NBER WORKING PAPER SERIES

# A FLEXIBLE MODEL OF FOOD SECURITY: ESTIMATION AND IMPLICATIONS FOR PREDICTION

Will Davis
Jose O. Xilau
Rusty Tchernis
Christian A. Gregory

Working Paper 32460
http://www.nber.org/papers/w32460

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2024

We would like to thank participants at the Batsheva de Rothschild Workshop on ”Avoiding the coming food security crisis: Novel solutions at the intersection of agriculture, environment and health”, the 2021 Agricultural & Applied Economics Association annual conference, and the 7th meeting of the Society of Economics of the Household for their thoughtful comments on earlier versions of this project. The findings and conclusions in this publication are those of the author(s) and should not be construed to represent any official USDA or U.S. Government determination or policy. This research was supported by the U.S. Department of Agriculture, Economic Research Service, cooperative agreement 58-4000-8-0027. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w32460

A Flexible Model of Food Security: Estimation and Implications for Prediction
Will Davis, Jose O. Xilau, Rusty Tchernis, and Christian A. Gregory
NBER Working Paper No. 32460
May 2024
JEL No. C11,I14,I32

**ABSTRACT**

We propose a novel Bayesian Graded Response Model (BGRM) for food security measurement. Our BGRM has several attractive features. It produces continuous food security estimates and measures of estimation uncertainty at the household level. Unlike the USDA's official measurement model, the BGRM can be used with binary and polytomous items. We further modify our BGRM to include any combination of binary, ordered polytomous, and continuous variables. With data from the 2017-18 National Health and Nutrition Examination Survey (NHANES), we estimate our BGRM for responses to the 10 adult core Food Security Module (FSM) questions. We find substantial uncertainty in household-level estimates, emphasizing the inherent uncertainty of latent trait estimation. We observe overlap in BGRM estimates across USDA-defined food security categories and significant variation within categories. We estimate our model using Current Population Survey (CPS) data as a robustness check. CPS results are qualitatively similar to those from the NHANES, highlighting possible implications for national USDA food security estimates. We explore the BGRM's ability to explain variations in health outcomes associated with food security and compare results to those produced using standard USDA category definitions. Finally, we demonstrate the BGRM's flexibility by incorporating an additional continuous variable, the Healthy Eating Index (HEI), into the model, capturing nutrition quality and food security information in a novel latent construct. The adaptability of our BGRM positions it as a versatile tool for measuring food security and other latent traits requiring a diverse range of variable types like nutrition security.

Will Davis
Department of Agricultural Economics
Mississippi State University
gwd53@msstate.edu

Jose O. Xilau
Georgia State University
jxilau1@student.gsu.edu

Rusty Tchernis
Department of Economics
Andrew Young School of Policy Studies
Georgia State University
P.O. Box 3992
Atlanta, GA 30302-3992
and NBER
rtchernis@gsu.edu

Christian A. Gregory
ERS
MS9999
Beacon Facility
P.O. Box 419205
Kansas City, MO 64113
cgregory@ers.usda.gov

# 1 Introduction

The United States Department of Agriculture (USDA) defines food security as "access by all people at all times to enough food for an active, healthy life" (Rabbitt et al. (2023)). In 2022, approximately 87.2% of U.S. households were food secure. The remaining 12.8%, roughly 17.0 million households, were food insecure, implying that these households experienced difficulty during the past year providing nutritionally adequate food for all household members (Rabbitt et al. (2023)). In addition to the immediate nutritional consequences of food insecurity, the condition represents a vital issue for public health given its associations with poor chronic health outcomes like obesity, diabetes, and hypertension (Gundersen and Ziliak (2015)). Accurate measurement of food security allows researchers, policymakers, and other stakeholders to assess the prevalence of food insecurity, evaluate the evolving needs of food insecure households, and measure the effectiveness of related programs like the Supplemental Nutrition Assistance Program (SNAP) (Bickel et al. (2000)).

Household food security is a complex construct, making it challenging to fully capture with a single variable. To determine household-level food security, researchers use multiple indicators related to the conditions, behaviors, and experiences associated with food security (Barrett (2010), Opsomer et al. (2002), National Research Council (2006)). To account for the domains of food security, the USDA uses a household-level survey instrument called the core Food Security Module (FSM). The FSM was first piloted in the 1995 Current Population Survey (CPS) Food Security Supplement (FSS) and is now the most commonly used survey instrument for U.S. food security measurement.[1] The FSM includes a 10-item adult food security questionnaire completed by every responding household and an additional 8-item questionnaire for households with children used to measure child food security. The FSM questions cover a range of conditions and behaviors that demonstrate varying levels of severity. These include household members being worried about

[1] For a brief history of the development of the module, see Smith and Gregory (2023).

food running out before having money to purchase more, not being able to afford balanced meals, and going a full day without eating due to inadequate food resources.[2] When analyzing FSM responses, many researchers use the USDA's food security scale to assign food security categories to households. These USDA categories are based on a household's number of affirmative FSM responses.[3] Each household is assigned to one of four adult food security categories using the 10 adult FSM questions. The scale is extended to include three child food security categories using the 8 child FSM questions.

There are two main benefits to measuring food security using the FSM questionnaire and USDA scale categories. First, since the FSM includes at most 18 items, it can be added to new or existing surveys without substantially increasing respondent and enumerator burden. This burden can be further reduced by screening households. Specifically, households that respond negatively to less severe items at the beginning of the FSM questionnaire are not asked more severe questions. In some surveys like the CPS, the inclusion of an additional external screener prevents households with certain characteristics from participating in the FSM entirely.[4] Second, the nature of the FSM scale

2 For a full list of FSM questions, *see* Rabbitt et al. (2023).

3 In this context, an affirmative response is one which indicates an undesirable outcome with regards to food security, implying that households with lower food security (greater food insecurity) are more likely to respond affirmatively to FSM questions. As discussed in Section 2.1, we code FSM responses in our model so that higher numerical values indicate higher food security.

4 The CPS FSS contains external and internal screening protocols to reduce respondent survey burden and avoid questions that may be inappropriate given information provided earlier in the survey. Households that pass the external screen are exempt from taking the entire FSM and they are automatically assigned to the USDA's full food security category. Furthermore, the adult FSM questionnaire in the CPS includes two internal screeners. Households that register no food stress in a given set of adult internal screener questions are not asked the remaining adult FSM questions, and their responses to skipped questions are assumed to be negative. The National Health and Nutrition Examination Survey does not employ an external screener, implying that no households are exempt from taking the entire FSM questionnaire. It does, however, employ two internal screeners in the adult food security questionnaire.

categories simplifies calculation, interpretation, and identification. For example, assigning categories to households simply involves the summation of affirmative responses. This user-friendly approach makes the concept of food security accessible and convenient for researchers, policy-makers, and other stakeholders. FSM scale categories also provide easily interpreted and relevant targets for policies focused on improving household food security.

While the USDA FSM scale categories are deterministic, the scale was developed using estimates from a probabilistic 1-parameter logistic Rasch model (1-PRM). The 1-PRM belongs to a class of methods from Item Response Theory (IRT) which are used to measure latent constructs like test performance and psychological traits (Rasch (1960), Hambleton (1989), Baker (2001), Andrich and Marais (2019)). The IRT modeling approach has several strengths. Firstly, the 1-PRM has been widely employed in the field of education research, resulting in a comprehensive understanding of its statistical properties. Researchers were therefore able to consult IRT experts to confirm that measuring food security with the 1-PRM was consistent with IRT methods and produced reasonable results (Ohls et al. (2001)). Second, using the 1995-1997 CPS waves, the scale's original creators demonstrated that 1-PRM parameter estimates were robust across time and population subgroups (Ohls et al. (2001)). Finally, the 1-PRM can be used when households have incomplete FSM responses, whether the missing responses are due to forced skip patterns from screeners or random non-response (Opsomer et al. (2002)).

Despite these strengths, the 1-PRM has limitations. First, the dichotomous 1-PRM can only be used with binary "affirmative/negative" type response variables. Some FSM questions like "In the last 12 months, did you ever cut the size of any of the children's meals because there wasn't enough money for food?" can only be answered as "yes" or "no", but 9 of the 18 FSM questions have three ordered response options.[5] These three-response FSM questions must first be converted into binary variables for use in the 1-PRM, removing potentially valuable information about the severity of each household's food security level. The implications of only using binary variables

[5] Throughout the paper, we refer to "yes/no" response option questions as binary variables/questions and use polytomous for FSM questions with three potential responses.

for all FSM responses has been discussed often in the food security measurement literature (most notably in National Research Council (2006) and Nord (2012)), but to the best of our knowledge, Nord (2012), Tanaka et al. (2020), and Opsomer et al. (2002) are the only other studies to estimate food security without converting polytomous FSM questions into binary variables. Second, while the FSM scale's creators showed that the 1995 estimates were stable across time using the two following waves of CPS data, the food security environment of U.S. households may have changed since. Acknowledging this point, the scale's original creators state "…we recommend continuing to estimate the IRT model item parameters each year, as data become available".[6] This recommendation implies that estimating the underlying measurement model once and then using parameter estimates to derive deterministic categories may be problematic if parameters need to be continually updated. This recommendation also highlights the need for ongoing refinement of food security measurement methods to ensure validity and relevance. Finally, as discussed in National Research Council (2006), the 1-PRM does not provide a direct measure of uncertainty for its food security estimates, ignoring the inherent uncertainty of latent trait measurement.

In light of the strengths and weaknesses of the USDA's FSM scale and 1-PRM, we propose an alternative food security measurement model which we estimate using responses to the 10 adult FSM questions.[7] Specifically, we estimate adult food security at the household level using a Bayesian Graded Response Model (BGRM).

Our BGRM has several attractive features applicable to food security measurement. First, the BGRM estimates latent food security as a continuous variable. Using a continuous measure allows for a more detailed understanding of each household's underlying food security level that can supplement valuable information from food security categories. Second, our BGRM accommodates both binary and polytomous ordered response variables, capturing information regarding response

---

6 For a detailed discussion of the 1-PRM's estimation and stability across time, *see* Ohls et al. (2001).

7 In this study, we use the 10 adult FSM questions to estimate households' adult food security. We do not include the 8 child FSM questions answered by households with children. All 18 FSM questions can be included in our model without modification, but adult and child food security are thought to represent distinct latent traits.

severity present in the 5 polytomous adult FSM questions. Third, as a Bayesian model, our BGRM estimates latent food security using draws from each household's conditional posterior distribution of food security. As directly recommended by National Research Council (2006), drawing from a posterior distribution provides an inherent way to measure uncertainty in food security estimates along with producing point estimates.

We use our BGRM to measure household-level adult food security with data from the 2017 - 2018 waves of the National Health and Nutrition Examination Survey (NHANES). We use the NHANES for two primary reasons in addition to it including the FSM. First, the NHANES includes continuous Healthy Eating Index (HEI) which is absent from other surveys. The HEI is a continuous scoring metric that provides information on overall diet quality for a given respondent (National Cancer Institute (2023)). In Section 4.2, we add HEI to our model as a measure of nutrition quality to augment information from the FSM, highlighting model flexibility. Second, the NHANES contains a number of observable health outcomes commonly associated with food security. We use these data to test the BGRM's ability to explain variation in health outcomes relative to measures from the USDA's FSM scale categories. Our NHANES estimates highlight several key findings. First, as expected, our results show significant uncertainty associated with household food security estimates. Second, we find overlap in BGRM food security point estimates across standard USDA-defined categories. These findings suggest that probabilistically assigning households to multiple categories may be appropriate (National Research Council (2006), Nord (2012)), something that cannot be done without a measure of uncertainty. Third, we find significant variation in BGRM food security point estimates within the USDA's FSM scale categories. This within-category variation suggests differences in the range of food security levels for households assigned to the same USDA category. Finally, we find statistically significant associations between our BGRM point estimates and the set of health outcomes. The directions of these associations match *a priori* expectations, and our estimates explain roughly similar amounts of health outcome variation relative to measures created using the USDA's FSM scale. This finding implies that no explanatory power is lost when using our more flexible, but complex, method.

While our primary results are estimated using data from the NHANES, the USDA's national food security estimates come from the CPS FSS. Given that both the CPS and NHANES collect nationally representative samples of U.S. households, our BGRM should produce qualitatively similar estimates using both surveys. Additionally, noteworthy results found using CPS data may have more direct implications for government agencies and policymakers in the U.S. In Appendix D, we estimate our BGRM using data from the 2018 CPS FSS. Qualitatively, we find BGRM estimates for the CPS that are strikingly similar to those from the NHANES, highlighting the robustness of our model and the FSM questionnaire.

We also perform an exercise with the NHANES data where HEI is included alongside the set of adult FSM questions, creating a novel latent construct. This exercise produces several interesting results. First, distributional characteristics of our estimates with HEI are qualitatively similar to those using just the FSM. Alternatively, we find that our new latent construct with HEI has marginally better explanatory power for all health outcome regressions relative to our estimates using just the FSM. These results indicate that the additional dietary quality information provided by the HEI may better explain variation in health outcomes commonly associated with food security. We caution readers regarding the interpretation of these results as more theoretical measurement work is needed to confidently interpret the appropriateness and meaning of the novel latent trait. Regardless, being able to add a continuous variable is, to the best of our knowledge, a unique feature among alternative IRT models used to measure food security.

Overall, using our BGRM to estimate food security produces several important insights. First, our findings indicate that there is significant uncertainty in household food security estimates. Our model's ability to provide household-level measures of uncertainty in food security estimation provides an attractive way to account for the uncertainty inherent to latent trait estimation. Second, the considerable overlap in our estimates across common USDA-defined food security categories highlights the importance of carefully considering the implications of deterministic, rather than probabilistic, category assignment for researchers and policymakers alike. Third, by exploiting the BGRM's ability to incorporate a variety of variable types, we can fully integrate information from all FSM questions, both binary and polytomous. Furthermore, with our exercise using both adult FSM responses and continuous HEI, we illustrate how our BGRM can effectively capture additional factors adjacent to food security like nutrition quality. For future research, our BGRM represents a valuable tool for integrating related concepts like local food access into augmented food security measures, or for estimation of new latent traits like nutrition security that may require continuous variables, all without necessitating significant adjustments to the model's structure.

## 2 The Bayesian Graded Response Model of Food Security

Before moving to the model's full mathematical description, we begin with an intuitive overview of our BGRM and its relation to the existing food security measurement literature.[8] For simplicity, we start with the case for binary response variables which shares the same logic as the polytomous case. Our BGRM assumes that each household's set of FSM responses is partially determined by their continuous latent food security variable, $\delta$. As a latent variable, $\delta$ is never directly observed. Given how we code FSM responses (see Footnote 15), households with higher levels of $\delta$ are more likely to report higher numerical response values across all FSM questions, indicative of better food security.

For the binary case, each household can only provide a response of "yes" or "no" to each FSM question. Similar to the latent variable specification of the standard probit model Muthén (1979), our BGRM assumes that the household's observed binary response, $y$, to each FSM question is determined by an unobserved (latent) continuous response variable, $y^*$. In turn, the value of $y^*$ is a function of the household's value of $\delta$, a question specific intercept parameter, $\mu$, and an error term, $e$. Each household has a $J$ length vector of $y^*$ values where $J$ is the total number of FSM questions. With a household's value of $y^*$, they respond to a FSM question with a "yes" or a "no"

[8] For readers generally interested in Graded Response models and Bayesian Item Response Theory models, *see* Samejima (1968), Fox (2010), and Johnson and Albert (2006).

depending on if their value of $y^*$ is above or below some interior response threshold parameter, $\gamma$. When including polytomous questions, the number of interior $\gamma$ threshold parameters to identify is increased by one, allowing for three response ranges in the distribution of $y^*$ rather than two.

As a Bayesian model, we use a set of prior distributions and the model's likelihood function to define each parameter's conditional posterior distribution. Each posterior distribution is conditional on the value of all other model parameters and the set of observed binary and polytomous FSM responses. Unlike the frequentist approach which estimates a single value for each parameter with an associated sampling distribution, the BGRM produces a series of draws from each parameter's conditional posterior distribution. Point estimates can then be calculated using the posterior distribution mean across all draws.

We use a Markov Chain Monte Carlo (MCMC) algorithm to estimate our model, specifically a Gibbs Sampler algorithm.[9] While the form of these conditional posterior distributions are known *a priori*, sampling from each first requires that we specify initial values for the model's parameters. Starting with these initial values, a series of draws from each posterior distribution is taken sequentially across many iterations.[10] In each iteration, a new draw is taken from the parameter's conditional posterior distribution. This new draw then serves as an updated value of the parameter in the following iteration. This sequential updating procedure is repeated until the desired number of total iterations is achieved and the set of conditional posterior distributions stabilize, known as convergence. We drop draws from early Gibbs Sampler iterations in a process known as "burn in" to ensure that the model's estimates have had adequate time to converge and stabilize. The set of post-burn-in draws from each parameter's conditional posterior distribution then serve as

---

9 Casella and George (1992) provide a more detailed, yet intuitive, overview of the Gibbs Sampler algorithm and the conditions required for model convergence.

10 Assuming that the conditions for convergence are satisfied, the only notable implication of using different starting parameter values is the number of iterations needed for the model to converge. Intuitively, choosing starting values closer to the parameters' post-convergence posterior means reduces the number of iterations required for convergence.

our estimates. The means of these conditional posterior distribution draws are used as parameter point estimates, and their 95% credible intervals are used as a common measure of estimation uncertainty.[11]

Our BGRM has several attractive features. First, our model estimates food security as a continuous latent trait, potentially capturing more variation in severity than the USDA's FSM scale categories alone. While a continuous measure of food security may provide more information than discrete categories, it comes at the cost of simplicity and, in some respects, ease of interpretation. More specifically, the four USDA FSM scale food security categories provide a straightforward method for assigning households to easily interpreted, policy relevant groups. This deterministic approach does, however, rely on the assumption that each category is mutually exclusive rather than "fuzzy", meaning that a household cannot be probabilistically assigned to multiple categories based on estimation uncertainty. In Figure 2, we find overlap in the distribution of BGRM posterior mean food security levels across three of the four USDA categories. While we do not conduct this specific analysis in our study, our BGRM estimates could also be used to create or modify food security categories based on distributional features of the latent trait estimates. Furthermore, as recommended by National Research Council (2006), uncertainty estimates from our model can be utilized to calculate the probability that a household falls within a specific food security category. Probabilistic category assignment with frequentist IRT methods has been explored in Nord (2012). While not probabilistically assigning food security categories, Gregory (2020) uses measures of uncertainty from a dichotomous Bayesian 4-parameter IRT model to probabilistically assign household misreporting probabilities with informative prior distribution.

Second, by not needing to convert polytomous FSM questions into binary variables, we can directly include information regarding the severity of related experiences. This issue, among others,

[11] In Bayesian statistics, a credible interval represents a range of credible parameter values calculated using draws from the parameter's conditional posterior distribution. The 95% credible interval specifically represents the range that includes 95% of post-burn-in parameter draws. In other words, the parameter's value falls outside the 95% credible interval for only 5% of draws. Larger 95% credible intervals imply greater levels of uncertainty.

was raised by National Research Council (2006) in their seminal report detailing potential modifications to improve USDA food security measurement. With respect to polytomous food security models, our work builds on Nord (2012), Tanaka et al. (2020), and Opsomer et al. (2002) who use a frequentist polytomous IRT model, specifically a Partial Credit Rash Model (PCRM), to estimate household food security. Similar to our BGRM, their PCRM can be used with a mix of ordered polytomous and binary outcome variables. While Nord (2012) suggests that directly modeling the FSM's polytomous responses leads to little improvement compared to the 1-PRM, Tanaka et al. (2020) find that the PCRM does improve estimation precision, highlighting the potential benefits of moving beyond binary IRT methods. Unlike the PCRM, which is built using the logistic distribution, our BGRM can be interpreted as an extension of the probit model based on the normal distribution.[12] In our Bayesian framework, the normalizing assumption allows us to use conjugate non-informative prior distributions, providing analytically tractable conditional posterior distributions for our model parameters. In turn, we are able to simulate posterior parameter draws from known statistical distributions, greatly improving computational efficiency and reducing estimation time. While we use non-informative priors in our study, informative prior distributions can be used to incorporate *a priori* information similar to Gregory (2020). This unique feature of Bayesian models is especially useful for estimation with small sample sizes.

Third, drawing samples from each household's posterior distribution of $\delta$ provides a direct way to measure estimation uncertainty. This approach aligns with a recommendation from National Research Council (2006) which states that the large amounts of uncertainty inherent in food security estimates should not be ignored. Instead, National Research Council (2006) recommends modeling food security values as a posterior distribution of potential values rather than a single point estimate. This recommendation is exactly what we do with our BGRM. Using 95% credible intervals

[12] As proposed by Bock and Aitkin (1981), general IRT models beyond the 1-parameter model cannot directly calculate the unit-level latent trait parameter without additional assumptions. They propose using Marginal Maximum Likelihood estimation after first assuming a distribution for the parameter. This approach is commonly used in the IRT literature, with many models assuming a normal distribution for the latent trait parameter.

as a measure of uncertainty for each household's food security posterior distribution, our model allows us to capture the general uncertainty present in latent trait estimation. This work builds on Nord (2012) who considers the potential role of uncertainty in his report responding to National Research Council (2006) .[13]

Finally, the purpose of our analysis is not ultimately to create a fixed food security scale or categories, but rather to estimate latent food security through full model estimation. By estimating the model in each application, we can capture potential changes in the underlying relationship between observable variables and latent food security occurring across time, groups, and data sets. This approach aligns with the suggestion from Ohls et al. (2001) that the underlying food security measurement model be estimated using different surveys and years to detect any notable changes.[14]

In addition to the features discussed above, our methodology shares two of the 1-PRM's greatest strengths. First, our BGRM only uses the set of FSM questions to which a household responds to estimate food security. Doing so allows us to include households with both complete and incomplete responses. This flexibility is valuable in capturing food security for a wider range of households rather than only for households with valid responses to all FSM questions. Second, similar to the 1-PRM, our BGRM determines the relative contribution of each manifest variable to the measurement of latent food security based on data rather than relying on expert opinion like many other scales/indices. This data-driven approach ensures that the food security measure is grounded in empirical evidence based on relationships present in the data.

---

13 The limited amount of information provided by binary and polytomous response variables can produce significant estimation uncertainty. As expected, the uncertainty shown for estimates from our model are significant. Regardless, since our model produces measures of uncertainty as a direct result of estimation, the BGRM can be used to examine the levels of estimation uncertainty present within and across households, supplementing information from point estimates.

14 Additionally, as discussed in a conference paper by Opsomer et al. (2002), the 1-PRM is overparametrized, implying that comparability of model parameter estimates across data sets requires some significant assumptions.

## 2.1 Model Specification

Moving to the full BGRM specification, we use households' FSM questionnaire responses to measure food security. Our BGRM includes the following parameters and variables:

1. $y_{ij}$ denotes the discrete response of household $i = 1, ..., I$, to FSM question $j = 1, ..., J$. Each question $j$ has a set of ordered, potential responses ranging from 1 to $C_j$, implying that $y_{ij} \in \{1, ..., C_j\}$.

2. The variable $\delta_i$ represents the latent food security level of household $i$ which is invariant across the set of FSM responses provided by the household.[15]

3. The set of question specific intercept parameters $\mu = \mu_1, ..., \mu_J$.

4. Response thresholds for each question $j$ denoted by $\gamma^j = \gamma_0^j,\ \gamma_1^j,\ ...,\ \gamma_{C_j}^j$, such that $\gamma_{k-1}^j \leq \gamma_k^j$ for all $k = 1, ..., C_j$.

5. Random variation in the elicited food security level of household $i$ with respect to question $j$, $e_{ij}$, which captures error in a question's ability to measure latent food security.

With these variables, the model of household FSM responses is defined as follows. Household $i$ answers FSM question $j$ with a response of $y_{ij} = k$ if and only if:

$$\gamma_{k-1}^j < \mu_j + \delta_i + e_{ij} \leq \gamma_k^j \tag{1}$$

[15] Affirmative responses to FSM questions (e.g., a "yes" response rather than a "no" response) are designed to be negatively related to food security. To interpret $\delta$ as a measure of latent food security rather than food insecurity, we code the numerical response options of all FSM questions such that higher numerical response values (i.e., higher values of $y_{ij}$) correspond to higher food security levels (higher $\delta$s). This approach aligns with the conceptualization of food security as a continuum. For USDA FSM categories, the continuum ranges from full food security (highest values of $\delta$ in our model) to very low food security (lowest values of $\delta$). Since this is only a matter of interpretation, however, response values could easily be coded such that higher values are indicative of lower food security rather than higher.

We assume that the distribution of each $e_{ij}$ is Normal with mean 0 and variance $\sigma_j^2$. Furthermore, for the purpose of identification, we restrict the value of the first and last response threshold parameter for each question $j$ such that $\gamma_0^j = -\infty$ and $\gamma_{C_j}^j = \infty$.

The model of FSM responses given by (1) implies that the probability household $i$ provides response $k$ to FSM question $j$ is such that:

$$P(y_{ij} = k) = \Phi\left(\frac{\gamma_k^j - (\mu_j + \delta_i)}{\sigma_j}\right) - \Phi\left(\frac{\gamma_{k-1}^j - (\mu_j + \delta_i)}{\sigma_j}\right) \tag{2}$$

where $\Phi()$ is the CDF of the standard Normal distribution.

The response probabilities given by (2) describe the likelihood function of a graded response model with a probit link function, given formally by:

$$L(\mu, \delta, \gamma, \sigma^2) = \prod_i \prod_{j \in C_i} \left[\Phi\left(\frac{\gamma_{y_{ij}}^j - (\mu_j + \delta_i)}{\sigma_j}\right) - \Phi\left(\frac{\gamma_{y_{ij}-1}^j - (\mu_j + \delta_i)}{\sigma_j}\right)\right] \tag{3}$$

The outer product of (3) extends over all households in the sample, and the inner product extends over the set of questions $j$ answered by household $i$, denoted by $C_i$.[16]

Examining the model's likelihood in (3), we see that not all model parameters can be identified. As is the case with many ordinal response models, the value of the likelihood does not change with affine transformations of each question's response threshold parameters, intercepts, error variances, or latent food security levels. To address this, we restrict $\gamma_1^j = 0$ for all FSM questions $j$ and make the normalizing assumption that $\delta_i \sim N(0, 1)$ which serves as the prior distribution for $\delta$.

[16] As indicated by $C_i$, our model does not assume that all households provide responses to all possible FSM questions. This allows for non-responses to certain questions which is a vital feature as many surveys screen respondents, allowing them to skip some or all FSM questions.

The prior distributions used in our model for $\mu$, $\sigma^2$, and $\gamma$ are given as:

$$\mu_j \sim N(\bar{\mu}, \sigma_\mu^2),\ \forall j = 1, ..., J \tag{4}$$

$$\sigma_j^2 \sim IG(\alpha, \beta),\ \forall j = 1, ..., J \tag{5}$$

$$\gamma_k^j \sim U(a, b)1(\gamma_{k-1}^j < \gamma_k^j \leq \gamma_{k+1}^j),\ \forall k = 2, ..., C_j - 1 \tag{6}$$

where $IG()$ denotes the inverse gamma distribution and $U()$ denotes the uniform distribution. Our set of prior distributions, including the normal prior placed on latent food security, represent conjugate prior distributions. In turn, conjugate prior distributions produce conditional posterior distributions that correspond to known probability distributions. This approach makes sampling from the posterior distributions substantially more tractable and computationally efficient.

Exploiting a unique feature of Bayesian models, we use a process known as data augmentation to define a new latent variable, $y^*$, such that for household $i$ and FSM question $j$:

$$y_{ij}^* = \mu_j + \delta_i + e_{ij} \tag{7}$$

As discussed in our intuitive overview, the value of $y_{ij}^*$ determines the value of observed response, $y_{ij}$. Specifically, from (1) and (7), $y_{ij} = k$ if and only if:

$$\gamma_{k-1}^j < y_{ij}^* \leq \gamma_k^j \tag{8}$$

which gives the following augmented likelihood function over our model parameters and augmented response variable, $y^*$:

$$L(\mu, \delta, \gamma, \sigma^2, y^*) = \prod_i \prod_{j \in C_i} \phi\left(\frac{y_{ij}^* - (\mu_j + \delta_i)}{\sigma_j}\right) 1(\gamma_{y_{ij}-1}^j < y_{ij}^* \leq \gamma_{y_{ij}}^j) \tag{9}$$

Estimation of the model's structural parameters and latent variables is performed using the Markov Chain Monte Carlo (MCMC) algorithm outlined in the following sub-section.

## 2.2 Estimation Algorithm

We use the following Gibbs Sampler algorithm to sample draws from the conditional posterior distributions of the model's parameters and latent variables. All steps of the algorithm are performed sequentially in each iteration. An iteration is completed after Step 5, at which point the next iteration begins starting at Step 1.

**Step 1. Sample $\delta$.**

For $\delta = [\delta_1, ..., \delta_I]'$, let the estimation equation be given as:

$$y^* - \mu \otimes 1_I = \Lambda\delta + e$$

where $y^* = [y_1^*, ..., y_J^*]'$ is a stacked vector of $y^*$'s, $1_I$ is a $I \times 1$ vector of 1's, $\Lambda = I_I \otimes 1_J$, $I_I$ is an $I \times I$ identity matrix, $1_J$ is a $J \times 1$ vector of 1's, $e = [e_1, ..., e_J]'$, and $\otimes$ denotes the Kronecker product.

All elements of $\delta$ are drawn simultaneously from the following full conditional posterior distribution:

$$\delta|\mu, \sigma^2, y^*, Y \sim N(d, D) \tag{10}$$

where $D = \left[I_I + \Lambda'(I_I \otimes \Sigma^{-1})\Lambda\right]^{-1}$, $d = D\left[\Lambda'(I_I \otimes \Sigma^{-1})(y^* - \mu \otimes 1_I)\right]$, and $\Sigma$ is a $J \times J$ variance-covariance matrix with diagonal elements $(\sigma_1^2, ..., \sigma_J^2)$ and zeros for the off diagonal elements.

**Step 2. Sample elements of $\mu$.**

For each element of $\mu = [\mu_1, ..., \mu_J]'$, the estimation equation is given as:

$$y_j^* - \delta = 1_I\mu_j + e_j$$

where $y_j^*$ is a stacked vector of length $I$ containing all households' continuous response variables for question $j$.

Each $\mu_j$ is then drawn from the following full conditional posterior distribution:

$$\mu_j|\delta, \sigma^2, y^*, Y \sim N(v, V) \tag{11}$$

where $V = \left(\frac{1}{\sigma_\mu^2} + \frac{1_I' 1_I}{\sigma_j^2}\right)^{-1}$ and $v = V\left[\frac{\bar{\mu}}{\sigma_\mu^2} + \frac{1_I'(y_j^* - \delta)}{\sigma_j^2}\right]$.

**Step 3. Sample elements of $\gamma$.**

Samples for each element $\gamma_k^j$ of $\gamma^j = [\gamma_0^j, \gamma_1^j, ..., \gamma_{C_j}^j]$, such that $k = 2, ..., C_j - 1$, are drawn from the following full conditional posterior distribution:

$$\gamma_k^j|\gamma_{-k}^j, y^*, Y \sim U(L, R) \tag{12}$$

where $L = max[max(y_j^*|y_{ij} = k), \gamma_{k-1}^j]$, and $R = min[min(y_j^*|y_{ij} = k+1), \gamma_{k+1}^j]$.

**Step 4. Sample elements of $y^*$.**

Samples for each element $y_{ij}^*$ of $y^* = [y_1^*, ..., y_J^*]'$ are drawn from the following full conditional posterior distribution:

$$y_{ij}^*|\mu, \delta, \sigma^2, \gamma, Y \sim N\left(\mu_j + \delta_i, \sigma_j^2\right) \tag{13}$$

truncated to the interval $(\gamma_{y_{ij}-1}^j, \gamma_{y_{ij}}^j)$.

**Step 5. Sample elements of $\sigma^2$.**

For each element of $\sigma^2 = [\sigma_1^2, ..., \sigma_J^2]'$, the estimation equation is given as:

$$y_j^* = 1_I \mu_j + \delta + e_j$$

Each $\sigma_j^2$ is then drawn from the following full conditional posterior distribution:

$$\sigma_j^2 | \mu, \delta, \gamma, Y \sim IG(a, b) \tag{14}$$

where $a = \alpha + \frac{I}{2}$, and $b = \left[\frac{1}{\beta} + \frac{(y_j^* - 1_I \mu_j - \delta)'(y_j^* - 1_I \mu_j - \delta)}{2}\right]^{-1}$.

With the model and Gibbs Sampler algorithm fully defined, we first verify model performance using simulated data.[17] Specifically, we first simulate a set of binary and polytomous response variables matching the properties of the FSM. To accomplish this, we define values for each parameter in our BGRM and use the model to generate simulated responses to 10 questions (5 polytomous and 5 binary) for 5,000 households. We then estimate our model using the simulated data, running the Gibbs Sampler algorithm for 30,000 iterations with the first 10,000 iterations removed for burn-in. After estimation, we test to see if our model accurately converges to the known data-generating parameter values used to create our simulated data. Detailed information regarding the simulated data exercise and its results can be found in Appendix A. The results of our simulated data exercise suggest that the BGRM performs well in standard parameter retrieval tests found in the related literature.

# 3 Data

Our primary analysis uses data from the 2017 - 2018 waves of the National Health and Nutrition Examination Survey (NHANES).[18] The NHANES is one of the major programs conducted by

---

[17] All model estimation is conducted in the MATLAB programming language. All data and code used in this study are available from the authors upon request.

[18] While more recent waves of the NHANES are available at the time of this study, we use the 2017-2018 waves to avoid potential data collection or response concerns resulting from the COVID-19 pandemic.

the National Center for Health Statistics (NCHS), providing data from a nationally representative sample of roughly 5,000 individuals from the U.S. in each wave and for 2017-18 includes oversamples of Hispanic persons, non-Hispanic black persons, non-Hispanc Asian persons, non-Hispanic whites with incomes less than 185% of the federal poverty line, and Non-Hispanic white persons and persons of other races and ethnicities aged 0–11 years or 80 years and over.[19] The NHANES includes questions regarding the demographic, dietary, socioeconomic, and health-related characteristics of adults and children. In addition to self-reported variables provided by NHANES respondents, trained NHANES medical personnel collect medical, dental, and physiological measurements and administer laboratory tests. The NHANES is used to assess the prevalence of major diseases, disease risk factors, nutritional status, and their associations with health promotion and disease prevention efforts.

Most importantly for our study, the NHANES includes the full FSM questionnaire, data on important health outcomes commonly associated with food security, and nutrition quality , namely the HEI. This combination of data allows us to estimate our BGRM (with and without HEI) and assign food security categories based on the USDA's FSM scale. We then estimate each measures' ability to capture variation in these associated health outcomes, specifically obesity, diabetes, high blood pressure, and a self-reported indicator variable of following a good diet.

Our final 2017-2018 NHANES sample consists of 4,738 adults, even though there is variation in the number of observations across variables due to missingness. Table 1 provides descriptive statistics for our NHANES sample. At 63%, most adults in the sample are considered fully food secure under the standard USDA FSM scale categories. The remaining shares of marginal, low, and very low food secure adults are 14%, 13%, and 10%, respectively. For health outcomes, 37% of adults are obese; 16% have been diagnosed with diabetes; 38% have been diagnosed with high blood pressure; and 67% report following a good diet. The mean HEI score for adults in our sample is roughly 51 with a standard deviation of 13.68 points. Females make up 51% of the

[19] For additional information regarding the 2017 - 2018 NHANES, see NHANES Documentation 2017 - 2018.

sample. White and Black are the most commonly reported races in our sample at 36% and 23%, respectively. Hispanics, Asians, and other races/ethnicities represent 22%, 13%, and 6% of our NHANES sample, respectively. 88% of adults in the sample are U.S. citizens. 24% of the sample hold a bachelors degree or higher. 33% have attended some college or hold a two-year associates degree. 24% of the sample have a high school degree or equivalent, and the remaining 19% have less than a high school degree. Finally, 50% of the adults in the sample are married. The second largest group by marital status are never married adults at 18%. The remaining respondents are either widowed (8%), divorced (11%), living apart (9%), or separated (4%).

# 4 Results

## 4.1 Main Results

Using the full NHANES sample, we run our BGRM Gibbs Sampler algorithm for 30,000 total iterations. To ensure adequate time for convergence, we remove the first 10,000 draws from the final set of estimates in a process known as "burn-in", leaving us with 20,000 post-burn-in draws from each conditional posterior distribution. Figure 1 displays the BGRM posterior mean food security (FSEC) estimates for each adult arranged in ascending order, along with colored lines representing the 95% credible intervals of each adult's posterior FSEC distribution.[20] We find that the level of uncertainty is lowest for households with mean FSEC values towards the middle of the distribution and highest for adults with mean FSEC values in the tails.

The "hourglass" pattern of uncertainty in the posterior mean adult FSEC values has two primary causes. First, our model contains responses from what are often referred to as "extreme response"

---

[20] Our BGRM produces 20,000 simulated FSEC draws from each household's conditional posterior distribution. We therefore use the posterior mean of each household's distribution as our point estimates. 95% credible intervals represent the range from which 95% of a household's 20,000 FSEC values are drawn, representing a measure of estimation uncertainty/variability.

households. Specifically, extreme response households are those with all negative ("no") or all affirmative ("yes") responses to the set of FSM questions.[21] Since most IRT models rely on variation in household responses across questions to identify the value of latent food security, extreme response households cannot be included. As discussed in Nord (2012) for example, the USDA's 1-PRM omits households with extreme FSM responses from the sample prior to estimation as the model cannot identify their latent trait without within-household-variation across responses. Instead, only households with at least some level of response variation are used in estimation. Alternatively, while our model produces point estimates for extreme response households that match *a priori* expectations, the credible intervals surrounding these point estimates are substantial given the limited amount of information available aside from which end of the distribution they fall in. Second, since the value of each extreme response household's $y^*$ is drawn from different portions of a truncated normal distribution based on observed responses, these households only have draws of $y^*$ from the left or right tails of the distribution rather than areas in the middle (for middle responses to polytomous items). This increases the likelihood of extreme $y^*$ draws as these portions of the distribution are bounded by either $-\infty$ or $\infty$. In turn, high values of $y^*$ affect variability in the posterior mean of latent food security across iterations, increasing overall uncertainty. Even with a significantly larger sample size, we find a qualitatively similar distribution of 95% credible intervals in our CPS exercise shown in Appendix D, though the credible intervals are more consistent across households in the CPS.

From an intuitive standpoint, the use and interpretation of our measure of uncertainty for extreme response households should be treated appropriately. Simply put, if an individual responds "yes" to all FSM questions, we are suitably confident that adults in their household are food insecure, and if "no" to all questions, food secure. This is a case where the statistical challenges inherent to working with limited response data clash with *a priori* beliefs. As opposed to the 1-PRM which

[21] For polytomous FSM questions, a negative extreme response household would have the lowest possible response to all questions while an affirmative extreme response household would have the highest possible response to all questions.

omits extreme response households, our BGRM can still provide point estimates for these households, even though the usefulness of their uncertainty estimates is limited. Where uncertainty is truly useful to consider, however, is for households with mixed responses towards the middle of the overall distribution. Uncertainty for these households helps us capture concerns regarding the relative impact of the FSM's polytomous items. For example, uncertainty if a response of "sometimes" to two FSM questions is more or less impactful than a response of "often" to one question. While we do not conduct this analysis, our uncertainty measures could easily be used to probabilistically assign households to various food security categories similar to the approach taken by Nord (2012). Extreme households would likely need to be deterministically assigned to the highest or lowest categories while households with mixed responses can be assigned a probability of being in each category based on the proportion of their 95% credible intervals that fall within the category's range.

To examine the relationship between our estimates and standard USDA food security categories, Figure 2 shows the distribution of BGRM posterior mean food security estimates separated by USDA-defined category. Each household in our NHANES sample is assigned to one of four categories following the USDA's methodology, specifically: full food secure, marginal food secure, low food secure, and very low food secure.[22] We then plot the histograms of BGRM posterior mean FSEC values for households in each category, showing how the distribution of BGRM estimates varies within and across official USDA categories. As expected, the rank order is consistent across our BGRM point estimates and the USDA food security categories. Adults assigned to the full food secure category under the FSM scale have the highest BGRM posterior mean values followed by the marginal, low, and very low food secure categories, respectively. The distribution of BGRM posterior mean estimates is roughly normal for full food secure adults given that their identical response patterns offer little information to differentiate households within the category. The dis-

[22] The USDA food security categories are defined such that Full Food Secure households have 0 affirmative responses to the adult FSM questionnaire, followed by 1-2, 3-5, and 6-10 affirmative responses for Marginal, Low and Very Low Food Secure households, respectively. For additional information, see Rabbitt et al. (2023).

tribution of BGRM estimates for full food secure households also has no overlap with the other three categories, implying that posterior means for households in the full category are distinctly separate from the lower food security categories.[23] We also find significant variation in BGRM posterior mean values within USDA food security categories, highlighting potential differences in severity levels for households assigned to the same category.

Perhaps most notably, we find meaningful overlap in the distributions of our BGRM posterior means across the USDA's marginal, low, and very low food security categories. The USDA categories implicitly assume no overlap across categories. While we do not find overlap between the full and marginal categories, overlap across the lower three categories suggests that the assumption of mutually independent categories may not hold. With overlap across categories, the food security status of households may be over or under estimated when relying on traditional USDA category assignment. For example, a nontrivial portion of the BGRM posterior means for households in the left tail of the marginal category fall below the means of some households in the low food secure category's right tail. Our results therefore suggest that certain households may be misclassified as belonging to a higher or lower USDA category than their posterior mean BGRM values would indicate. Furthermore, this overlap supports the potential benefits of probabilistic, rather than deterministic, food security category assignment, particularly for households with point estimates near the category thresholds. Finally, Figure 2 shows some visually identifiable areas of high density in our BGRM estimates within existing USDA categories which could prove meaningful. For example, the distribution of BGRM posterior mean estimates for households in the marginal

[23] In most applications, households are considered food secure if they have full or marginal food security and food insecure with low or very low food security. The empty region between the marginal and food secure group shown in Figure 2, however, suggests that marginal food secure households have posterior means closer to the low group than the full group. This finding aligns with Cook et al. (2013), Coleman-Jensen (2010), and the related literature cited within finding that health outcomes for households in the marginal group may be closer to the outcomes for low food secure households than full food secure households. Additionally, as the 95% credible intervals shown in Figure 1 indicate, while no households have posterior mean values in the empty range between full and marginal in Figure 2, some latent food security values are sampled from the range across iterations.

category appears qualitatively bimodal. Examining these areas of high density may be a promising avenue for future research identifying new potential food security categories or refining existing ones. Furthermore, as shown in Figure D.2, we find a similar bimodal quality in the marginal food secure category using our CPS sample. This consistency is key as the USDA uses the CPS to calculate national food security rates each year. While doing so is outside the scope of this study, careful consideration of these distributional features and how they relate to category definitions is valuable given their critical research and policy relevance.

Finally, as shown in Appendix B, we examine the degree to which our BGRM posterior mean FSEC estimates capture variation in health and diet outcomes commonly associated with food security. Specifically, we estimate regressions of four binary variables, self-reported obesity, diagnosis of diabetes, diagnosis of high blood pressure, and following a good diet, on our BGRM posterior mean point estimates and a set of variables controlling for gender, age, race, level of education, marital status, citizenship status, household size, and annual household income. We compare the resulting BGRM regression estimates to those from regressions of health outcomes on binary category indicators of marginal, low, and very low food security defined using the USDA's FSM scale, leaving full food security as the reference group; the number of raw affirmative FSM responses for each adult in the sample; and the respondent-level probabilities of falling within a given quartile of the overall BGRM FSEC posterior distribution.

Results from these regressions are shown in Table B.1. We find that the direction of associations between our BGRM posterior means and the four health outcomes of interest matches *a priori* expectations. Specifically, higher BGRM posterior mean values (indicative of better food security) are associated with decreased probabilities of obesity, diabetes, and high blood pressure, and an increased probability of following a good diet. These coefficients are also statistically significant at the 1% level for all outcomes. When comparing the BGRM regression results to those using USDA food security category indicators, the explanatory power measured using adjusted R-squared is quite similar. Therefore, the point estimates generated by our BGRM produce similar results to those using the set of USDA categories, implying that there is no notable loss in explanatory power from our more flexible, but complex, BGRM.

We use the NHANES to estimate our primary results as the survey includes both the health outcome variables and HEI (which we utilize in Section 4.2 below). As previously discussed, however, official USDA national food security estimates come from the CPS FSS. To evaluate the robustness of BGRM estimates and examine how results for the CPS differ from those of the NHANES, we estimate our BGRM using the 2018 CPS FSS. We discuss this CPS exercise in more detail in Appendix D.

From Figure D.1, a CPS analog to Figure 1, while we find significant uncertainty in the CPS estimates, the distribution of uncertainty is more consistent across households in the CPS relative to the NHANES. This improved consistency is most likely due to the CPS's increased sample size. While more consistent, our results generally suggest that national estimates of food security are likely subject to non-trivial levels of uncertainty, a point raised by previous studies like the National Research Council (2006). Perhaps more notably, however, are the distributional features of our CPS point estimates separated by USDA-defined food security categories shown in Figure D.2. Overall, these features are very qualitatively similar to the NHANES estimates shown in Figure 2 which is expected as both surveys provide large, nationally representative samples. Like our NHANES estimates, we find overlap in BGRM posterior means across USDA categories in the CPS. Depending on their level of uncertainty, many households have 95% credible intervals that stretch across thresholds of USDA categories. Again, this finding supports using measures of estimation uncertainty to probabilistically assign food security categories as discussed in National Research Council (2006) and Nord (2012). Probabilistic category assignment could in turn have direct implications for targeted interventions and policies based on food security categories, either at the household or aggregate-level. Finally, as discussed previously, the overall distributional features of BGRM point estimates from the CPS in Figure D.2 are strikingly similar to the NHANES results shown in Figure 2. We view these similarities as a demonstration of robustness, both for our model and the FSM questionnaire in general.

## 4.2 Estimation Including a Continuous Variable: The Healthy Eating Index

One of the main advantages of our BGRM is its ability to include a mix of different observable variable types, even continuous variables. For example, the 1-PRM restricts the set of potential variables to binary variables. While extensions of the 1-PRM like the PCRM used in Nord (2012), Tanaka et al. (2020), and Opsomer et al. (2002) allow for combinations of binary and ordered polytomous variables, it cannot be estimated using continuous variables. With some minor modifications, however, our BGRM can be used to estimate latent traits with any mix of binary, ordered polytomous, and continuous variables. We highlight this flexibility by estimating our model using both the set of adult FSM questions and an additional continuous variable: the Healthy Eating Index (HEI). This application produces a new construct that combines adult food security and diet quality information into a single latent trait.

The HEI is a scoring metric that provides information regarding individuals' overall diet quality as well as the quality of specific dietary components. The HEI can be used regardless of food consumption quality, and is adequate to assess compliance with U.S. Dietary Guidelines for Americans (DGAs). Since 2005, HEI has been continuously updated by researchers from the U.S. Department of Health and Human Services' (DHHS's) National Cancer Institute (NCI) and the USDA to account for DGA revisions (National Cancer Institute (2023)).

We calculate composite HEI for all adults in our NHANES sample. We then estimate our BGRM using the 10 adult FSM questions and continuous HEI. Since HEI is a continuous variable, including it in our BGRM requires slight modifications to the model and Gibbs Sampler algorithm. We discuss these changes in detail in Appendix C. Similar to our primary results, the BGRM Gibbs Sampler algorithm with FSM and HEI is run for 30,000 iterations, removing the first 10,000 draws for burn-in.

Figure 3 shows the distribution of BGRM latent trait estimates for our novel FSM/HEI construct separated by USDA-defined FSM categories. Similar to Figure 2 made without HEI, Figure 3 shows overlap across USDA food security categories and significant within-category variation in BGRM means. Additionally, the distribution of our new latent construct is qualitatively similar to the distribution of BGRM food security means without HEI. For example, we find that the distribution of BGRM means for the new latent construct appears bimodal for households in the USDA's marginal food secure group after adding HEI.[24]

We then estimate associations between our new latent construct with HEI and the set of health outcomes using the same set of regressions shown in Table B.1. Assuming that the HEI captures variation in diet quality not already reflected in the FSM, we expect these new BGRM estimates to better explain variation in health outcomes. These results are included in Table C.1.[25]

The estimates shown in Table C.1 provide several implications of note. First, the BGRM posterior mean coefficients for our new latent construct with HEI increase in magnitude across all health outcome regressions and remain statistically significant at the 1% level of significance. The coefficients on our BGRM posterior means for obesity and good diet are statistically different from those found without HEI shown in Table B.1 at the 10% and 5% level, respectively. Second, we find a notable change in the magnitude and level of statistical significance for BGRM quartile probability coefficients after including HEI. The quartile probability coefficients for obesity more than double in magnitude with HEI and become statistically significant. We detect similar changes in these coefficients for high blood pressure and good diet. BGRM quartile probability coefficients are statistically insignificant for diabetes with and without HEI. We also find significant increases in the explanatory power of our BGRM posterior mean and quartile probabilities with HEI for all health outcome regressions measured using adjusted R-squared. These results indicate that the additional information provided by the HEI accounts for more of the variation in our health outcomes of interest than the adult FSM questions alone. The increase in explanatory power is non-trivial, especially for the regressions of obesity and following a good diet.

---

24 We also compute a counterpart to Figure 1 using our estimates with HEI. The counterpart to Figure 1 is shown in Appendix C as Figure C.1. This figure does not change meaningfully after including HEI.

25 In table C.1, we compare the BGRM FSEC estimates from Table B.1 with those produced using the set of adult FSM questions and HEI.

While adding HEI to our BGRM in this study primarily serves to showcase model flexibility, estimating a new latent construct using both the discrete FSM and continuous HEI serves two additional purposes. First, as emphasized in sources like National Research Council (2006) and Opsomer et al. (2002), a common critique of the FSM is its lack of a direct nutrition quality measure. While the HEI is a strong choice for researchers who want to create a measure of food security with a nutrition quality component, it cannot be included in binary/polytomous IRT food security models like the 1-PRM and PCRM. Alternatively, our BGRM can be used with any mix of binary, ordered polytomous, and continuous variables which we see as a major advantage to future latent trait measurement. Second, given the overlapping domains of food security and nutrition security, we propose our BGRM as one suitable method for future nutrition security measurement research. For example, Seligman et al. (2023) create the first conceptual framework for nutrition security which they explicitly state is informed by the robust food security literature. Relatedly, Seligman et al. (2023) discuss how information from the HEI and FSM should be jointly considered for nutrition security, though no measurement tool currently exists. Given the continuous nature of HEI and other common nutrition variables, however, traditional IRT models used to measure food security like the 1-PRM may prove too restrictive. That said, we highlight the need for caution when interpreting the results of our HEI exercise. Specifically, while this new latent trait includes a composite measure of nutrition quality absent from the FSM, far more work is needed to define the latent trait and gauge its appropriateness as a measure; a task that is outside the scope of this study.[26] Furthermore, to reiterate, we do not propose this novel latent construct as an acceptable measure of nutrition security. As the theoretical nutrition security literature develops further (e.g., Seligman et al. (2023)), however, designing a statistical measure of nutrition security using flexible IRT models like the BGRM represents a promising avenue for future research.

[26] As an example of the similar work required to design the FSM questionnaire and USDA FSM scale, *see* Ohls et al. (2001).

# 5 Conclusion

In this study, we propose a novel measurement model for household food security and similar latent constructs: the Bayesian Graded Response Model (BGRM). Our BGRM is well suited to food security measurement for several reasons. First, the model provides continuous estimates of latent food security. This continuous measure provides more detailed information regarding households' food security levels which can support information from discrete food security categories like those from the USDA FSM scale. Second, the BGRM allows us to estimate food security using both binary and polytomous ordered response variables. Unlike the USDA's 1-PRM, we do not need to first convert each polytomous FSM question into a binary "affirmative/negative" type variable before estimating our BGRM. Importantly, this allows us to incorporate information about response severity at the intensive margin for the set of polytomous FSM questions. Finally, as a Bayesian model, the BGRM estimates draw from a posterior distribution of latent food security for each household. Using these posterior distributions, we produce both food security point estimates and measures of uncertainty at the household-level.

We estimate our BGRM with the set of 10 adult FSM questions using data from the 2017 - 2018 NHANES. We find significant uncertainty in our estimates across the distribution of households, with higher uncertainty in the lower and upper tails. As discussed in National Research Council (2006), food security measures should account for the large amount of uncertainty inherent to latent trait estimation, something our BGRM is well suited to. Examining the distribution of our BGRM posterior mean estimates across and within USDA-defined food security categories, our results suggest that there is significant variation in latent food security within USDA categories. Furthermore, we find nontrivial overlap in BGRM food security posterior means across USDA food security categories, implying that some households may be misclassified under current USDA categories. Given the significant uncertainty in our estimates, the 95% credible intervals of many households stretch across USDA category thresholds, supporting the use of probabilistic category assignment as proposed by National Research Council (2006). Building on Nord (2012), using our BGRM estimates to assign household-level food security category probabilities is a promising avenue for future research. Using data from the CPS FSS as a robustness check, we find qualitatively similar results as those from the NHANES. The results of our CPS exercise have direct implications for food security related policy as the data set is used to calculate official USDA food security/insecurity rates each year.

We also examine how well our BGRM food security estimates explain variation in health outcomes commonly associated with food security. Using our NHANES sample, we estimate regressions of obesity, diabetes, high blood pressure, and following a good diet on our BGRM posterior mean estimates and a set of variables measuring the probability of each household belonging to specific BGRM food security quartiles. We compare these results to those from regressions of the same health outcomes on binary indicators of marginal, low, and very low food security categories calculated using official USDA definitions, as well as the raw number of affirmative FSM responses. We find a statistically significant relationship between the BGRM posterior mean food security estimates and all health outcomes. Additionally, regression results are similar across the BGRM and USDA food security category regressions with improvements from the BGRM in some specifications, suggesting that no explanatory power is lost when using our more flexible, but complex, BGRM.

To the best of our knowledge, only three previous studies Nord (2012), Tanaka et al. (2020), and Opsomer et al. (2002) have used polytomous IRT models to estimate continuous food security. While the model used in those studies can accommodate a mix of binary and ordered polytomous variables, it cannot be used with continuous variables. Our flexible BGRM does not share this limitation. We demonstrate the flexibility of our model by estimating a new latent construct with an additional continuous variable. Specifically, with data from the NHANES, we estimate our model using both the 10 binary/ordered polytomous adult FSM questions and continuous HEI, a composite measure of nutrition quality. By adding or removing variables, researchers can evaluate different dimensions of food security or create novel latent constructs. As mentioned, however, many models previously used in the food security measurement literature place heavy restrictions on the types of variables that can be included. One of the main strengths of our BGRM is its ability to easily add or remove different types of observable variables with only simple modifications to the underlying model framework, allowing for any mix of binary, polytomous, and continuous variables.

Our HEI exercise produces several interesting insights. Adding HEI to our model increases the explanatory power of the BGRM posterior means and quartile probabilities for all health outcomes regressions relative to the BGRM estimates without HEI. Comparative performance between our BGRM estimates with HEI also improves relative to the affirmative FSM count and USDA food security category indicators. Alternatively, the distributional features of our estimates with and without HEI are quite similar, including the levels of household-level estimation uncertainty. In summary, our new latent construct created using the adult FSM questions and HEI better accounts for variation in health outcomes of interest relative to estimates produced using just the FSM. As discussed in Section 4.2, however, we caution interpretation of these results as more work needs to be done to accurately identify the meaning of this new latent trait and its appropriateness from a theoretical measurement perspective. Therefore, we primarily see the HEI exercise as a demonstration of model flexibility. Regardless, adding a nutrition quality dimension to food security measurement does address a common critique of the FSM as discussed in sources like National Research Council (2006).

While we view this paper as a strong initial application of our BGRM to measuring latent traits like food security, additional work is needed in a few major areas. First, the current BGRM Gibbs Sampler estimation algorithm is computationally intensive. While we were able to estimate our model without special computing resources using large samples of data from the NHANES and CPS, improving computational efficiency would prove beneficial. This is especially true for future users with limited computing resources. If the model can be estimated more efficiently, the method would be increasingly approachable for the target audience of food security researchers. Additionally, while the MATLAB code used to estimate the BGRM is readily available from the authors upon request, building packages for more common statistical programs with Bayesian functionality like R, Python, or STATA would further expand adoption.

The second area where additional work is needed relates to expanding BGRM analysis for food security measurement. Our BGRM can be used to measure food security with any survey that includes the FSM questionnaire. Applying our model to other surveys would help determine variation in measurement and parameter estimates across data sets. Additionally, more work is needed to understand how BGRM FSEC estimates vary by individual, household, and community-level characteristics. For example, related to previous work by Tanaka et al. (2019) future work could estimate our BGRM for SNAP and non-SNAP households, identifying if both groups produce significantly different latent trait estimates. Furthermore, our study only estimates adult food security using the 10 adult FSM questions. Future work should measure child food security separately or create a combined measure which includes both the set of adult and child FSM questions. This avenue for future research builds on previous work regarding the dimensionality of latent food security (Rabbitt et al. (2020) , Tanaka et al. (2020) , Coleman-Jensen et al. (2017)). Additional study is also needed to understand how BGRM-derived estimates of food security relate to outcomes outside of physical health, including economic outcomes, mental health, and food purchasing behaviors. While we estimate regressions of health outcomes on our BGRM food security estimates, identifying the effects of various factors on latent trait estimates is key to understanding causal factors that affect a household's food security level. The most straightforward approach for this analysis is to first estimate the latent trait with an IRT model like the BGRM and then including the trait estimates as the dependent variable of a regression. Alternatively, the results of Rabbitt (2018) suggest that jointly estimating the latent trait and a latent regression model in a single IRT framework improves precision and accuracy for cases with endogenous covariates. Rabbitt (2018) demonstrate this approach for endogenous household SNAP participation, but future work should consider the joint estimation strategy for similar applications.

Finally, using our BGRM to create a new latent construct with both the adult FSM questions and HEI provides a potentially suitable measurement method for other latent traits, especially noteworthy for the rapidly developing nutrition security literature. For example, Seligman et al. (2023)

create the first conceptual model of nutrition security which they explicitly state is informed by the robust food security literature. At the time of this study, no statistical measurement model exists for nutrition security. While we in no way claim that including HEI alongside the set of adult FSM questions represents a valid measure of nutrition security, our results do suggest that a measure combining information on food security and overall diet quality may capture additional variation in health outcomes relative to measures of food security alone. Seligman et al. (2023) discuss how information from the HEI and FSM should be jointly considered for nutrition security. Regardless, the similarities between food and nutrition security suggest that future nutrition security researchers will likely look to the robust food security literature for possible measurement models. Similar to food security, any eventual measure of nutrition security must distill multiple observable variables into a single latent trait. Unlike the FSM's binary and polytomous questions, however, many nutrition variables like HEI are continuous. These variables can therefore not be included in IRT models previously used to measure food security like the binary 1-PRM and polytomous PCRM, heavily restricting the set of usable variables. Alternatively, any number of continuous variables can be included in our flexible BGRM, making it well suited to measuring a wide array of latent traits like nutrition security.

# References


Andrich, D. (1988). *Rasch models for measurement*. SAGE.

Andrich, D. and Marais, I. (2019). *A course in Rasch measurement theory*. Springer.

Baker, F. B. (2001). *The basics of item response theory*. ERIC.

Barrett, C. B. (2010). Measuring food insecurity. *Science*, 327(5967):825–828.

Bickel, G., Nord, M., Price, C., Hamilton, W., Cook, J., et al. (2000). Guide to measuring household food security.

Bickel, G. W., Hamilton, W. L., et al. (1997). Household food security in the united states in 1995: technical report of the food security measurement project.

Bock, R. D. and Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an em algorithm. *Psychometrika*, 46(4):443–459.

Broomell, S. B. and Bhatia, S. (2014). Parameter recovery for decision modeling using choice data. *Decision*, 1(4):252.

Casella, G. and George, E. I. (1992). Explaining the gibbs sampler. *The American Statistician*, 46(3):167–174.

Coleman-Jensen, A., Rabbitt, M. P., and Gregory, C. A. (2017). Examining an" experimental" food security status classification method for households with children.

Coleman-Jensen, A. J. (2010). U.S. food insecurity status: toward a refined definition. *Social Indicators Research*, 95:215–230.

Cook, J. T., Black, M., Chilton, M., Cutts, D., De Cuba, S. E., Heeren, T. C., Rose-Jacobs, R., Sandel, M., Casey, P. H., Coleman, S., et al. (2013). Are food insecurity's health impacts underestimated in the U.S. population? marginal food security also predicts adverse health outcomes in young US children and mothers. *Advances in Nutrition*, 4(1):51–61.

Fox, J.-P. (2010). *Bayesian item response modeling: Theory and applications*. Springer.

Gregory, C. A. (2020). Are we underestimating food insecurity? partial identification with a bayesian 4-parameter irt model. *Journal of Classification*, 37:632–655.

Gundersen, C., Kreider, B., and Pepper, J. (2011). The economics of food insecurity in the united states. *Applied Economic Perspectives and Policy*, 33(3):281–303.

Gundersen, C., Waxman, E., and Crumbaugh, A. S. (2019). An examination of the adequacy of supplemental nutrition assistance program (snap) benefit levels: impacts on food insecurity. *Agricultural and Resource Economics Review*, 48(3):433–447.

Gundersen, C. and Ziliak, J. P. (2015). Food insecurity and health outcomes. *Health affairs*, 34(11):1830–1839.

Hambleton, R. K. (1989). Principles and selected applications of item response theory. *Educational Measurement Third Edition.*

Johnson, V. E. and Albert, J. H. (2006). *Ordinal data modeling*. Springer Science & Business Media.

Kieftenbeld, V. and Natesan, P. (2012). Recovery of graded response model parameters: A comparison of marginal maximum likelihood and markov chain monte carlo estimation. *Applied Psychological Measurement*, 36(5):399–419.

Moffitt, R. A. and Ribar, D. C. (2016). Rasch analyses of very low food security among households and children in the three city study. *Southern Economic Journal*, 82(4):1123–1146.

Muthén, B. (1979). A structural probit model with latent variables. *Journal of the American Statistical Association*, 74(368):807–811.

National Cancer Institute, N. (2023). Overview & background of the healthy eating index. https://epi.grants.cancer.gov/hei/. Accessed: 2023-06-23.

National Research Council, N. (2006). Food insecurity and hunger in the united states: An assessment of the measure. Washington, DC. The National Academies Press. https://doi.org/10.17226/11578.

Nord, M. (2012). Assessing potential technical enhancements to the U.S. household food security measure. Technical report, United States Department of Agriculture.

Ohls, J., Radbill, L., and Schirm, A. (2001). Household food security in the United States, 1995-1997: Technical issues and statistical report. Technical report, Mathematica Policy Research.

Opsomer, J. D., Jensen, H. H., Nusser, S. M., Drignei, D., and Amemiya, Y. (2002). Statistical considerations for the usda food insecurity index.

Rabbitt, M. P. (2014). Measuring the effect of supplemental nutrition assistance program participation on food insecurity using a behavioral rasch selection model. *Unpublished paper, University of North Carolina at Greensboro*.

Rabbitt, M. P. (2018). Causal inference with latent variables from the rasch model as outcomes. *Measurement*, 120:193–205.

Rabbitt, M. P. and Coleman-Jensen, A. (2017). Rasch analyses of the standardized spanish translation of the U.S. household food security survey module. *Journal of Economic and Social Measurement*, 42(2):171–187.

Rabbitt, M. P., Engelhard, G., and Jennings, J. K. (2020). Assessing the dimensionality of food-security measures. *Journal of Economic and Social Measurement*, 45(3-4):183–213.

Rabbitt, M. P., Hales, L. J., Burke, M. P., and Coleman-Jensen, A. (2023). Household food security in the United States in 2022.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Danmarks pædagogiske Institut.

Samejima, F. (1968). Estimation of latent ability using a response pattern of graded scores. *ETS Research Report Series*, 1968(1).

Seligman, H. K., Levi, R., Adebiyi, V. O., Coleman-Jensen, A., Guthrie, J. F., and Frongillo, E. A. (2023). Assessing and monitoring nutrition security to promote healthy dietary intake and outcomes in the United States. *Annual Review of Nutrition*, 43:409–429.

Smith, T. A. and Gregory, C. A. (2023). Food Insecurity in the United States: Measurement, Economic Modeling, and Food Assistance Effectiveness. *Annual Review of Resource Economics*, 15(Volume 15, 2023):279–303. Publisher: Annual Reviews Type: Journal Article.

Tanaka, V. T., Engelhard, G., and Rabbitt, M. P. (2019). Examining differential item functioning in the household food insecurity scale: Does participation in snap affect measurement invariance. *Journal of Applied Measurement*, 20(1):100–111.

Tanaka, V. T., Engelhard, G., and Rabbitt, M. P. (2020). Modeling household food insecurity with a polytomous rasch model. In *Quantitative Psychology: 84th Annual Meeting of the Psychometric Society, Santiago, Chile, 2019 84*, pages 319–331. Springer.

Zhu, X. and Stone, C. A. (2011). Assessing fit of unidimensional graded response models using bayesian methods. *Journal of Educational Measurement*, 48(1):81–97.

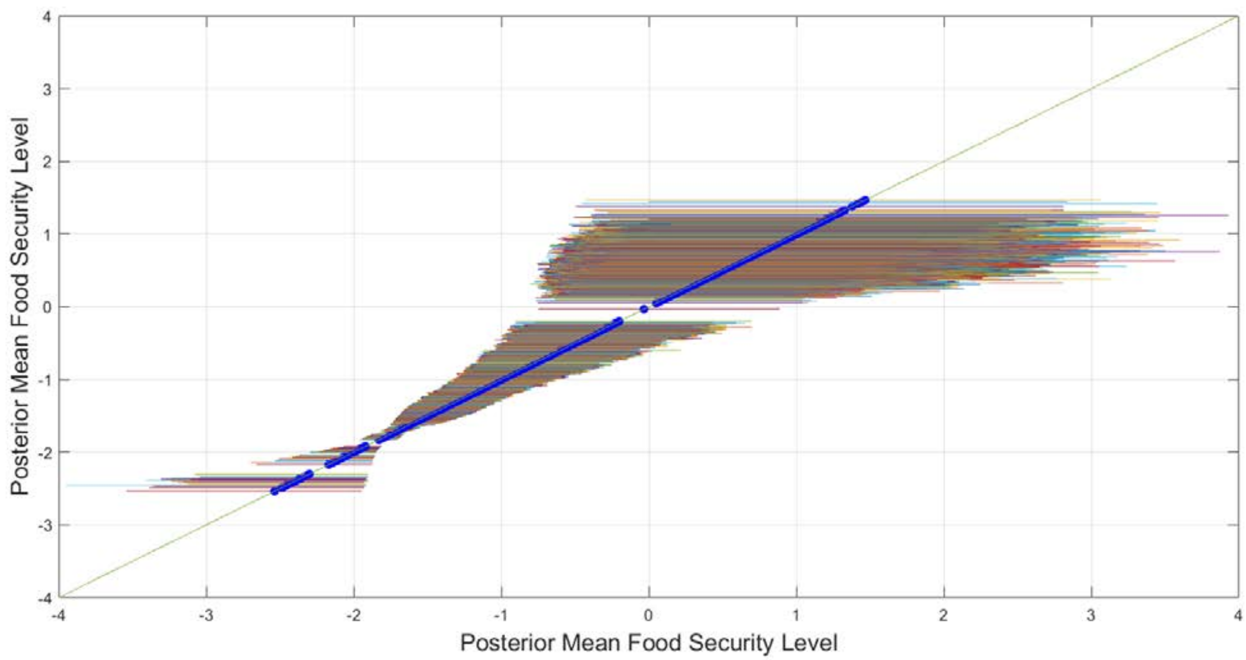
# TABLES & FIGURES

**Table 1:** Summary Statistics 2017 - 2018 National Health and Nutrition Examination Survey (NHANES) Sample

| | **Years:** 2017 - 2018 | | |
|---|---|---|---|
| | Mean | Standard Deviation | Count |
| **Food Security** | | | |
| Full Food Secure Households | 0.63 | 0.48 | 4738 |
| Marginal Food Secure Households | 0.14 | 0.35 | 4738 |
| Low Food Secure Households | 0.13 | 0.34 | 4738 |
| Very Low Food Secure Households | 0.10 | 0.29 | 4738 |
| **Health Outcomes** | | | |
| Obese | 0.37 | 0.48 | 4059 |
| Diabetes | 0.16 | 0.37 | 3967 |
| High Blood Pressure | 0.38 | 0.49 | 4082 |
| Good Diet | 0.67 | 0.47 | 4090 |
| **Healthy Eating Index (HEI)** | | | |
| HEI | 51.39 | 13.68 | 4738 |
| **Gender** | | | |
| Female | 0.51 | 0.50 | 4738 |
| **Race** | | | |
| Hispanic | 0.22 | 0.42 | 4738 |
| White | 0.36 | 0.48 | 4738 |
| Black | 0.23 | 0.42 | 4738 |
| Asian | 0.13 | 0.34 | 4738 |
| Other | 0.06 | 0.22 | 4738 |
| **Citizenship** | | | |
| Citizen | 0.88 | 0.33 | 4726 |
| **Education Level** | | | |
| Less than High School | 0.19 | 0.39 | 4505 |
| High School | 0.24 | 0.43 | 4505 |
| Some College & Associate Degree | 0.33 | 0.47 | 4505 |
| College Graduate & Above | 0.24 | 0.43 | 4505 |
| **Marital Status** | | | |
| Married | 0.50 | 0.50 | 4508 |
| Widowed | 0.08 | 0.27 | 4508 |
| Divorced | 0.11 | 0.32 | 4508 |
| Separated | 0.04 | 0.19 | 4508 |
| Never Married | 0.18 | 0.38 | 4508 |
| Living Apart | 0.09 | 0.29 | 4508 |

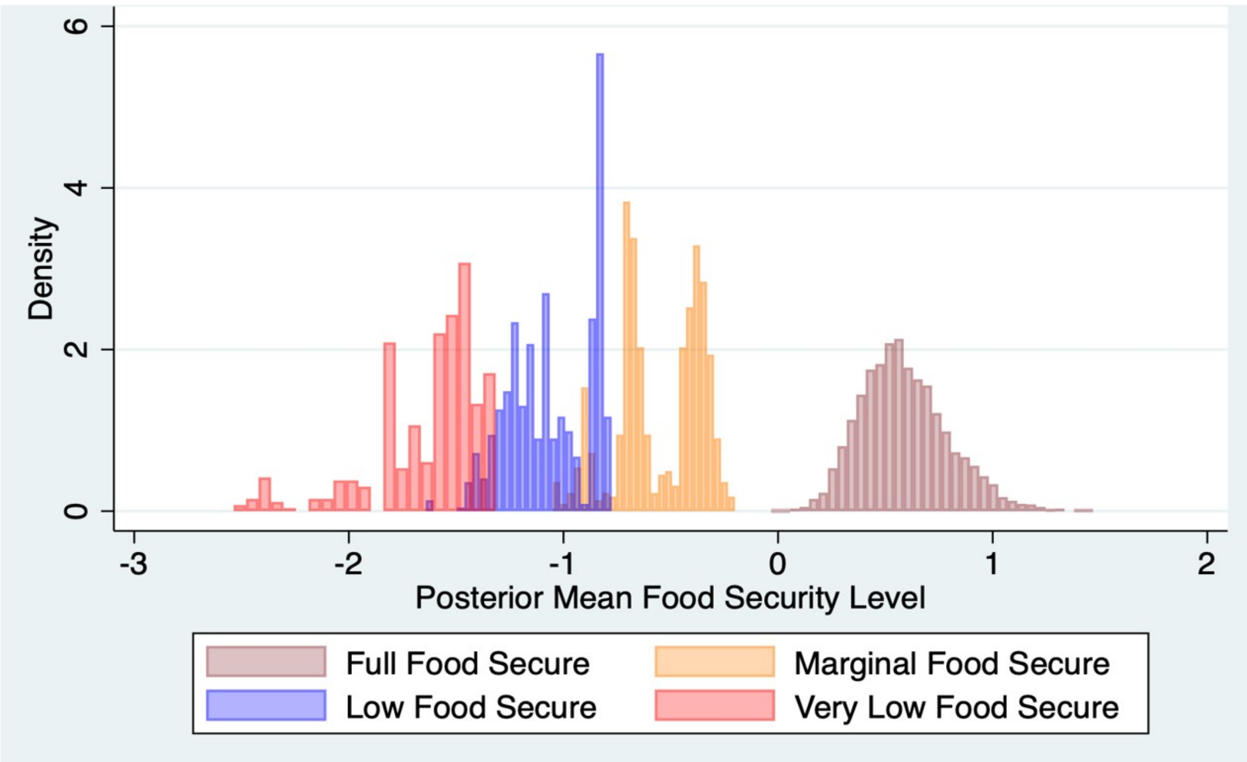
**Notes:** The sample consists of adults from the 2017 - 2018 National Health and Nutrition Examination Survey (NHANES). The entire sample contains 4,738 adults. Number of observations vary across variables due to missing values. Food security categories follow USDA definitions. Full Food Secure households have 0 affirmative responses to the adult FSM questionnaire, followed by 1-2, 3-5, and 6-10 affirmative responses for Marginal, Low and Very Low Food Secure households, respectively. HEI is a continuous variable. All other variables are indicator variables.

**Figure 1:** Posterior Mean Values of BGRM Latent Food Security with 95% Credible Intervals using NHANES 2017 - 2018



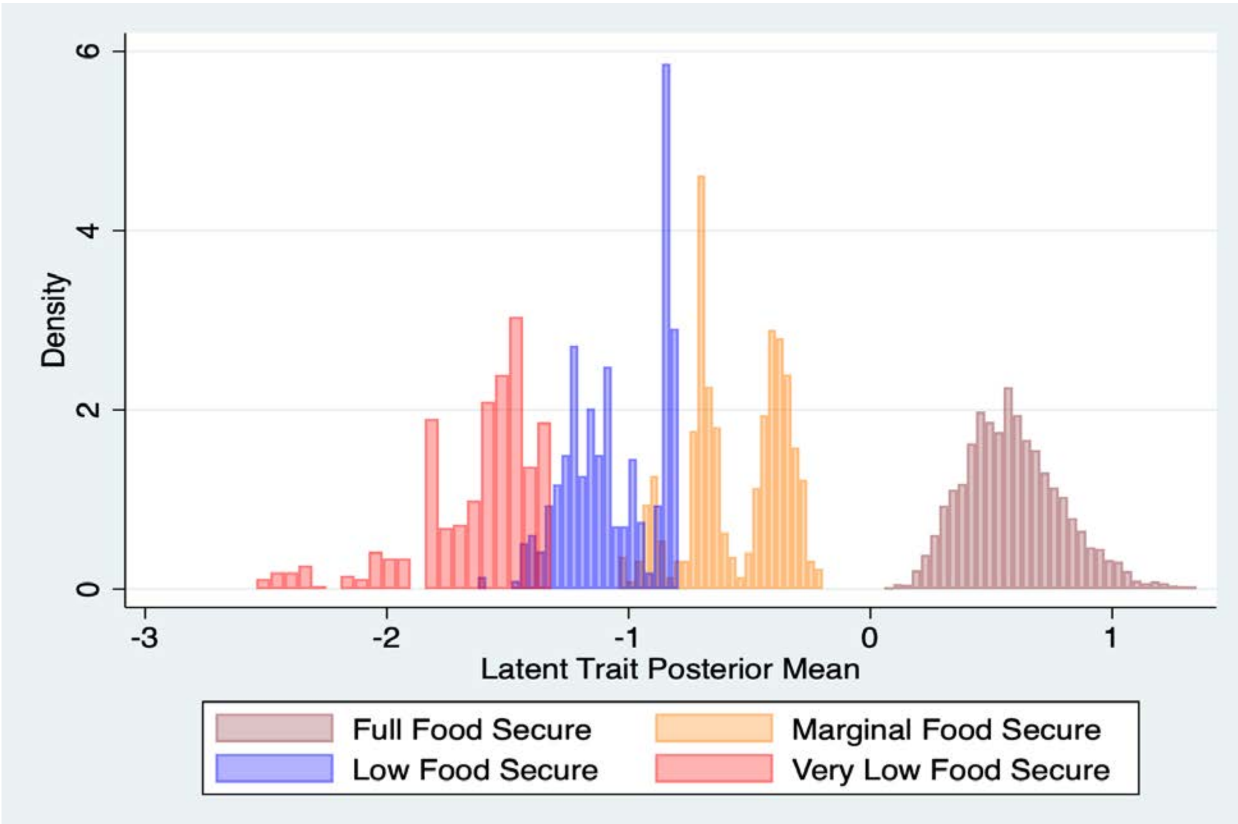

**Note:** The sample consists of households with responses to the 10 adult FSM questions from the 2017 - 2018 National Health and Nutrition Examination Survey (NHANES). The entire sample contains 4,738 households. Each blue dot on the 45$^o$ degree line represents the BGRM posterior mean latent food security value for a given household in the sample. The colored lines running through the blue dots represent the magnitude of the 95% credible intervals for each household.

**Figure 2:** Distribution of BGRM Posterior Mean Food Security by USDA FSM Scale Food Security Categories using NHANES 2017 - 2018




**Note:** The sample consists of households with responses to the 10 adult FSM questions from the 2017 - 2018 National Health and Nutrition Examination Survey (NHANES). The entire sample contains 4,738 households. Each household in our sample has a BGRM posterior mean value of food security that is estimated by the model. Using the responses of the same households to the adult food security questionnaire questions, households are categorized into 1 of 4 food security categories based on the standard USDA food security category definitions. Full Food Secure households have 0 affirmative responses to the adult FSM questionnaire, followed by 1-2, 3-5, and 6-10 affirmative responses for Marginal, Low and Very Low Food Secure households, respectively. Histograms depicting the household distribution of BGRM posterior mean food security values for households belonging to the same FSM food security category are displayed using the 4 colors above.

**Figure 3:** Distribution of BGRM Posterior Mean Latent Construct Using FSM and continuous HEI by USDA FSM Scale Food Security Category using NHANES 2017 - 2018




**Note:** The sample consists of households with responses to the 10 adult FSM questions from the 2017 - 2018 National Health and Nutrition Examination Survey (NHANES). The entire sample contains 4,738 households. For each household in our sample a BGRM posterior mean latent construct using FSM and continuous HEI is estimated by the model. Using the responses of the same households to the adult food security questionnaire, households are categorized into 1 of 4 food security categories based on the standard USDA food security category definitions. Full Food Secure households have 0 affirmative responses to the adult FSM questionnaire, followed by 1-2, 3-5, and 6-10 affirmative responses for Marginal, Low and Very Low Food Secure households, respectively. Histograms depicting the household distribution of BGRM posterior mean latent construct using FSM and HEI values for households belonging to the same FSM food security category are displayed using the 4 colors above.

# APPENDIX

# A Simulated Data Exercise

Prior to estimating the BGRM with real data, we test model performance in a simulated data exercise. The goal of this exercise is to ensure that our model can accurately retrieve known data generating parameters used to simulate a set of household responses. Data are simulated for a set of $I = 5,000$ households answering a set of $J = 10$ questions.[27] The potential responses for each question is set to mimic those of the household food security component of the FSM, implying that $C = [C_1, C_2, ..., C_{10}]$ is defined as $C = [3, 3, 3, 2, 3, 2, 2, 2, 2, 3]$. Our BGRM is well suited to this mix of binary and polytomous ordered response data as opposed to strictly binary or polytomous alternative IRT methods.

The set of data generating structural parameters is given in Table A.1 below. The food security level of each household $i$ is constructed such that $\delta_i \sim N(0, 1)$. The error term for each household $i$'s response to question $j$ is constructed such that $e_{ij} \sim N(0, \sigma_j^2)$, and each $y_{ij}^*$ and $y_{ij}$ are constructed using equations (7) and (1), respectively. The Gibbs Sampler algorithm outlined in Section 2.2 was run for 30,000 iterations with the first 10,000 draws removed for burn-in.

Table A.2 compares the data generating values of our parameters to their estimated posterior mean and 95% credible intervals.[28] Beginning with our estimates of $\mu$, we see that the posterior mean value is qualitatively similar to the data generating value in all cases. Furthermore, the true

[27] In this study, we only consider the set of 10 FSM questions answered by all households, ignoring the set of 8 additional child food security questions for households with children. In the future, the set of 8 child food security questions can be added to the model, though adult and child food security are thought to be separate latent constructs.

[28] The values in parentheses in Table A.2 show 95% credible intervals given by the post-burn-in posterior draw of each parameter corresponding to the specified percentile value. In this specific case, the 20,000 post-burn-in iterations are first ordered from smallest to largest. The left credible interval value is then given by the 500th draw and the right interval as the 19,500th draw. The "NE" designation implies that the specified parameter is fixed and therefore not estimated by the model.

data generating value of each $\mu_j$ falls within the posterior 95% credible interval in all cases. Looking to our estimates of $\sigma^2$, we again see that the posterior mean and data generating values are reasonably close to one another. Alternatively, the 95% credible interval of $\sigma^2$ does not cover the true value in the single case of question 7. Finally, moving to the estimates of $\gamma_2$, we find that while qualitatively similar to their data generating values, the 95% credible intervals of $\gamma$ only cover the true parameter value in two out of the five total cases. This again may be due to not having enough iterations to achieve convergence, but it may also be the result of utilizing limited response data.

As an alternative to checking whether the true value falls within the estimated 95% credible interval of each parameter, similar studies often rely on various forms of the Root Mean Squared Error (RMSE) calculated using the data generating and estimated parameter values (Zhu and Stone (2011), Kieftenbeld and Natesan (2012), Broomell and Bhatia (2014)). Taking this approach, we calculate the RMSE of our estimates for each parameter using the full set of post-burn-in draws from the posterior distribution. The RMSE of each parameter type is then averaged, giving us the average RMSE of $\mu$, $\sigma^2$, and $\gamma_2$. We find that the average RMSE of $\mu$, $\sigma^2$, and $\gamma_2$ are 0.029, 0.0432, and 0.0437, respectively. In line with other common statistical tests, there is no single RMSE value that signifies adequate parameter retrieval. However, we do find that our RMSE's fall well below the thresholds for convergence used in similar studies (Zhu and Stone (2011), Kieftenbeld and Natesan (2012)).

We now discuss our model's ability to accurately predict each household's latent food security variable. Figure A.1 shows the posterior mean value of $\delta$ for each household along the x-axis and true data generating values of $\delta$ on the y-axis, along with a 45 degree line representing perfect agreement between the two measures. The results shown in Figure A.1 provide two key insights. First, our estimated values of $\delta$ correspond well with their true values. Households with higher posterior mean values of $\delta$ are associated with higher data generating food security levels. Second, Figure A.1 shows a substantial amount of binning in our estimates of $\delta$. Specifically, the posterior mean values of $\delta$ fall into visually separable groups with discrete jumps across most groups. This binning is a byproduct of relying on discrete ordinal data. For a given set of households with identical responses in the observed data, the model can only partially distinguish between the relative food security level of households within the set. This limitation leads to households with the same responses having similar posterior mean values of $\delta$, forming bins.

Furthermore, in this specific exercise, values of $\mu$, $\sigma^2$, and $\gamma$ are constructed so that they are the same across all questions. With equal parameter values across questions, questions are also treated as equal in so far as their relative importance in determining $\delta$. For example, an individual who answers "3" to question 1 and "1" to all other questions will be placed into the same bin of posterior mean $\delta$'s as another individual who answers "3" to question 2 and "1" to all others. Alternatively, variation in parameter values across questions increases both the predictive power of $\delta$ and the number of final bins since questions are now distinct from one another in how they relate to $\delta$. Given that assigning identical parameters across questions can therefore be seen as a particularly challenging scenario, we still find that our model produces estimates that generally converge to their true value.

Finally, we compare the relative abilities of both the BGRM and official USDA FSM scale to accurately predict household food security status. First, we designate the 20th percentile of our data generating $\delta$'s as the cutoff separating food security and food insecurity, implying that households with values of $\delta$ below the 20th percentile are categorized as "food insecure" while households with $\delta$'s above the 20th percentile are "food secure". Next we categorize households as either food secure or food insecure using the posterior mean estimates of $\delta$ from our BGRM. To then categorize households using the FSM scale, we use the standard approach where households are considered food insecure if they have 3 or more responses indicative of food insecurity.[29] Finally, we redefine the food insecurity cutoff percentile to the 50th and 5th percentiles to evaluate the sensitivity of both measures to changes in underlying food insecurity rates. This involves changing the percentile used for both the data generating $\delta$'s and posterior mean $\delta$'s while keeping the FSM scale

[29] Since responses to each of the 10 questions are positively related to food security, responses indicative of food insecurity include a response of 1 or 2 to questions 1, 2, 3, 5, and 10, and a response of 1 to questions 4, 6, 7, 8, and 9.

cutoff values fixed.

We measure classification accuracy using the proper match rate and mismatch rate. To define these rates, note that each household $i$ has some true food security status $F_i$ such that $F_i = 1$ if the household is food secure and 0 otherwise based on the previously described threshold percentiles. Given each household's estimated food security status, $\hat{F}_i$, from either the BGRM or FSM scale cutoffs, we then define a proper match $P_i = 1$ as the case where $F_i = \hat{F}_i$ and a mismatch $P_i = 0$ as the case where $F_i \neq \hat{F}_i$. The proper match rate is then calculated as $(\sum_{i=1}^{I} P_i)/I$ and the mismatch rate is given by 1 minus the proper match rate. Table A.3 shows the proper match rates and mismatch rates for both the BGRM and FSM scale under our three percentile thresholds of food security. Beginning with the 20th percentile definition in Panel A of Table A.3, we see that the proper match rate for the BGRM and FSM scale are 0.94 and 0.52, respectively. This implies that while food security categories estimated using the BGRM were correct 94% of the time, categorizations from the FSM scale were only correct for roughly 52% of households. We find that the BGRM similarly outperforms the FSM scale when the food security cutoff is set to the 50th and 5th percentile in Panel B and Panel C, respectively.

While the BGRM's assignments of food security categories outperform those produced using the USDA's FSM scale with the 3 response rule, we can also adjust the FSM scale such that both measures more closely match the true underlying data. Specifically, we adjust the FSM scale's cutoff of "food insecurity indicative responses" needed to classify a household as food insecure until the share of food secure and food insecure households produced by the FSM scale most closely matches the true data. For example, with the food security threshold set to the 20th percentile, adjusting the FSM scale to require 7 indicative responses produces shares of food secure and food insecure households equal to roughly 77% and 23% of the sample, respectively. This same process is repeated for our alternative thresholds and the match/mismatch rates are then given in Table A.4. As expected, Table A.4 shows that adjusting the number of responses needed in the FSM scale to mirror the data increases its performance considerably. While the gap in performance between the BGRM and FSM scale categorizations is smaller, the BGRM still outperforms the FSM scale in all cases. Therefore, while the vast majority of studies rely on the traditional 3 response FSM scale to define food security/insecurity regardless of the sample, the BGRM is better able to assign households food security status in our simulated data exercise even in the case where the FSM scale is adjusted to most closely match a known threshold.

## A.1 Appendix A Tables & Figures

**Table A.1:** Simulated Data Exercise Data Generating Parameters

| $j$ | $\mu$ | $\sigma^2$ | $\gamma_2$ |
|---|---|---|---|
| **1** | 1 | 0.5 | 1.5 |
| **2** | 1 | 0.5 | 1.5 |
| **3** | 1 | 0.5 | 1.5 |
| **4** | 1 | 0.5 | $\infty$ |
| **5** | 1 | 0.5 | 1.5 |
| **6** | 1 | 0.5 | $\infty$ |
| **7** | 1 | 0.5 | $\infty$ |
| **8** | 1 | 0.5 | $\infty$ |
| **9** | 1 | 0.5 | $\infty$ |
| **10** | 1 | 0.5 | 1.5 |

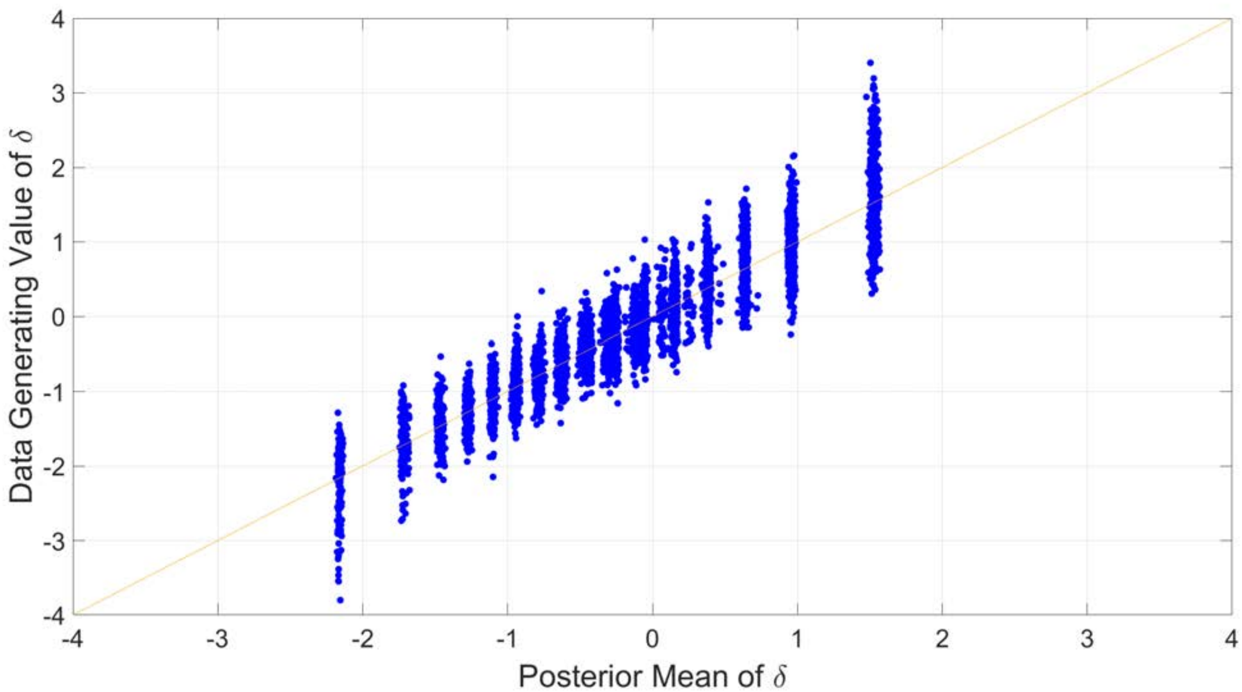
**Note:** Each $j$ represents a FSM response question that households without children must answer to be assigned a domestic household food security category. Each $\mu$ represents question $j$ specific intercept parameter. Each $\sigma^2$ represents question $j$'s specific variance associated with the elicited food security level of household $i$. Each $\gamma_2$ represents a response threshold for question $j$.

**Table A.2:** Simulated Data Exercise Data Generating Parameters vs. Posterior Means and 95% Credible Intervals

| $j$ | $\mu$ | $\hat{\mu}$ | $\sigma^2$ | $\hat{\sigma}^2$ | $\gamma_2$ | $\hat{\gamma}_2$ |
|---|---|---|---|---|---|---|
| **1** | 1 | 0.9972 | 0.5 | 0.4736 | 1.5 | 1.4572 |
| | | *(0.9470,1.0468)* | | *(0.4268,0.5224)* | | *(1.4050,1.5021)* |
| **2** | 1 | 0.9922 | 0.5 | 0.5048 | 1.5 | 1.4921 |
| | | *(0.9499,1.0360)* | | *(0.4613,0.5508)* | | *(1.4553,1.5304)* |
| **3** | 1 | 0.9754 | 0.5 | 0.4751 | 1.5 | 1.4396 |
| | | *(0.9335,1.0178)* | | *(0.4339,0.5175)* | | *(1.3986,1.4745)* |
| **4** | 1 | 1.0029 | 0.5 | 0.4585 | $\infty$ | NE |
| | | *(0.9505,1.0579)* | | *(0.3980,0.5252)* | | |
| **5** | 1 | 0.9970 | 0.5 | 0.5315 | 1.5 | 1.4632 |
| | | *(0.9504,1.0441)* | | *(0.4841,0.5825)* | | *(1.4226,1.4998)* |
| **6** | 1 | 0.9684 | 0.5 | 0.4502 | $\infty$ | NE |
| | | *(0.9169,1.0209)* | | *(0.3877,0.5181)* | | |
| **7** | 1 | 0.9820 | 0.5 | 0.4186 | $\infty$ | NE |
| | | *(0.9312,1.0354)* | | *(0.3608,0.4854)* | | |
| **8** | 1 | 1.0096 | 0.5 | 0.4918 | $\infty$ | NE |
| | | *(0.9557,1.0634)* | | *(0.4244,0.5651)* | | |
| **9** | 1 | 1.0090 | 0.5 | 0.5125 | $\infty$ | NE |
| | | *(0.9577,1.0619)* | | *(0.4451,0.5885)* | | |
| **10** | 1 | 1.0247 | 0.5 | 0.5080 | 1.5 | 1.5368 |
| | | *(0.9825,1.0670)* | | *(0.4654,0.5530)* | | *(1.5062,1.5669)* |

**Note:** Posterior mean value of $\mu$ is qualitatively close to the data generating value in all cases and the true data generating value of $\mu$ falls within the posterior 95% credible interval in all cases. Posterior mean and data generating values of $\sigma^2$ are reasonably close to one another. The 95% credible intervals of $\sigma^2$ cover the true value in all but one case. Posterior mean and data generating values of $\gamma_2$ are qualitatively similar. The 95% credible intervals of $\gamma^2$ cover the true value in 2 out of 5 cases.

**Figure A.1:** Data Generating Values of $\delta$ vs. Posterior Mean Draws of $\delta$




**Note:** Estimated values of $\delta$ correspond well with their true values. Households with higher posterior mean values of $\delta$ are associated with higher data generating food security levels. There is considerable binning in our estimates of $\delta$. This binning is a byproduct of the fact that we rely on ordinal polytomous data.

**Table A.3:** Food Security Categorization Accuracy of BGRM and FSM Scale

| **Panel A: Food Insecurity Defined as $\delta$ below the 20th Percentile** | | |
|---|---|---|
| | Proper Match Rate | Mismatch Rate |
| BGRM | 0.9416 | 0.0584 |
| FSM Scale | 0.5194 | 0.4806 |
| **Panel B: Food Insecurity Defined as $\delta$ below the 50th Percentile** | | |
| | Proper Match Rate | Mismatch Rate |
| BGRM | 0.9080 | 0.0920 |
| FSM Scale | 0.8078 | 0.1922 |
| **Panel C: Food Insecurity Defined as $\delta$ below the 5th Percentile** | | |
| | Proper Match Rate | Mismatch Rate |
| BGRM | 0.9720 | 0.0280 |
| FSM Scale | 0.3694 | 0.6306 |

**Note:** In panel A, we detect a large difference between the BGRM and FSM scale proper match rates. This result implies that food security categories estimated using the BGRM were correct 94% of the time, compared to only 52% of the time for the FSM scale. The scale of the difference between the BGRM scale and the FSM scale proper match rate stays relatively the same in both panel B and C where the security cutoffs are set to the 50th and 5th percentile respectively. Thus, generally, food security cutoffs estimated using the BGRM outperform the ones estimated using the FSM scale.

**Table A.4:** Food Security Categorization Accuracy of BGRM and Adjusted FSM Scale

| | | |
|---|---|---|
| **Panel A: Food Insecurity Defined as $\delta$ below the 20th Percentile** | | |
| | Proper Match Rate | Mismatch Rate |
| BGRM | 0.9416 | 0.0584 |
| FSM Scale | 0.9158 | 0.0842 |
| **Panel B: Food Insecurity Defined as $\delta$ below the 50th Percentile** | | |
| | Proper Match Rate | Mismatch Rate |
| BGRM | 0.9080 | 0.0920 |
| FSM Scale | 0.8810 | 0.1190 |
| **Panel C: Food Insecurity Defined as $\delta$ below the 5th Percentile** | | |
| | Proper Match Rate | Mismatch Rate |
| BGRM | 0.9720 | 0.0280 |
| FSM Scale | 0.9654 | 0.0346 |

**Note:** We adjust the FSM scale's number of "food insecurity indicative responses" needed to classify a household as food secure until the share of food secure and food insecure households produced by the FSM scale most closely matches the true data. The performance of the FSM scale improves considerably, however the BGRM still outperforms the FSM scale in all cases.

# B Health Outcomes Regressions

We examine the ability of our BGRM posterior mean FSEC measure to capture variation in health outcomes commonly associated with food security. Specifically, we use data from the 2017-18 NHANES to estimate regressions of four binary health outcomes: self-reported obesity, diagnosis of diabetes, diagnosis of high blood pressure, and following a good diet, on our BGRM posterior mean estimates. In each regression, we employ a vector of controls which capture gender, age, race, level of education, marital status, citizenship status, household size, and annual household income. We compare the resulting BGRM regression coefficients to regressions of our health outcomes of interest on binary indicators of marginal, low, and very low food security defined using the USDA's FSM scale, leaving full food security as the reference group; the number of affirmative FSM responses for each adult in the sample; and the respondent-level probabilities of falling within a given quartile of the BGRM posterior distribution.

Given this set of explanatory variables, the estimated regressions for each binary health outcome are represented through the linear model,

$$H_i = \beta_0 + \boldsymbol{\beta_1} \boldsymbol{Z}_i + \boldsymbol{\beta_2} \boldsymbol{X}_i + \epsilon_i \tag{1}$$

where $H_i$ represents a given binary health outcome for adult $i$, $Z_i$ represents the vector of explanatory variables, $X_i$ represents the vector of control variables, and $\epsilon_i$ is a stochastic error term. The parameters of interest are represented by $\boldsymbol{\beta_1}$.

Results from these regressions are shown in Table B.1. Table B.1 provides several key insights. The BGRM posterior means have a statistically significant association with all health outcomes. A one standard deviation increase in adult BGRM posterior mean FSEC value decreases the probability of obesity, being diagnosed with diabetes, and being diagnosed with high blood pressure, by 2.3, 2.1, and 2.6 percentage points, respectively, and increases the probability of an adult following a good diet by 4.9 percentage points. These results align with *a priori* expectations as higher

BGRM posterior mean food security values correspond to higher levels of food security. With each additional affirmative response to the set of FSM questions, adults are 0.6 percentage points more likely to be obese, 0.8 percentage points more likely to be diagnosed with diabetes, 1.2 percentage points more likely to have high blood pressure, and 1.9 percentage points less likely to follow a good diet. The BGRM posterior mean values explain more of the variation in obesity, diabetes, and diet quality than the raw affirmative FSM count according to adjusted R-squared. BGRM posterior means also explain more of the variation in obesity than the set of USDA-defined FSM food security category indicator variables.

While each respondent's posterior mean BGRM FSEC value has a statistically significant association with adult health outcomes, we find little evidence of a statistically significant relationship between the BGRM quartile probabilities and our set of adult health outcomes. A one standard deviation increase in the probability of an individual falling within the first quartile of the BGRM posterior distribution decreases the probability of that individual having a good diet by 5.2 percentage points. Conversely, adults that are very low food secure according to the FSM scale are 4.2 percentage points more likely to have diabetes, 6 percentage points more likely to have high blood pressure, and 15 percentage points less likely to follow a good diet compared to full food secure adults. Low food secure adults as defined by the USDA are 5.9 percentage points more likely to be obese, 8.4 percentage points more likely to have diabetes, 6.8 percentage points more likely to have high blood pressure, and 8.9 percentage points less likely to have a good diet compared to full food secure adults. Marginally food secure adults are 5.3 percentage points more likely to be obese, and 5.6 percentage points less likely to have a good diet compared to full food secure adults. The relationship between USDA FSM scale food security categories and the set of health outcomes is non-linear. This result is clearly illustrated for diabetes and high blood pressure where we find that very low food secure adults are less likely to be diagnosed with diabetes or high blood pressure compared to low food secure adults.

Taken together, the results of our regression analyses suggest that the associations between our BGRM posterior mean estimates and health outcomes of interest are highly statistically and economically significant in all cases. The adjusted R-squared of our regressions using BGRM posterior mean are similar to those of regressions using FSM scale indicator variables and higher in the case of obesity. Given the similar health outcome regression results produced using both BGRM and FSM scale indicators, the flexibility, ability to measure uncertainty in estimation, and other attractive features of our BGRM provide additional capabilities to those of more common methods with relatively little loss in predictive power and gains for certain outcomes.

## B.1 Appendix B Tables & Figures

**Table B.1:** Coefficient estimates of the BGRM Posterior Mean of Food Security vs. FSM Scale Food Security Categories on Adult Households Health Outcomes

| | **Obesity** | | | | Diabetes | | | | **High Blood Pressure** | | | | Good Diet | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **(1)** | **(2)** | **(3)** | **(4)** | (5) | (6) | (7) | (8) | **(9)** | **(10)** | **(11)** | **(12)** | (13) | (14) | (15) | (16) |
| BGRM Posterior Mean | -0.027*** | | | | -0.025*** | | | | -0.031*** | | | | 0.058*** | | | |
| | (0.010) | | | | (0.008) | | | | (0.009) | | | | (0.010) | | | |
| Affirmative FSM Count | | 0.006* | | | | 0.008*** | | | | 0.012*** | | | | -0.019*** | | |
| | | (0.003) | | | | (0.003) | | | | (0.003) | | | | (0.003) | | |
| 1st Quartile | | | 0.078 | | | | 0.058 | | | | 0.043 | | | | -0.134*** | |
| | | | (0.054) | | | | (0.041) | | | | (0.050) | | | | (0.052) | |
| 2nd Quartile | | | 0.070 | | | | 0.020 | | | | -0.060 | | | | -0.052 | |
| | | | (0.054) | | | | (0.041) | | | | (0.050) | | | | (0.052) | |
| 3rd Quartile | | | 0.022 | | | | -0.013 | | | | -0.056 | | | | -0.015 | |
| | | | (0.104) | | | | (0.079) | | | | (0.097) | | | | (0.100) | |
| Very Low Food Secure | | | | 0.034 | | | | 0.042** | | | | 0.060** | | | | -0.154*** |
| | | | | (0.028) | | | | (0.021) | | | | (0.026) | | | | (0.027) |
| Low Food Secure | | | | 0.059** | | | | 0.084*** | | | | 0.068*** | | | | -0.089*** |
| | | | | (0.025) | | | | (0.018) | | | | (0.023) | | | | (0.023) |
| Marginal Food Secure | | | | 0.053** | | | | 0.019 | | | | -0.006 | | | | -0.056** |
| | | | | (0.023) | | | | (0.017) | | | | (0.022) | | | | (0.022) |
| Adj. R-squared | 0.0579 | 0.0569 | 0.0577 | 0.0578 | 0.1129 | 0.1124 | 0.1130 | 0.1148 | 0.1931 | 0.1937 | 0.1939 | 0.1933 | 0.0921 | 0.0919 | 0.0911 | 0.0924 |
| Observations (N) | 4,059 | 4,059 | 4,059 | 4,059 | 3,967 | 3,967 | 3,967 | 3,967 | 4,082 | 4,082 | 4,082 | 4,082 | 4,090 | 4,090 | 4,090 | 4,090 |
| Mean DV | 0.37 | 0.37 | 0.37 | 0.37 | 0.16 | 0.16 | 0.16 | 0.16 | 0.38 | 0.38 | 0.38 | 0.38 | 0.67 | 0.67 | 0.67 | 0.67 |

**Notes:** There are 4 binary outcomes employed: self-reported obesity, diagnosis of diabetes, diagnosis of high blood pressure, and good diet quality. For each of these outcomes four regressions are displayed respectively: (1) A regression where the outcome is regressed on the BGRM Posterior Means of Food Security for the sample; (2) A regression where the outcome is regressed on the FSM Scale Categories of Food Security; (3) A regression where the outcome is regressed on the number of FSM questions adults answered affirmatively; (4) A regression where the outcome is regressed on the probabilities of individuals' draws of latent food security falling within a given quartile. The same controls are used accross all regressions: gender, age, education, marital status, citizenship, race, household size and annual household income. Mean DV stands for mean of dependent variable. The sample consists of the 2017 - 2018 adult food security data of the National Health and Nutrition Examination Survey (NHANES). (* 10%, ** 5%, and *** 1%).

# C Estimation Including Healthy Eating Index

In the primary BGRM specification, our estimation procedure solely relied on respondents' answers to the adult FSM food security questions. For each question in the survey, we estimated up to three parameters: $\mu_j$, $\sigma_j$, $\gamma_j$.[30] For these questions, we could not simultaneously estimate $\sigma_j$ and $\lambda_j$, as the $y^*_{ij}$ were not observed and therefore $\sigma_j$ and $\lambda_j$ were not identified simultaneously.

We now add the Healthy Eating Index (HEI) associated with each participant in our sample. The estimation procedure for the parameters corresponding to the original 10 FSM questions remains unchanged. However, for the HEI, we can now estimate $\mu_j$, $\sigma_j$, and $\lambda_j$ simultaneously. Estimation of $\sigma_j$ and $\lambda_j$ is possible because the value of the HEI for the $i$'th individual is treated as the true value of $y^*_{ij}$ for that individual. Thus for the HEI variable, the conditional posterior distributions of $\mu_j$, $\sigma_j$, and $\lambda_j$, differ slightly from those for FSM-related parameters. We display these conditional posterior distributions below. These conditional posterior distributions are then sampled from directly in the Gibbs Sampler algorithm. While we use HEI as our continous variable of interest in this study, the same approach allows for any additional continuous variable(s), highlighting the BGRM's flexibility.

## C.1 Conditional Posteriors

**1. Conditional Posterior of $\mu_j$.**

The estimation equation for the HEI variable $j$ is given as:

$$y^*_j - \lambda_j \delta = 1_I \mu_j + e_j$$

[30] Gammas were only estimated for FSM questions with more than two potential responses.

The $\mu_j$ is then drawn from the following full conditional posterior distribution:

$$\mu_j|\delta, \sigma^2, y^*, Y \sim N(v, V) \tag{2}$$

where $V = \left(\frac{1}{\sigma_\mu^2} + \frac{1_I' 1_I}{\sigma_j^2}\right)^{-1}$ and $v = V\left[\frac{\bar{\mu}}{\sigma_\mu^2} + \frac{1_I'(y_j^* - \lambda_j \delta)}{\sigma_j^2}\right]$.

**2. Conditional Posterior of $\sigma_j^2$**

The estimation equation is given as:

$$y_j^* = 1_I \mu_j + \lambda_j \delta + e_j$$

The $\sigma_j^2$ is then drawn from the following full conditional posterior distribution:

$$\sigma_j^2|\mu, \delta, \gamma, Y \sim IG(a, b) \tag{3}$$

where $a = \alpha + \frac{I}{2}$, and $b = \left[\frac{1}{\beta} + \frac{(y_j^* - 1_I\mu_j - \lambda_j\delta)'(y_j^* - 1_I\mu_j - \lambda_j\delta)}{2}\right]^{-1}$.

**3. Conditional Posterior of $\lambda_j$.**

The estimation equation is given as:

$$y_j^* - 1_I \mu_j = \lambda_j \delta + e_j$$

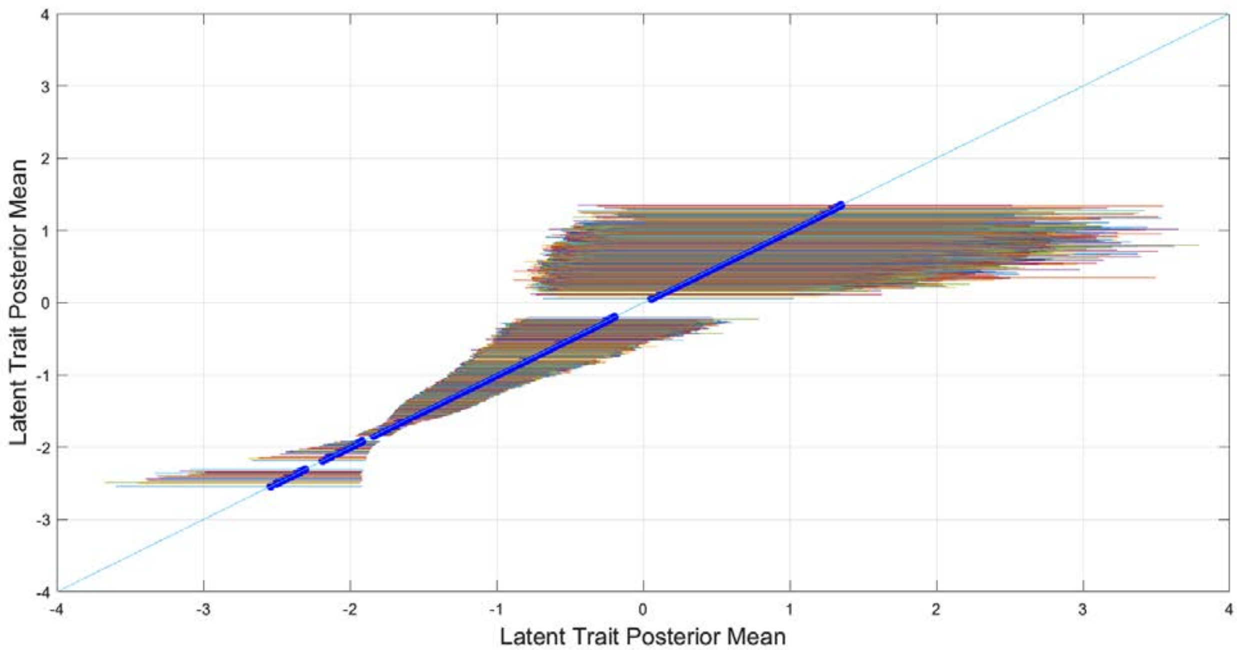
The $\lambda_j$ is then drawn from the following full conditional posterior distribution:

$$\lambda_j|\delta, \sigma^2, y^*, Y \sim \mathcal{N}(G, g) \tag{4}$$

where $g = \left(\frac{\delta'\delta}{\sigma_j^2} + \frac{1}{\sigma_\lambda^2}\right)^{-1}$, and $G = g\left[\frac{\delta'(y_j^* - 1_I\mu_j)}{\sigma_j^2} + \frac{\lambda}{\sigma_\lambda^2}\right]$.

## C.2 Appendix C Tables & Figures

**Figure C.1:** Posterior Mean Values of BGRM Latent Construct using FSM and continuous HEI with 95% Credible Intervals using NHANES 2017 - 2018




**Note:** The sample consists of households with responses to the 10 adult FSM questions from the 2017 - 2018 National Health and Nutrition Examination Survey (NHANES). The entire sample contains 4,738 households. Each blue dot on the $45^o$ degree line represents the BGRM posterior mean latent construct using FSM and continuous HEI value for a given household in the sample. The colored lines running through the blue dots represent the magnitude of the 95% credible intervals for each household.

**Table C.1:** Coefficient estimates of the BGRM Posterior Mean of Food Security vs HEI BGRM Latent Construct on Adult Households Health Outcomes

| | **Obesity** | | | | Diabetes | | | | **High Blood Pressure** | | | | Good Diet | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **(1)** | **(2)** | **(3)** | **(4)** | (5) | (6) | (7) | (8) | **(9)** | **(10)** | **(11)** | **(12)** | (13) | (14) | (15) | (16) |
| BGRM Posterior Mean | -0.027$^{***}$ | | | | -0.025$^{***}$ | | | | -0.031$^{***}$ | | | | 0.058$^{***}$ | | | |
| | (0.010) | | | | (0.008) | | | | (0.009) | | | | (0.010) | | | |
| BGRM Posterior Mean | | -0.033$^{***}$ | | | | -0.027$^{***}$ | | | | -0.034$^{***}$ | | | | 0.065$^{***}$ | | |
| | | (0.010) | | | | (0.008) | | | | (0.010) | | | | (0.010) | | |
| 1st Quartile | | | 0.078 | | | | 0.058 | | | | 0.043 | | | | -0.134$^{***}$ | |
| | | | (0.054) | | | | (0.041) | | | | (0.050) | | | | (0.052) | |
| 2nd Quartile | | | 0.070 | | | | 0.020 | | | | -0.060 | | | | -0.052 | |
| | | | (0.054) | | | | (0.041) | | | | (0.050) | | | | (0.052) | |
| 3rd Quartile | | | 0.022 | | | | -0.013 | | | | -0.056 | | | | -0.015 | |
| | | | (0.104) | | | | (0.079) | | | | (0.097) | | | | (0.100) | |
| 1st Quartile | | | | 0.171$^{***}$ | | | | 0.037 | | | | 0.167$^{***}$ | | | | -0.261$^{***}$ |
| | | | | (0.056) | | | | (0.042) | | | | (0.052) | | | | (0.053) |
| 2nd Quartile | | | | 0.149$^{***}$ | | | | 0.007 | | | | 0.033 | | | | -0.177$^{***}$ |
| | | | | (0.055) | | | | (0.041) | | | | (0.051) | | | | (0.053) |
| 3rd Quartile | | | | 0.211$^{**}$ | | | | -0.062 | | | | 0.211$^{**}$ | | | | -0.260$^{**}$ |
| | | | | (0.107) | | | | (0.081) | | | | (0.100) | | | | (0.102) |
| Healthy E. Index | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| Adj. R-squared | 0.0579 | 0.0586 | 0.0577 | 0.0589 | 0.1129 | 0.1131 | 0.1130 | 0.1133 | 0.1931 | 0.1936 | 0.1939 | 0.1947 | 0.0921 | 0.0939 | 0.0911 | 0.0935 |
| Observations (N) | 4,059 | 4,059 | 4,059 | 4,059 | 3,967 | 3,967 | 3,967 | 3,967 | 4,082 | 4,082 | 4,082 | 4,082 | 4,090 | 4,090 | 4,090 | 4,090 |
| Mean DV | 0.37 | 0.37 | 0.37 | 0.37 | 0.16 | 0.16 | 0.16 | 0.16 | 0.38 | 0.38 | 0.38 | 0.38 | 0.67 | 0.67 | 0.67 | 0.67 |

**Notes:** There are 4 binary outcomes employed: self-reported obesity, diagnosis of diabetes, diagnosis of high blood pressure, and good diet quality. For each of these outcomes four regressions are displayed respectively: (1) A regression where the outcome is regressed on the BGRM Posterior Means of Food Security for the sample; (2) A regression where the outcome is regressed on the BGRM Healthy Eating Index latent construct; (3) A regression where the outcome is regressed on the probabilities of individuals' draws of latent food security falling within a given quartile; (4) A regression where the outcome is regressed on the probabilities of individuals' draws of latent healthy eating index construct falling within a given quartile. The same controls are used accross all regressions: gender, age, education, marital status, citizenship, race, household size and annual household income. Mean DV stands for mean of dependent variable. The sample consists of the 2017 - 2018 adult food security data of the National Health and Nutrition Examination Survey (NHANES). (* 10%, ** 5%, and *** 1%).

# D Estimation Using Current Population Survey (2018)

The FSM was first piloted in the 1995 Current Population Survey (CPS) Food Security Supplement (FSS). Ever since, the USDA has used annual CPS data to calculate official national food security rates in the United States annually. Consequently, to examine the robustness of our BGRM and identify potential implications for national food security rates, we estimate latent food security using FSM responses in the 2018 CPS FSS.[31] The 2018 CPS FSS includes 37,300 interviewed households with 89,665 person records. The CPS FSS contains external and internal screening protocols to reduce respondent survey burden and avoid questions that may be inappropriate for some respondents given information provided earlier in the survey. Households that pass the external screen are exempt from taking the entire FSM and they are automatically assigned to the USDA's full food security category. Furthermore, the adult FSM questionnaire in the CPS includes two internal screeners. Households that register no food stress in a given set of adult internal screener questions are not asked the remaining adult FSM questions, and their responses to skipped questions are assumed to be negative.[32]

Unlike the CPS, the NHANES does not use an external screener for the FSM, but it does use internal screeners. To make our CPS results more comparable to the NHANES, we exclude all households from our CPS sample that were screened out of completing the full survey by the external screener. This restriction reduces our CPS sample size from 37,300 households to 14,326 households. Using this reduced sample, we estimate our model and produce Figures D.1 and D.2 shown below, counterparts to Figures 1 and 2 in the main text produced using the NHANES sample.

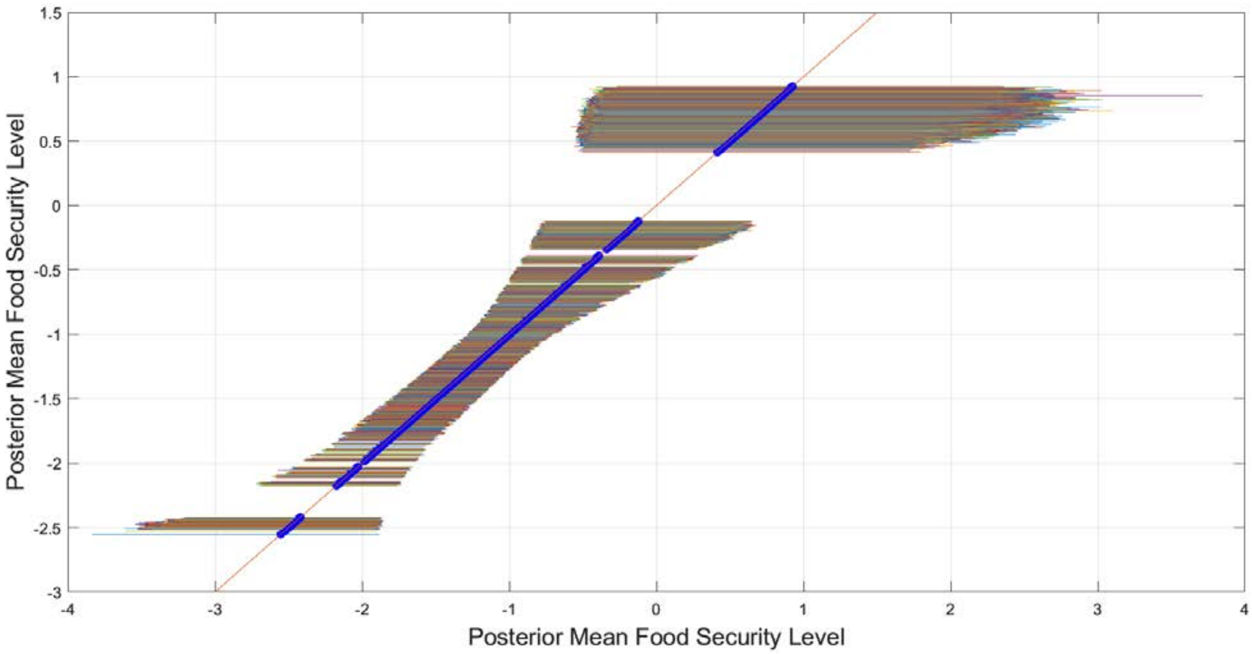
---

[31] We use the 2018 FSS for consistency with our 2017-18 NHANES results.

[32] For more information regarding screening procedures for the 2018 CPS FSS, see 2018 CPS FSS Technical Documentation.

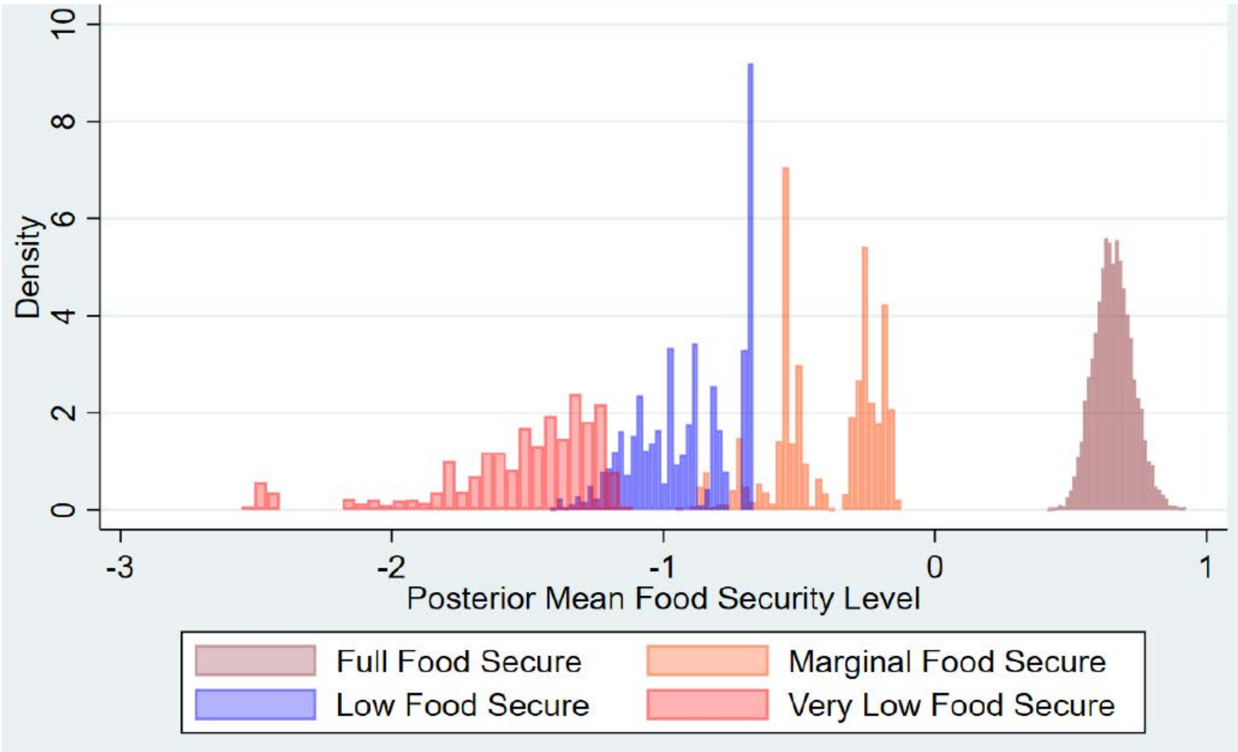
## D.1 Appendix D Tables & Figures

**Figure D.1:** Posterior Mean Values of BGRM Latent Food Security with 95% Credible Intervals using CPS 2018 Food Security Supplement




**Note:** The sample consists of households with responses to the 10 adult FSM questions from the 2018 Current Population Survey (CPS) Food Security Supplement. The entire sample contains 14,326 households. Each blue dot on the 45$^o$ degree line represents the BGRM posterior mean latent food security value for a given household in the sample. The colored lines running through the blue dots represent the magnitude of the 95% credible intervals for each household.

**Figure D.2:** Distribution of BGRM Posterior Mean Food Security by USDA FSM Scale Food Security Categories using CPS 2018 Food Security Supplement




**Note:** The sample consists of households with responses to the 10 adult FSM questions from the 2018 Current Population Survey (CPS) Food Security Supplement. The entire sample contains 14,326 households. Each household in our sample has a BGRM posterior mean value of food security that is estimated by the model. Using the responses of the same households to the adult food security questionnaire questions, households are categorized into 1 of 4 food security categories based on the standard USDA food security category definitions. Full Food Secure households have 0 affirmative responses to the adult FSM questionnaire, followed by 1-2, 3-5, and 6-10 affirmative responses for Marginal, Low and Very Low Food Secure households, respectively. Histograms depicting the household distribution of BGRM posterior mean food security values for households belonging to the same FSM food security category are displayed using the 4 colors above.