## APPENDIX A. DETAILED OUTPUTS

|                      | Overall | <$35k | $35–55k | $55–100k | >$100k |
|----------------------|---------|-------|---------|----------|--------|
| **Estimate**         |         |       |         |          |        |
| Unweighted estimate  | 0.101   | 0.192 | 0.070   | 0.048    | 0.023  |
| [0.088, 0.114]       |         |       | [0.045, 0.107] | [0.032, 0.072] | [0.014, 0.038] |
| Weighted estimate    | 0.093   | 0.165 | 0.048   | 0.059    | 0.019  |
| [0.078, 0.111]       |         |       | [0.026, 0.086] | [0.035, 0.097] | [0.010, 0.035] |
| WFPBB                | 0.114   | 0.205 | 0.046   | 0.078    | 0.039  |
| [0.084, 0.160]       |         |       | [0.026, 0.073] | [0.034, 0.159] | [0.014, 0.093] |
| Classical MRP        | 0.090   | 0.164 | 0.067   | 0.048    | 0.024  |
| [0.076, 0.106]       |         |       | [0.041, 0.099] | [0.031, 0.069] | [0.014, 0.038] |
| WFPBB-MRP            | 0.091   | 0.163 | 0.072   | 0.052    | 0.023  |
| [0.076, 0.108]       |         |       | [0.044, 0.107] | [0.031, 0.079] | [0.013, 0.036] |
| Multinomial MRP      | 0.091   | 0.162 | 0.072   | 0.052    | 0.023  |
| [0.077, 0.107]       |         |       | [0.044, 0.106] | [0.033, 0.074] | [0.013, 0.035] |
| Two-stage MRP        | 0.091   | 0.162 | 0.065   | 0.045    | 0.022  |
| [0.077, 0.105]       |         |       | [0.041, 0.096] | [0.028, 0.064] | [0.012, 0.034] |
| **Standard error**   |         |       |         |          |        |
| Unweighted estimate  | 0.006   | 0.013 | 0.015   | 0.010    | 0.006  |
| Weighted estimate    | 0.008   | 0.017 | 0.015   | 0.015    | 0.006  |
| WFPBB                | 0.020   | 0.040 | 0.012   | 0.033    | 0.021  |
| Classical MRP        | 0.008   | 0.015 | 0.015   | 0.010    | 0.006  |
| WFPBB-MRP            | 0.008   | 0.016 | 0.017   | 0.012    | 0.006  |
| Multinomial MRP      | 0.008   | 0.014 | 0.016   | 0.011    | 0.006  |
| Two-stage MRP        | 0.007   | 0.014 | 0.014   | 0.009    | 0.006  |
| **95% CI length**    |         |       |         |          |        |
| Unweighted estimate  | 0.025   | 0.052 | 0.060   | 0.039    | 0.024  |
| Weighted estimate    | 0.033   | 0.067 | 0.057   | 0.060    | 0.023  |
| WFPBB                | 0.076   | 0.156 | 0.046   | 0.125    | 0.079  |
| Classical MRP        | 0.031   | 0.059 | 0.057   | 0.038    | 0.025  |
| WFPBB-MRP            | 0.032   | 0.061 | 0.063   | 0.047    | 0.023  |
| Multinomial MRP      | 0.030   | 0.056 | 0.062   | 0.042    | 0.023  |
| Two-stage MRP        | 0.028   | 0.056 | 0.055   | 0.036    | 0.022  |

*Note:* We compare the point estimate, standard error, and 95% confidence interval (CI) length between the unweighted and survey-weighted estimators, direct imputation of the outcome using the weighted finite population Bayesian bootstrap (WFPBB), classical multilevel regression and poststratification (MRP), and embedded multilevel regression and poststratification (EMRP) methods (WFPBB-MRP, Multinomial MRP, Two-stage MRP).
Table A2: Results from estimating food insecurity prevalence across subgroups defined by interactions of annual income and agency visitor status.

|                | Overall | <$35k | $35–55k | $55–100k | >$100k |
|----------------|---------|-------|---------|----------|--------|
|                | Nonvisitor | Visitor | Nonvisitor | Visitor | Nonvisitor | Visitor | Nonvisitor | Visitor | Nonvisitor | Visitor | Nonvisitor | Visitor |
| Unweighted estimate | 0.060  | 0.206 | 0.136  | 0.274    | 0.020  | 0.153 | 0.022 | 0.025 |
| Weighted estimate | 0.065 | 0.201 | 0.132  | 0.247    | 0.027  | 0.234 | 0.019 | 0.014 |
| WFPBB-MRP       | 0.081  | 0.189 | 0.177  | 0.246    | 0.016  | 0.249 | 0.042 | 0.024 |
| Multinomial MRP | 0.053  | 0.178 | 0.109  | 0.248    | 0.038  | 0.121 | 0.024 | 0.041 |
| Two-stage MRP   | 0.054  | 0.191 | 0.111  | 0.246    | 0.035  | 0.130 | 0.024 | 0.041 |
| Classical MRP   | 0.090  | 0.164 | 0.067  | 0.048    | 0.024  |       |        |        |

Note: We compare the point estimates between the unweighted and survey-weighted estimators, direct imputation of the outcome using the weighted finite population Bayesian bootstrap (WFPBB), and embedded multilevel regression and poststratification (EMRP) methods (WFPBB-MRP; Multinomial MRP, Two-stage MRP) with the marginalized classical multilevel regression and poststratification (MRP) estimate for reference.

Appendix B. Simulation Study with Only Main Effects in the \((X|Z)\) Model

We consider the case that the \((X|Z)\) model under Two-stage MRP is correctly specified. During the simulation study, we assume that Model (6) only has main effects of \(Z\) (MAIN) with \(\beta_0 = -0.5, \beta^Z_a = (1.7, 0.25, 0.2, -0.75, -1.7)^T, \beta^Z_b = (2.3, 1.5, 0.15, 0.2, 0.9)^T, \beta^Z_c = (0, -1)^T,\) and null values for the interaction terms \(\beta_Z^aZ_b\) and \(\beta_Z^aZ_c\). Other settings are the same as those in Section 3.1. The simulated population cell frequencies range from 5 to 470, with an average of 100. The average sampled cell sizes in the \((Z, X)\) cross-tabulation table for subgroups \(p_i: 0.03-0.27, p_i: 0.21-0.38, p_i: 0.30-0.54,\) and \(p_i: 0.40-0.95\) are \((9, 24, 50, 89)\) with total subgroup sample sizes \((194, 499, 1017, 1785)\), respectively. The results are shown in Table B1. The conclusions comparing EMRP, MRP and weighted estimators stay the same. Among the three EMRP methods, the CIs from Multinomial MRP and Two-stage MRP are similar and narrower than those under WFPBB-MRP. Multinomial MRP and Two-stage MRP have coverage rates of 94.5% or above, as an improvement over the case when the \((X|Z)\) model is misspecified. The performance of WFPBB-MRP is competitive when Multinomial MRP and Two-stage MRP work well.
TABLE B1  Simulation results for the overall and subdomain mean estimates, where the \((X|Z)\) relationship contains main effects only.

|               | Overall | \( p_1: 0.40–0.95 \) | \( p_1: 0.30–0.54 \) | \( p_1: 0.21–0.38 \) | \( p_1: 0.03–0.27 \) |
|---------------|---------|----------------------|----------------------|----------------------|----------------------|
| Population \(Y\) | 0.572   | 0.567                | 0.441                | 0.568                | 0.600                |
| rMSE          |         |                      |                      |                      |                      |
| Unweighted Est.| 0.015   | 0.034                | 0.025                | 0.027                | 0.041                |
| Weighted Est. | 0.010   | 0.007                | 0.012                | 0.017                | 0.055                |
| WFPBB         | 0.010   | 0.006                | 0.011                | 0.017                | 0.054                |
| Classical MRP | 0.007   | 0.048                | 0.076                | 0.060                | 0.081                |
| WFPBB-MRP     | 0.007   | 0.005                | 0.010                | 0.010                | 0.018                |
| Multinomial MRP| 0.006  | 0.005                | 0.011                | 0.010                | 0.017                |
| Two-stage MRP | 0.007   | 0.005                | 0.010                | 0.008                | 0.016                |
| Bias          |         |                      |                      |                      |                      |
| Unweighted Est.| 0.014   | 0.034                | −0.022               | −0.019               | 0.027                |
| Weighted Est. | −0.001  | <0.001               | <0.001               | 0.001                | −0.006               |
| WFPBB         | −0.001  | −0.001               | 0.002                | −0.003               | 0.004                |
| Classical MRP | −0.002  | −0.048               | 0.075                | −0.060               | −0.080               |
| WFPBB-MRP     | <0.001  | 0.002                | −0.006               | −0.003               | 0.003                |
| Multinomial MRP| <0.001 | 0.002                | −0.007               | −0.003               | 0.001                |
| Two-stage MRP | −0.001  | 0.003                | −0.007               | −0.002               | 0.003                |
| 95% CI length |         |                      |                      |                      |                      |
| Unweighted Est.| 0.023   | 0.035                | 0.046                | 0.067                | 0.103                |
| Weighted Est. | 0.035   | 0.037                | 0.048                | 0.068                | 0.166                |
| WFPBB         | 0.045   | 0.049                | 0.062                | 0.087                | 0.204                |
| Classical MRP | 0.033   | 0.033                | 0.037                | 0.040                | 0.062                |
| WFPBB-MRP     | 0.038   | 0.041                | 0.050                | 0.059                | 0.107                |
| Multinomial MRP| 0.032  | 0.034                | 0.041                | 0.040                | 0.072                |
| Two-stage MRP | 0.033   | 0.034                | 0.041                | 0.041                | 0.073                |
| Coverage rate |         |                      |                      |                      |                      |
| Unweighted Est.| 0.325   | 0.015                | 0.535                | 0.755                | 0.775                |
| Weighted Est. | 0.925   | 0.995                | 0.950                | 0.960                | 0.860                |
| WFPBB         | 0.955   | 1.000                | 1.000                | 0.985                | 0.925                |
| Classical MRP | 0.985   | 0.000                | 0.000                | 0.000                | 0.000                |
| WFPBB-MRP     | 0.995   | 1.000                | 0.995                | 1.000                | 1.000                |
| Multinomial MRP| 0.980  | 0.995                | 0.945                | 0.950                | 0.960                |
| Two-stage MRP | 0.980   | 0.995                | 0.945                | 0.990                | 0.970                |

Note: Subdomains are defined by inclusion probability ranges and the joint distribution of \((Z, X)\). We report root mean squared error (rMSE), absolute bias, average 95% confidence interval (CI) length, and 95% CI coverage rate from the direct imputation of the outcome using the weighted finite population Bayesian bootstrap (WFPBB), classical multilevel regression and poststratification (MRP), embedded multilevel regression and poststratification (EMRP) methods (WFPBB-MRP, Multinomial MRP, Two-stage MRP), and the unweighted and survey-weighted estimators.

APPENDIX C. MODEL DIAGNOSTICS FOR THE APPLICATION STUDY

We conduct a posterior predictive check and Bayesian leave-one-out cross-validation to evaluate model fit and performance. For the posterior predictive check, we generate posterior predictive distributions of food insecurity prevalence for each analysis subgroup and compare them against the observed values from the LSW sample. Figure C1 includes the histogram of 5000 estimates of \( \hat{\mu}_k^{\text{pred}} \), \( k = 1, \ldots , 5000 \), which are derived from taking the respective subgroup means of the last 5000 posterior predictive outcome draws:

\[
\hat{\mu}_k^{\text{pred}} = \frac{1}{n_k} \sum_{i \in k} \hat{y}_i^{\text{pred}}. \tag{C1}
\]

The EMRP posterior predictive distributions are centered around the observed values in the relevant subgroups, indicating that the predictive distribution captures the structure of the real data.
FIGURE C1 Model diagnostics from applied analysis (posterior predictive distribution and ROC). Posterior predictive distributions are plotted for the food insecurity prevalence for the overall population and (income) or (visitor) analysis sets. Each density consists of 5000 estimates of $\hat{\mu}_k$ based on draws of $\hat{y}$ from the posterior predictive distribution of the LSW outcome model used by either the EMRP or classical MRP methods. The vertical line indicates the value of the direct estimate obtained from the sample.

For Bayesian leave-one-out cross-validation, we compare the pointwise out-of-sample prediction accuracy between the EMRP and MRP models. Pareto $k$ estimates are less than 0.7 for both models, which means that all leave-one-out posteriors are similar to the full posterior. The expected log predictive density for the EMRP model is greater than that of the MRP model (18.4 difference); the EMRP model is preferred for prediction.