothers. We use inclusion to represent their status—both selected and responded—a unit is included only when it has been selected and also responds. We will adjust for the exclusion for the child samples considering the cases with Year 9 mother weight.

Beside these home visit samples, there are 98 case with in-home observations or child survey data, but without mother weights. The mothers are either unlocated or non-responded.

Table 2: Sample classification for Year 9 child weighting

Classification for weighting				Child	
				No HV	HV
Eligible	located	response (with mom weight)	-7 NA (with survey data)	239	3276
			1 Mother died	14	18
			3 Adopted/Ne parent has legal custody	48	1
		nonresponse	5 Refusal	325	25
			7 Other non-response	313	42
	unlocated		6 Could Not Locate	404	31
Ineligible			2 Child died	46	0
			4 Other Ineligible	116	0

A subset of these children do not live with their mother (either with dad or another non-parental caregiver), and the person who cared for the child most of the time is interviewed. And the home visit activities are conducted at the child's (new) home.

Therefore, our weighting process has two main steps:

1. Starting from cases with Year 9 mother weight, we adjust for the exclusion of the home visit samples.

2. Bring in the home visit samples for which no Year 9 mother weights were assigned but with allocated weights from the nearest previous waves; poststratify the combined home visit samples to match the population totals.

To account for the exclusion, we build a regression model with the binary inclusion indicator as the outcome for the cases with Year 9 mother weight. The covariates in the regression model are collected from mother wave 5 survey variables, listed in Appendix A. The covariates are available for samples both with home visits and without home visits. We did a preliminary selection by excluding the variables with more than 20% item missingness, more than 11 possible values.[1] We filled in the missing items by random draws from the corresponding observed frequency distributions. The predictors after dummy coding are used as covariates.

We used the predicted inclusion propensity scores to form deciles for the national weights, and quintiles within city for the city-level weights. We used these deciles to form the weighting cells for the exclusion adjustments. Therefore, each weighting cell comprises sample members who have similar inclusion 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.

After the exclusion adjustment, we bring back the home visit samples for which no Year 9 mother weights were assigned. We rake the weights to wave 5 mother weight totals. The raking variables include mother's age, education, ethnicity and marital status. See the

[1] we did not include the continuous variables here, but we included city and hospitals as covariates. We will use dummy coding the regression to recode these categorical variables. Only using categorical variables helps implement the R package *glmnet*.

Table 3: Summary of Year 9 national sample weights.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
mother	46.17	70.40	144.30	425.00	404.10	5181.00	2237
child	17.02	90.12	169.50	454.20	434.20	5130.00	2408

frequency distributions of the raking variables in Appendix D, for national and city weights, respectively.[2] Then we trim the large weights and re-rake.

2.2.1 National weighting

We collect the sample dispositions based on the flag *cm5samp* and *Y9childsamp*. We start with mother national weights at wave 5 for these home visit cases. If these units were not assigned wave 5 weights, we move on to incorporate the weights in previous waves sequentially. This results in sample size 2623 (2409 included and 214 excluded cases). The sample size of home visits being representative of national samples is 2490.

We build the logistic regression model under **Lasso** (Friedman *et al.* , 2010) for regularization with the inclusion indicator as the outcome and variables in Appendix A as covariates and use the predicted propensity scores to form deciles for the national weights. After the exclusion adjustment, we bring back the home visits with mother weights from previous waves. We rake the weights to mother wave 5 weight totals. The raking variables include mother's age, education, ethnicity and marital status. We implement the raking process utilizing commands from the R package *survey* (Lumley, 2013). The complex survey design of the FF studies involves cluster sampling and requires corresponding specification when defining the survey subject. The variable "natpsu" represents the primary sampling unit (PSU), and "natstratum" represents the strata structure. Hence we define the survey object with the one stage cluster sampling, nested stratified sampling and without replacement. We use the summation of the wave 5 national weights to approximate the population size and incorporate it for the finite population correction factor.

Then we rake the exclusion-adjusted weights to the mother wave 5 weight totals, trim any outlier weights, and rake the weights. After raking, there are some extremely large weights (95% percentile: 1747; 99% percentile: 5290). We trim the large weights to remove the outliers. We choose a different trimming rule to achieve better control of the extreme weights by marital status. We set the 97.5% quantile of weights after raking for unmarried families as their upper truncation level and 95% quantile of weights for married families as their upper truncation level. Then we re-rake the weights to match the wave 5 totals. The summaries are in Table 3.

To construct replicate weights for variance estimation, we use the Jackknife schemes for stratified designs in the R package *survey*. The number of sets of replicate weights is equal

[2]For the raking variables of city weights, in Year 5, we collapse ethnicity as white and non-Hispanic verse others; in Year 3, we collapse ethnicity as white and non-Hispanic verse others, and age as ≤ 19, 20–24 and 25+.

Table 4: Summary of Year 9 national sample weights (exclude City X).

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
mother	48.64	79.69	165.50	466.40	457.80	5520.00	2473
child	19.56	101.90	186.90	497.20	488.80	5366.00	2623

to the number of PSUs, where the random subsamples exclude one PSU at each time. These subsamples were selected so that no case could appear in more than one excluded random group. Then the replicate weights for those remaining in the subsamples are adjusted by raking mothers' demographics to match the known total. For trimming on each replicate weights, we set the 97.5% quantile of weights after raking for unmarried families as their upper truncation level and 95% quantile of weights for married families as their upper truncation level. This resulted trimming values are different across the 26 replicate weights. The trimmed weights are calibrated by raking with the same factors again to match the mother weight totals in wave 5.

2.2.2 National weighting (exclude City X)

We collect the sample dispositions based on the flag *cm5samp* and *Y9childsamp*. We start with mother national weights (exclude City X) *m5natwtx* at wave 5 for these home visit cases. If these units were not assigned wave 5 weights, we move on to incorporate the weights in previous waves sequentially. This results in sample size 2389 (2200 included and 189 excluded cases). The sample size of home visits being representative of national samples is 2275.

We build the logistic regression model and use the predicted propensity scores to form deciles for the national weights. After the exclusion adjustment, we bring back the home visits with mother weights (exclude City X) *m5natwtx* from previous waves. We rake the weights to mother wave 5 weight *m5natwtx* totals. The raking variables include mother's age, education, ethnicity and marital status.

Then we rake the exclusion-adjusted weights to the mother wave 5 weight totals, trim any outlier weights, and rake the weights. We trim the large weights to remove the outliers. We set the 95% quantile of weights after raking for unmarried families as their upper truncation level and 95% quantile of weights for married families as their upper truncation level. Then we re-rake the weights to match the wave 5 mother weight totals. The summaries are in Table 4.

Finally, we construct the replicate weights for variance estimation. The number of sets of replicate weights is equal to the number of PSUs, where the random subsamples exclude one PSU at each time. The replicate weights for those remaining in the subsamples are adjusted by raking mothers' demographics to match the known total. For trimming on each replicate weights, we set the 95% quantile of weights after raking for unmarried families as their upper truncation level and 95% quantile of weights for married families as their upper truncation level. This resulted trimming values are different across the 23 replicate weights.

Table 5: Summary of Year 9 city sample weights.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
mother	26.82	40.95	58.76	95.10	94.46	830.90	1249
child	4.26	47.43	65.86	102.30	105.60	819.30	1505

The trimmed weights are calibrated by raking with the same factors again to match the mother weight totals in wave 5.

2.2.3 City weighting

We start with mother city weights *m5citywt* at wave 5 for these home visit cases. If these units were not assigned wave 5 weights, we move on to incorporate the weights in previous waves sequentially. This results in sample size 3595 (3295 included and 300 excluded cases). The sample size of home visits being representative of city samples is 3393. The adjustment for exclusion and poststratification is done city by city.

We build the logistic regression model and use the predicted propensity scores to form quintiles for the city weights. After the exclusion adjustment, we bring back the home visits with mother city weights from previous waves. We rake the weights to mother wave 5 city weight totals. The raking variables include mother's age, education, ethnicity and marital status.

Then we rake the exclusion-adjusted weights to the mother wave 5 city weight totals, trim any outlier weights, and rake the weights. We set the 95% quantile of weights after raking for unmarried families as their upper truncation level and 95% quantile of weights for married families as their upper truncation level. Then we re-rake the weights to match the wave 5 mother weight totals. The summaries are in Table 5.

Finally, we construct the replicate weights for variance estimation. The number of sets of replicate weights is equal to the number of PSUs, where the random subsamples exclude one PSU at each time. The variables "citypsu" and "citystratum" indicator the PSU and strata structure for the city weights, where hospitals are the PSU. The replicate weights for those remaining in the subsamples are adjusted by raking mothers' demographics to match the known total.

3 Five-year home visit

3.1 Overview of the Five-Year follow-up data collection

The Year 5 In-Home Longitudinal Study of Pre-School Aged Children (LSPAC) is a collaborative research of the FF study. The LSPAC collects information on a variety of domains of the child's environment, including: (i) Physical Environment—through quality of housing, nutrition and food security, health care, adequacy of clothing and supervision; (ii) Parenting—through parental discipline, parental attachment, and cognitive stimulation.

Table 6: Sample classification for Year 5 child weighting

Classification for weighting				Child	
				No HV	HV
Eligible	located	response (with mom weight)	-7 NA (with survey data)	1723	2377
			1 Mother died	16	0
			3 Adopted/Ne parent has legal custody	77	14
		nonresponse	5 Refusal	180	0
			7 Other non-response	148	0
	unlocated		6 Could Not Locate	316	0
Ineligible			2 Child died	42	0
			4 Other Ineligible	5	0

In addition, the LSPAC also collects information on several important child outcomes, including anthropometrics, child behaviors, and cognitive ability. This information has been collected through interviews with the child's primary caregiver, administration of standard tests; direct observation of the child's home environment and the child's interactions with the caregiver. The Five-Year survey collects data when the children are about five years old and was completed in 2006.

The survey instrument composes of two components: a parent survey questionnaire and an activity booklet. Slightly over 91% of the respondents of the Five-Year Core mother survey were contacted and invited to participate in the In-home survey. Among people contacted, about 81% completed the Five-year In-Home study. About 78% of the Five-Year In-Home respondents completed both components of the survey. Most of the remaining participants completed only the parent interview over the telephone either because the parent or the care giver refused a home visit or such visit could not be conducted because the family had moved away from the last located residence without leaving any new contact information. A very small fraction of the respondents completed only a part of the activity assessment.

Respondents of the Fragile Families Baseline survey were located and screened for eligibility for inclusion in the succeeding waves of the core survey and collaborative studies of the core survey. The survey administration process allows all still eligible respondents of the Baseline survey to participate in any follow-up surveys of the Fragile Families Study. As such, eligible respondents who could not participate in a prior wave of the follow-up survey, because of reasons other than permanent refusal, may still participate in the current or future wave of the follow-up survey. Only respondents of the Five-year Core survey, however, were invited to participate in the Five-year In-Home survey. Hence, we start from the Five-year bio mother weights to construct the weights for the Five-year home visits.

3.2 Child weighting

In Year 5, we will construct weights for every child who participated in any assessment (n=2391—the indicator variable is *Year5CHILD*).

Table 6 presents the classification for Year 5 biological mother samples and child samples

with home visits or not. The Year 5 mother weights are assigned to those who responded in the survey (with survey data, mother died, and child adopted or neither parent has legal custody). For children from these 4207 families, 2391 of them participated the home visit while 1816 did not. The 1816 non-participants neither were not sampled for the home visit or refused to participate by the mothers. We use inclusion to represent their status—both selected and responded—a unit is included only when it has been selected and also responds. We will adjust for the exclusion for the child samples considering the 4207 cases with Year 5 mother weight.

To account for the exclusion, we build a regression model with the binary inclusion indicator as the outcome for the cases with Year 5 mother weight. The covariates in the regression model are collected from mother wave 4 survey variables, listed in Appendix B. The covariates are available for samples both with home visits and without home visits. We did a preliminary selection by excluding the variables with more than 20% item missingness, more than 11 possible values.[3] We filled in the missing items by random draws from the corresponding observed frequency distributions. The predictors after dummy coding are used as covariates.

We used the predicted inclusion propensity scores to form deciles for the national weights, and quartiles within city for the city-level weights. We used these deciles to form the weighting cells for the exclusion adjustments. Therefore, each weighting cell comprises sample members who have similar inclusion 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.

After the exclusion adjustment, we rake the weights to wave 4 mother weight totals. The raking variables include mother's age, education, ethnicity and marital status. Then we trim the large weights and re-rake.

3.2.1 National weighting

We start with mother national weights at wave 4 for these home visit cases. This results in sample size 2976 (1728 included and 1248 excluded cases). The sample size of home visits being representative of national samples is 1728.

We build the logistic regression model under **Lasso** (Friedman *et al.* , 2010) for regularization with the inclusion indicator as the outcome and variables in Appendix B as covariates and use the predicted propensity scores to form deciles for the national weights. We rake the weights to mother wave 4 weight totals. The raking variables include mother's age, education, ethnicity and marital status. We implement the raking process utilizing commands from the R package *survey* (Lumley, 2013). The complex survey design of the FF studies involves cluster sampling and requires corresponding specification when defining the survey subject. The variable "natpsu" represents the primary sampling unit (PSU), and "natstratum" represents the strata structure. Hence we define the survey object with the one stage

[3]we did not include the continuous variables here, but we included city and hospitals as covariates. We will use dummy coding the regression to recode these categorical variables. Only using categorical variables helps implement the R package *glmnet*.

Table 7: Summary of Year 5 national sample weights.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
mother	1.71	24.26	96.75	376.30	329.80	8005.00	1892
child	117.10	152.50	262.80	654.50	652.30	6447.00	3170

Table 8: Summary of Year 5 national sample weights (exclude City X).

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
mother	1.78	30.60	113.30	416.40	387.40	8329.00	2182
child	122.50	161.10	278.90	678.90	692.00	6317.00	3232

cluster sampling, nested stratified sampling and without replacement. We use the summation of the wave 4 national weights to approximate the population size and incorporate it for the finite population correction factor.

Then we rake the exclusion-adjusted weights to the mother wave 4 weight totals, trim any outlier weights, and rake the weights. We trim the large weights to remove the outliers. We choose a different trimming rule to achieve better control of the extreme weights by marital status. We set the 95% quantile of weights after raking for unmarried families as their upper truncation level and 92.5% quantile of weights for married families as their upper truncation level. Then we re-rake the weights to match the wave 4 totals. The summaries are in Table 7.

To construct replicate weights for variance estimation, we use the Jackknife schemes for stratified designs in the R package *survey*. The replicate weights for those remaining in the subsamples are adjusted by raking mothers' demographics to match the known total.

3.2.2 National weighting (exclude City X)

We start with mother national weights (exclude City X) *m4natwtx* at wave 4 for these home visit cases. This results in sample size 2688 (1666 included and 1022 excluded cases). The sample size of home visits being representative of national samples is 1666.

We build the logistic regression model and use the predicted propensity scores to form deciles for the national weights. We rake the weights to mother wave 4 weight *m4natwtx* totals. The raking variables include mother's age, education, ethnicity and marital status.

Then we rake the exclusion-adjusted weights to the mother wave 4 weight totals, trim any outlier weights, and rake the weights. We trim the large weights to remove the outliers. We set the 95% quantile of weights after raking for unmarried families as their upper truncation level and 92.5% quantile of weights for married families as their upper truncation level. Then we re-rake the weights to match the wave 4 mother weight totals. The summaries are in Table 8.

Finally, we construct the replicate weights for variance estimation. The replicate weights for those remaining in the subsamples are adjusted by raking mothers' demographics to

Table 9: Summary of Year 5 city sample weights.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
mother	1.08	13.39	30.84	83.36	64.04	4927.00	735
child	37.17	58.93	86.95	146.00	138.60	1583.00	2522

match the known total.

3.2.3 City weighting

We start with mother city weights *m4citywt* at wave 4 for these home visit cases. This results in sample size 4122 (2376 included and 1746 excluded cases). The sample size of home visits being representative of city samples is 2376. The adjustment for exclusion and poststratification is done city by city.

We build the logistic regression model and use the predicted propensity scores to form quartiles for the city weights. We rake the weights to mother wave 4 city weight totals. The raking variables include mother's age, education, ethnicity and marital status. Different from national weighting, we aggregate the ethnicity categories as: white and non-hispanic or others. This is done to make sure no empty categories in each city.

Then we rake the exclusion-adjusted weights to the mother wave 4 city weight totals, trim any outlier weights, and rake the weights. We set the 95% quantile of weights after raking for unmarried families as their upper truncation level and 95% quantile of weights for married families as their upper truncation level. Then we re-rake the weights to match the wave 4 mother weight totals. The summaries are in Table 9.

Finally, we construct the replicate weights for variance estimation. The number of sets of replicate weights is equal to the number of PSUs, where the random subsamples exclude one PSU at each time. The variables "citypsu" and "citystratum" indicator the PSU and strata structure for the city weights, where hospitals are the PSU. The replicate weights for those remaining in the subsamples are adjusted by raking mothers' demographics to match the known total.

4 Three-year home visit

4.1 Overview of the Three-Year follow-up data collection

The Year 3 LSPAC samples cover more than 79% of the respondents of the Three-Year Core survey. Of these, about 78% of the participants completed both components of the survey. Most of the remaining participants completed only the parent interview over the telephone because the parent or the care giver refused a home visit or such visit could not be conducted because the family had moved away from the city where the child was born. A very small fraction of the respondents completed only a part of the activity component.

Table 10: Sample classification for Year 3 child weighting. We treat the one HV case in the category of "7 Other non-response" as 0.

Classification for weighting				Child	
				No HV	HV
Eligible	located	response (with mom weight)	-7 NA (with survey data)	1559	2646
			1 Mother died	9	0
			3 Adopted/Ne parent has legal custody	44	11
		nonresponse	5 Refusal	174	0
			7 Other non-response	118	1
	unlocated		6 Could Not Locate	288	0
Ineligible			2 Child died	42	0
			4 Other Ineligible	6	0

Eligible respondents who could not participate in a prior wave of the follow-up survey, because of reasons other than permanent refusal, may still participate in the current or future wave of the follow-up survey. Only respondents of the Three-Year Core survey, however, were invited to participate in the Three-Year In-Home survey.

4.2 Child weighting

In Year 3, we will construct weights for every child who participated in any assessment (n=2658—the indicator variable is *Year3CHILD*).

Table 10 presents the classification for Year 3 biological mother samples and child samples with home visits or not. The Year 3 mother weights are assigned to those who responded in the survey (with survey data, mother died, and child adopted or neither parent has legal custody). For children from these 4269 families, 2657 of them participated the home visit while 1612 did not. The 1612 non-participants neither were not sampled for the home visit or refused to participate by the mothers. We use inclusion to represent their status—both selected and responded—a unit is included only when it has been selected and also responds. We will adjust for the exclusion for the child samples considering the 4269 cases with Year 3 mother weight.

To account for the exclusion, we build a regression model with the binary inclusion indicator as the outcome for the cases with Year 3 mother weight. The covariates in the regression model are collected from mother wave 3 survey variables, listed in Appendix C. The covariates are available for samples both with home visits and without home visits. We filled in the missing items by random draws from the corresponding observed frequency distributions. The predictors after dummy coding are used as covariates.

We used the predicted inclusion propensity scores to form deciles for the national weights, and quartiles within city for the city-level weights. We used these deciles to form the weighting cells for the exclusion adjustments. Therefore, each weighting cell comprises sample members who have similar inclusion 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.

Table 11: Summary of Year 3 national sample weights.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
mother	1.35	24.07	94.23	373.00	327.60	8427.00	1866
child	100.00	134.70	253.80	620.40	607.60	5828.00	3075

After the exclusion adjustment, we rake the weights to wave 3 mother weight totals. The raking variables include mother's age, education, ethnicity and marital status. Then we trim the large weights and re-rake.

4.2.1 National weighting

We start with mother national weights at wave 3 for these home visit cases. This results in sample size 3002 (1823 included and 1179 excluded cases). The sample size of home visits being representative of national samples is 1823.

We build the logistic regression model under **Lasso** (Friedman *et al.* , 2010) for regularization with the inclusion indicator as the outcome and variables in Appendix C as covariates and use the predicted propensity scores to form deciles for the national weights. We rake the weights to mother wave 3 weight totals. The raking variables include mother's age, education, ethnicity and marital status. We implement the raking process utilizing commands from the R package *survey* (Lumley, 2013). The complex survey design of the FF studies involves cluster sampling and requires corresponding specification when defining the survey subject. The variable "natpsu" represents the primary sampling unit (PSU), and "natstratum" represents the strata structure. Hence we define the survey object with the one stage cluster sampling, nested stratified sampling and without replacement. We use the summation of the wave 3 national weights to approximate the population size and incorporate it for the finite population correction factor.

Then we rake the exclusion-adjusted weights to the mother wave 3 weight totals, trim any outlier weights, and rake the weights. We trim the large weights to remove the outliers. We choose a different trimming rule to achieve better control of the extreme weights by marital status. We set the 95% quantile of weights after raking for unmarried families as their upper truncation level and 92.5% quantile of weights for married families as their upper truncation level. Then we re-rake the weights to match the wave 3 totals. The summaries are in Table 11.

To construct replicate weights for variance estimation, we use the Jackknife schemes for stratified designs in the R package *survey*. The replicate weights for those remaining in the subsamples are adjusted by raking mothers' demographics to match the known total.

4.2.2 National weighting (exclude City X)

We start with mother national weights (exclude City X) *m4natwtx* at wave 3 for these home visit cases. This results in sample size 2714 (1654 included and 1060 excluded cases). The

Table 12: Summary of Year 3 national sample weights (exclude City X).

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
mother	1.40	30.31	111.50	412.50	381.50	8824.00	2156
child	103.90	147.60	285.60	683.80	691.50	6107.00	3244

sample size of home visits being representative of national samples is 1654.

We build the logistic regression model and use the predicted propensity scores to form deciles for the national weights. We rake the weights to mother wave 3 weight *m3natwtx* totals. The raking variables include mother's age, education, ethnicity and marital status.

Then we rake the exclusion-adjusted weights to the mother wave 3 weight totals, trim any outlier weights, and rake the weights. We trim the large weights to remove the outliers. We set the 95% quantile of weights after raking for unmarried families as their upper truncation level and 92.5% quantile of weights for married families as their upper truncation level. Then we re-rake the weights to match the wave 3 mother weight totals. The summaries are in Table 12.

Finally, we construct the replicate weights for variance estimation. The replicate weights for those remaining in the subsamples are adjusted by raking mothers' demographics to match the known total.

4.2.3 City weighting

We start with mother city weights *m3citywt* at wave 3 for these home visit cases. This results in sample size 4177 (2603 included and 1574 excluded cases). The sample size of home visits being representative of city samples is 2603. The adjustment for exclusion and poststratification is done city by city.

We build the logistic regression model and use the predicted propensity scores to form quartiles for the city weights. We rake the weights to mother wave 3 city weight totals. The raking variables include mother's age, education, ethnicity and marital status. Different from national weighting, we aggregate the ethnicity categories as: white and non-hispanic or others and the age categories as: < 20, 20–24 and 25+. This is done to make sure no empty categories in each city.

Then we rake the exclusion-adjusted weights to the mother wave 3 city weight totals, trim any outlier weights, and rake the weights. We set the 95% quantile of weights after raking for unmarried families as their upper truncation level and 95% quantile of weights for married families as their upper truncation level. Then we re-rake the weights to match the wave 3 mother weight totals. The summaries are in Table 13.

Finally, we construct the replicate weights for variance estimation. The number of sets of replicate weights is equal to the number of PSUs, where the random subsamples exclude one PSU at each time. The variables "citypsu" and "citystratum" indicator the PSU and strata structure for the city weights, where hospitals are the PSU. The replicate weights for those remaining in the subsamples are adjusted by raking mothers' demographics to match

Table 13: Summary of Year 3 city sample weights.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
mother	1.23	13.73	29.79	82.27	64.56	3973.00	680
child	29.96	50.86	73.84	133.30	125.40	1478.00	2295

the known total.

A Covariate list for the response propensity score regression model for Year 9 weighting

```
 [1] "m5a2"     "m5a4"     "m5a4m"    "m5a5"     "m5a5b01"  "m5a5c01a" "m5a5b02"
 [8] "m5a5c02a" "m5a5c03a" "m5a5c04a" "m5a5c05a" "m5a5c06a" "m5a5c07a" "m5a6"
[15] "m5a7"     "m5a8"     "m5a8f01"  "m5a8f02"  "m5a8f03"  "m5a8f04"  "m5a8f05"
[22] "m5a8f06"  "m5a8f07"  "m5a8f08"  "m5a8f09"  "m5a8f10"  "m5a10"    "m5b2e"
[29] "m5b4"     "m5b4a"    "m5b23"    "m5b31"    "m5b31a"   "m5b32"    "m5c1"
[36] "m5c1a"    "m5c1b"    "m5c1c"    "m5c1d"    "m5c1e"    "m5c1f"    "m5c7"
[43] "m5c8a"    "m5d7a"    "m5e1g"    "m5e1i"    "m5e1j"    "m5e1k"    "m5e2"
[50] "m5e3"     "m5e3a"    "m5e4"     "m5e5"     "m5e6"     "m5e6b"    "m5f1"
[57] "m5f4a"    "m5f7a"    "m5f7c"    "m5f8a1"   "m5f8a2"   "m5f8a3"   "m5f12"
[64] "m5f18"    "m5f18d"   "m5f21"    "m5f22"    "m5f23a"   "m5f23b"   "m5f23c"
[71] "m5f23d"   "m5f23e"   "m5f23f"   "m5f23g"   "m5f23h"   "m5f23i"   "m5f23j"
[78] "m5f23k"   "m5f24"    "m5f25"    "m5g0"     "m5g1"     "m5g2"     "m5g2a3"
[85] "m5g2b"    "m5g2c"    "m5g2e"    "m5g3"     "m5g7"     "m5g16a"   "m5g16b"
[92] "m5g16c"   "m5g16d"   "m5g16e"   "m5g17"    "m5g19"    "m5g21a"   "m5g21b"
[99] "m5g21c"   "m5g21d"   "m5g21e"   "m5g21f"   "m5g21g"   "m5g21i"   "m5g23"
[106] "m5g24"    "m5g25"    "m5g30"    "m5g31"    "m5g32"    "m5g33"    "m5h1"
[113] "m5h2"     "m5h3"     "m5i1"     "m5i3"     "m5i3b"    "m5i3c"    "m5i4"
[120] "m5i8"     "m5i9"     "m5i11"    "m5i13p"   "m5i14a1"  "m5i14a2"  "m5i14a3"
[127] "m5i14a4"  "m5i14a5"  "m5i14b1"  "m5i14b2"  "m5i14b3"  "m5i14b4"  "m5i14c"
[134] "m5i16a"   "m5i16b"   "m5i16c"   "m5i17"    "m5i19"    "m5i24a"   "m5i25a"
[141] "m5i26a"   "m5j2"     "m5j6"     "m5j6b"    "m5j9"     "m5j9b"    "m5k8"
[148] "cm5gmom"  "cm5gdad"  "m5e8_0"   "m5e8_1"   "m5e8_2"   "m5e8_3"   "m5e8_4"
[155] "m5e8_5"   "m5e8_6"   "m5e8_7"   "m5e9_0"   "m5e9_1"   "m5e9_2"   "m5e9_3"
[162] "m5e9_4"   "m5e9_5"   "m5e9_6"   "m5e9_7"   "city"     "hospital"
```

B Covariate list for the response propensity score regression model for Year 5 weighting

```
[1] "m4a2"  "m4a4"  "m4a7"  "m4a8"  "m4a8c"  "m4a10b1"  "m4a12e"
```

```
[8] "m4a16"  "cm4relf"  "cm4marf"  "cm4cohf"  "m4b0"  "m4b1"  "m4b2"
[15] "m4b2a"  "m4b2b"  "m4b4a1"  "m4b4a2"  "m4b4a3"  "m4b4a4"  "m4b4a5"
[22] "m4b4a6"  "m4b4a7"  "m4b4a8"  "m4b4b1"  "m4b4b2"  "m4b4b3"  "m4b4b4"
[29] "m4b4b5"  "m4b4b6"  "m4b4b7"  "m4b4b8"  "m4b4b9"  "m4b4b10"  "m4b4b11"
[36] "m4b4b12"  "m4b4b13"  "m4b4b14"  "m4b4b15"  "m4b4b16"  "m4b4b17"  "m4b4b18"
[43] "m4b4b19"  "m4b5"  "m4b6a"  "m4b6b"  "m4b6c"  "m4b6d"  "m4b7"
[50] "m4b8"  "m4c1"  "m4c5a"  "m4c6"  "m4c6a"  "m4c7"  "m4c7a"
[57] "m4c7b"  "m4c7c"  "m4c7d"  "m4c7e"  "m4c8"  "m4c11"  "m4c27"
[64] "m4c30"  "m4c33"  "m4c37"  "m4c38"  "m4c39"  "m4c40a"  "m4c40b"
[71] "m4c41a"  "m4c41b"  "m4c41c"  "m4c41d"  "m4c42b"  "m4c43a"  "m4c44a"
[78] "m4d1"  "m4d1a"  "m4d1b"  "m4d1c"  "m4d1d"  "m4d1e"  "m4d1f"
[85] "m4d1g"  "m4d1h"  "m4d2"  "m4d3"  "m4d4"  "m4d4a"  "m4d5"
[92] "m4d8"  "m4d10"  "m4d10a"  "m4e1"  "cm4marp"  "cm4cohp"  "m4f2b1"
[99] "m4f2b2"  "m4f3"  "cm4gdad"  "cm4gmom"  "m4h1"  "m4h1g"  "m4h1i"
[106] "m4h1j"  "m4h1l"  "m4h1m"  "m4h2"  "m4h3"  "m4h4"  "m4h5"
[113] "m4h6"  "m4i0"  "m4i0k"  "m4i0l"  "m4i0m1"  "m4i0m2"  "m4i0m3"
[120] "m4i0m4"  "m4i0m5"  "m4i0n1"  "m4i0n2"  "m4i0n3"  "m4i0n4"  "m4i0n5"
[127] "m4i0o"  "m4i0p"  "m4i1"  "m4i7a"  "m4i7b"  "m4i7c"  "m4i7d"
[134] "m4i7e"  "m4i7f"  "m4i7h"  "m4i8a1"  "m4i8a2"  "m4i8a3"  "m4i9"
[141] "m4i15"  "m4i18d"  "m4i19"  "m4i21"  "m4i23a"  "m4i23b"  "m4i23c"
[148] "m4i23d"  "m4i23e"  "m4i23f"  "m4i23g"  "m4i23h"  "m4i23i"  "m4i23j"
[155] "m4i23k"  "m4i23l"  "m4i23m"  "m4i23n"  "m4i23p1"  "m4i23p2"  "m4i23p3"
[162] "m4i23p4"  "m4i23p5"  "m4i23p6"  "m4i24"  "m4i25"  "m4j0"  "m4j1"
[169] "m4j2"  "m4j2b"  "m4j2c"  "m4j3"  "m4j5"  "m4j9"  "m4j18"
[176] "m4j20"  "m4j22a"  "m4j22b"  "m4j22c"  "m4j22d"  "m4j22e"  "m4j22f"
[183] "m4j22g"  "m4j22i"  "m4j22j"  "m4j24a"  "m4j25a1"  "m4j25a2"  "m4j25b1"
[190] "m4j25b2"  "m4j25b3"  "m4j25b4"  "m4j25c"  "cm4md_case_con" "cm4md_case_lib"
[196] "m4r1"  "m4r2"  "m4r3"  "m4k1"  "m4k3"  "m4k3b"  "m4k3c"  "m4k4"
[204] "m4k11"  "m4k12"  "m4k13p"  "m4k14a1"  "m4k14a2"  "m4k14a3"  "m4k14a4"
[211] "m4k14a5"  "m4k14b3"  "m4k14b4"  "m4k15"  "m4k16a"  "m4k16b"  "m4k16c"
[218] "m4k17"  "m4k24a"  "m4k25a"  "m4k26a"  "m4l2"  "m4l3"  "city"  "hospital"
```

C Covariate list for the response propensity score regression model for Year 3 weighting

```
 [1] "m3a2"  "m3a4"  "m3a7"  "m3a8"  "m3a8c"  "m3a10"  "m3a11a"
 [8] "m3a12"  "m3a12d"  "m3a16"  "cm3relf"  "cm3marf"  "cm3cohf"  "m3b0"
[15] "m3b1"  "m3b2"  "m3b4a"  "m3b4b"  "m3b4c"  "m3b4d"  "m3b4e"
[22] "m3b4f"  "m3b4g"  "m3b4h"  "m3b4i"  "m3b4j"  "m3b4k"  "m3b4l"
[29] "m3b4m"  "m3b5"  "m3b6a"  "m3b6b"  "m3b6c"  "m3b6d"  "m3b7"
[36] "m3c1"  "m3c5a"  "m3c6"  "m3c7a"  "m3c7b"  "m3c7c"  "m3c7d"
```

```
 [43] "m3c8"  "m3c11"  "m3c31"  "m3c34"  "m3c39"  "m3c41a"  "m3c43"
 [50] "m3c44"  "m3d0"  "m3d1"  "m3d1a"  "m3d1b"  "m3d1c"  "m3d1d"
 [57] "m3d1e"  "m3d1f"  "m3d2"  "m3d3"  "m3d4"  "m3d4a"  "m3d4b"
 [64] "m3d5"  "m3d6"  "m3e1"  "cm3marp"  "cm3cohp"  "m3f2b1"  "m3f2b2"
 [71] "m3f3"  "cm3gdad"  "cm3gmom"  "m3h1"  "m3h2"  "m3h3"  "m3h4"
 [78] "m3h5"  "m3h6"  "m3h7"  "m3h8"  "m3i0a"  "m3i0b"  "m3i0c"
 [85] "m3i0d"  "m3i0e"  "m3i0f"  "m3i0g"  "m3i0i"  "m3i0l"  "m3i0m"
 [92] "m3i0n"  "m3i0o"  "m3i0p"  "m3i0q"  "m3i1"  "m3i6a"  "m3i6c"
 [99] "m3i6e"  "m3i6h"  "m3i6j"  "m3i7a"  "m3i7b"  "m3i7c"  "m3i7d"
[106] "m3i7e"  "m3i7f"  "m3i7g"  "m3i7i"  "m3i7j"  "m3i8a1"  "m3i8a2"
[113] "m3i8a3"  "m3i9"  "m3i14"  "m3i15"  "m3i19"  "m3i21"  "m3i23a"
[120] "m3i23b"  "m3i23c"  "m3i23d"  "m3i23e"  "m3i23f"  "m3i23g"  "m3i23h"
[127] "m3i23i"  "m3i23j"  "m3i24"  "m3i25"  "m3j0a"  "m3j0b1"  "m3j0b2"
[134] "m3j0b3"  "m3j0b4"  "m3j0b5"  "m3j0b6"  "m3j0b7"  "m3j1"  "m3j2"
[141] "m3j2a"  "m3j2c"  "m3j3"  "m3j5"  "m3j9"  "m3j18"  "m3j28"
[148] "m3j36a"  "m3j36b"  "m3j36c"  "m3j36d"  "m3j36e"  "m3j36f"  "m3j36g"
[155] "m3j36h"  "m3j36i"  "m3j36j"  "m3j43a"  "m3j44a"  "m3j44b"  "m3j44c"
[162] "m3j44d"  "m3j44e"  "m3j44f"  "m3j45"  "m3j48"  "m3j50"  "m3j51"
[169] "m3j52"  "m3j52a"  "m3j52b"  "m3j53"  "m3j54"  "cm3alc_case"  "cm3drug_case"
[176] "cm3gad_case" "cm3md_case_con" "cm3md_case_lib" "m3r0a" "m3r0b" "m3r1" "m3r9"
[183] "m3r10"  "m3r11"  "m3k1"  "m3k3"  "m3k3b"  "m3k3c"  "m3k4"
[190] "m3k11"  "m3k12"  "m3k13p"  "m3k14a1"  "m3k14a2"  "m3k14a3"  "m3k14a4"
[197] "m3k14a5"  "m3k14b3"  "m3k15"  "m3k16a"  "m3k16b"  "m3k16c"  "m3k17"
[204] "m3k24a"  "m3k25a"  "m3k26a"  "m3k27a"  "m3l2"  "m3l3"  "city" "hospital"
```

D Frequency of raking variables for Year 9

Table 14: Frequency of the mother's demographic information used for raking national weights; FF–Fragile Families samples.

MSN:	married	unmarried	NA
FF	827	2615	1456

EDUN:	<8th grade	Some HS	HS or equiv	Some College	College+	NA
FF	193	972	1026	852	399	1456

ETHN:	white, non-hispanic	black, non-hispanic	hispanic	other	NA
FF	1020	845	1430	147	1456

AGEN:	<18	18-19	20-24	25-29	30-34	35-30	40+	NA
FF	108	514	1262	757	487	239	75	1456

Table 15: Frequency of the mother's demographic information by used for raking city weights; FF–Fragile Families samples.

MS:	married	unmarried	NA
FF	1155	3634	109

EDU:	< HS	HS or equiv	Some College	College+	NA
FF	1679	1214	1240	656	109

ETH:	white, non-hispanic	black, non-hispanic	other	NA
FF	998	2257	1534	109

AGE:	≤19	20-24	25-34	35+	NA
FF	841	1724	1777	447	109

References

Friedman, J., Hastie, T., & Tibshirani, R. 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, **33**(1), 1–22.

Lumley, Thomas. 2013. *Analysis of complex survey samples.* http://cran.r-project.org/web/packages/survey.