Measurement 120 (2018) 193–205

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Causal inference with latent variables from the Rasch model as outcomes

Matthew P. Rabbitt

Economic Research Service, U.S. Department of Agriculture, 355 E Street SW, Washington, DC 20024-3221, USA

ARTICLE INFO

Keywords:
Causal inference
Food insecurity
Item response theory
Latent regression
Rasch model
Selection
SNAP

ABSTRACT

This article discusses and compares several methods for estimating the parameters of a latent regression model when one of the explanatory variables is an endogenous binary (treatment) variable. Traditional methods based on two-stage least squares and the Tobit selection model where the dependent variable is an estimate of the latent variable from the Rasch model are compared to the behavioral Rasch selection model. The properties of these methods are examined using simulated data and empirical examples are included to demonstrate the usefulness of the behavioral Rasch selection model for research in the social sciences. The simulations suggest the latent regression model parameters are more accurately and precisely estimated by the behavioral Rasch selection model than by two-stage least squares or the Tobit selection model. The empirical examples demonstrate the importance of addressing endogenous explanatory variables in latent regressions for Item Response Theory (IRT) models when estimating causal differences in the latent variable or examining differential item functioning.

1. Introduction

Inference in modern empirical research is often based on parameters from regression models of outcomes represented by a latent variable, such as a person's ability, health, food insecurity, or well-being, on a set of explanatory variables. Often these relationships are modeled indirectly using linear or censored regression models where the dependent variable is an estimate of the latent variable. Estimates of the latent variable can be obtained by combining item responses to an instrument using models from Item Response Theory (IRT; [36]), such as the Rasch model [26]. While this approach is straightforward, modeling the desired relationships between the latent variable and a set of explanatory variables directly using an IRT model offers greater precision and accuracy [9,38] because the measurement and latent regression (behavioral) models are jointly estimated, rather than estimated in two steps.

Early research on modeling latent regressions in IRT models focused on the validity of the distributional assumptions of the latent variable [2] and comparing the distribution of ability across groups [29]. Multivariate latent regressions models were later developed for the dichotomous Rasch model [18,19,20,38], polytomous Rasch model [39], and loglinear Rasch model [10]. For a comprehensive discussion of these models, see De Boeck and Wilson [13]. Even though these models rely on observational data to estimate the parameters of the latent regression, they do not address endogeneity in the latent regression model's explanatory variables. Endogeneity can occur in observational data because of omitted variables and measurement error. Failure to account for these sources of bias will render estimates of the latent regression model parameters biased and inconsistent. The behavioral Rasch selection model (BRSM; [22]) addresses endogenous binary variables in the latent regression model for the dichotomous Rasch model using an instrumental variables approach.

This article describes and compares several methods for estimating the parameters of a latent regression model with a binary endogenous variable. The methods considered include the BRSM, two-stage least squares (2SLS) and Tobit [34] selection model (TSM). The properties of these methods are compared using simulated item response data. Empirical examples are also considered to demonstrate the usefulness of the BRSM for estimating the causal effect of a binary endogenous variable on a latent variable, and correcting for the endogeneity of the group indicator variable in analyses of differential item functioning (DIF). The simulated data and empirical analyses make two notable contributions. First, this is the first time simulated data has been used to compare the BRSM, 2SLS, and TSM under the assumption that the data is generated with an endogenous binary variable. Second, the empirical analysis is the first time the endogeneity of the DIF group indicator has been addressed.

The remainder of this article is organized as follows. In the next

Abbreviations: BRM, behavioral Rasch model; BRSM, behavioral Rasch selection model; CPS-FSS, Current Population Survey Food Security Supplement; DIF, differential item functioning; EAP, expected a posteriori; HFSSM, household food security survey module; IRT, Item Response Theory; MML, marginal maximum likelihood; POM, potential outcomes model; RMSE, root mean square error; RCM, Rubin causal model; SNAP, Supplemental Nutrition Assistance Program; TSM, Tobit selection model

E-mail address: matthew.rabbitt@ers.usda.gov.

https://doi.org/10.1016/j.measurement.2018.01.044
Received 5 May 2017; Received in revised form 22 January 2018; Accepted 23 January 2018
Available online 02 February 2018
0263-2241/ Published by Elsevier Ltd.

section, I discuss how the BRSM is developed from the Rasch model and other methods for estimating the latent regression model parameters. The following section describes the simulation analyses and discusses the results. Next, I describe the data used in the empirical examples and how the BRSM is useful in two particular settings: estimating the causal effect of a binary endogenous variable on a latent variable, and correcting for endogeneity in analyses of DIF. Results from the empirical examples are discussed with emphasis on their implications for future applications of the BRSM. Lastly, I discuss the implications of my findings from the simulations and empirical examples.

2. An illustrative Rasch model

For illustrative purposes, I consider the Rasch model for dichotomous responses in this article; however, the methods presented can be modified to include more complex models, such as the two-parameter IRT logistic model [5] or models that allow for polytomous responses [1,3]. The model is framed to measure a person's ability; however, it could just as easily be modified to measure a person's health, food insecurity, or any other latent trait. The model assumes person i's underlying latent index of ability, denoted by θ_i, represents the person's location on the continuum of ability. Higher values of θ_i are associated with greater ability. Assuming there are J binary items administered to the person that capture different levels of ability, then the probability that the person affirms the j^{th} item is

$$P(Y_{ij}=1|\theta_i)=\frac{\exp\left(\theta_i-\sum_{k=1}^{J}\delta_k X_{kj}\right)}{1+\exp\left(\theta_i-\sum_{k=1}^{J}\delta_k X_{kj}\right)}, \tag{1}$$

where exp(•) is the exponential function, δ_k is an item-difficulty parameter, and X_{kj} equals one if k = j, and zero otherwise. Item-difficulty parameters represent each items location on the continuum of ability, and are assumed to take on different values for each item. Higher item-difficulty parameter values are consistent with items that capture greater ability. The model further assumes that the person's responses are independent, conditional on latent ability, and the item-discrimination parameters are constrained to be equal across all items and normalized to one.

The item-difficulty parameters can be estimated using a maximum likelihood method called marginal maximum likelihood (MML; [6]) because the parameters of the distribution of latent ability are integrated out. For a detailed discussion of the methods for estimating the Rasch and other IRT model parameters, see [36]. The person's ability parameter can be estimated, conditional on the values of the item-difficulty parameters, using maximum likelihood or Bayesian methods. Bayesian methods for obtaining estimates of the person's ability parameters are particularly useful when these estimates are to be used in linear or censored regression models, where the dependent variable is the estimate of person ability, since estimates of ability can be obtained for persons with all zero and perfect scores on an instrument. Persons with extreme responses to an instrument are particularly important in these regressions since their responses may reveal important differences in behavior that are associated with ability. Maximum likelihood methods have been developed that "assign" values of ability to persons with extreme responses; however, they depend on which statistical software is used [12].

The most commonly used methods for estimating IRT model parameters are MML for estimating the item-difficulty parameters and the Bayesian expected a posteriori (EAP) method for estimating the person-ability parameters [36]. A benefit of the Bayesian EAP method for obtaining estimates of the person-ability parameters is that linear and censored regression models using them as dependent variables do not have to be adjusted for estimation error from the measurement model prior to testing for differences in ability based on the explanatory variables.

All analyses contained in this article were performed using Stata 15 Multiprocessor. The number of quadrature points was set at 15 for the numerical methods required to estimate the item-difficulty and person-ability parameters. When the numerical methods required evaluating integrals of more than one dimension, 15 quadrature points were used for each dimension. Using 15 quadrature points has been shown to produce reasonably accurate parameter estimates in this type of analysis [7,31].

2.1. Incorporating person-level covariates into the Rasch model

The Rasch model, described above, has been used extensively for scale development in the measurement sciences; however, it can also be used to directly examine the relationship between latent ability and a series of explanatory variables. A multivariate behavioral component, consisting of person-level explanatory variables, can be incorporated into the Rasch model by respecifying the person's latent ability index (θ_i) as

$$\theta_i=\beta_T T_i+\beta_X' X_i+e_i,\ \text{with } e_i\sim \text{i. i. d. } N(0,\sigma^2), \tag{2}$$

where T_i is an observed treatment indicator and X_i is a vector of observed person-level explanatory variables. The term treatment is used very broadly in economics and other fields. Essentially, it covers any variable whose effect on some outcome is the object of study. Substituting Eq. (2) into Eq. (1) yields

$$P(Y_{ij}=1|T_i,X_i,e_i)=\frac{\exp\left(\beta_T T_i+\beta_X' X_i+e_i-\sum_{k=1}^{J}\delta_k X_{kj}\right)}{1+\exp\left(\beta_T T_i+\beta_X' X_i+e_i-\sum_{k=1}^{J}\delta_k X_{kj}\right)}. \tag{3}$$

The model described in Eq. (3) is referred to as the behavioral Rasch model (BRM) in this article. The term "behavioral" is used to emphasize that the BRM includes a latent regression model in addition to the measurement model. Alternatively, the BRM has also been described as a generalized Rasch model [38] and a person-explanatory Rasch model [13]. The BRM parameters can be estimated using the MML method.

The latent regression specified in Eq. (2) can also be estimated using the predicted person-ability parameters as the dependent variable in a regression model. Any measurement error resulting from the estimation of the person-ability parameters must be addressed for this approach to be feasible if the ability parameters are estimated using a maximum likelihood method. A common method for addressing this measurement error is to assume it is orthogonal to the person-level explanatory variables (i.e., classical measurement error). Since person-ability parameters can theoretically take on any value on the real line it is common for these models to be estimated using a linear regression model [9,38]. Yet, data limitations or poor instrument design may constrain the estimates of the person-ability parameters to a smaller interval on the real line.

If an instrument is poorly designed or survey data contains a selective sample of persons with high or low levels of ability, then there may be restricted coverage of items or persons on the continuum of latent ability, respectively. If this occurs, person-ability parameter estimates can be grouped at the bounds of the feasible range for estimation. Censoring will occur resulting in biased and inconsistent estimates of the regression model parameters. This can be addressed by estimating a censored regression model, such as the Tobit model. Under the Tobit model with censoring at the lower level, values of latent ability at or below the censoring threshold are unobservable and assigned a value of c, while latent ability above the threshold is observed and set equal to the person's ability parameter estimate.

2.2. Correcting for the endogeneity of a binary (treatment) variable

The BRM can be used for inference based on the model parameters if the explanatory variables are exogenous. Endogeneity of one or more of the explanatory variables can arise when using observational data because of omitted variables or measurement error. Consequently, inference based on the BRM parameters for any endogenous variable would be problematic because they are biased and inconsistent. This occurs because of observable and unobservable differences between persons who are self-selecting into the treatment. The BRM controls for selection on the observable explanatory variables, but it cannot account for selection on unobservable variables. To address this key limitation of the BRM, Rabbitt [22] proposed the BRSM, which can address the endogeneity of a binary variable in the BRM using cross sectional data.

The BRSM assumes person's i's decision to participate in the treatment is determined by

$$T_i = \begin{cases} 1 & \text{if } \alpha_X' X_i + \alpha_Z' Z_i + u_i > 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where T_i and X_i are defined above, Z_i is a set of instrumental variables, and u_i is an unobservable variable that is assumed to be standard normally distributed ($u_i \sim N(0,\ 1)$). The normality assumption is merely an assumption of convenience. Any symmetric distribution would be suitable, but non-normal distributions would require different numerical methods to evaluate the integrals. This is usually accomplished using Monte Carlo integration. The resulting model is consistent with a standard probit model for the decision to participate in the treatment.

Endogeneity of the treatment variable can be formalized in the BRSM, as in Terza [33], by assuming the error-component in Eq. (2) can be decomposed into u_i and e_i^*, such that $e_i = \lambda u_i + e_i^*$. Then the person's latent ability index (θ_i) can be respecified as

$$\theta_i^* = \beta_T T_i + \beta_X' X_i + \lambda u_i + e_i^*, \text{ with } e_i^* \sim \text{i. i. d. } N(0,\eta^2), \tag{5}$$

where λ is a factor loading (selection) parameter and e_i^* represents the residual unobserved heterogeneity in latent ability after controlling for observable and unobservable variables. Correlation between the treatment variable (T_i) and latent ability is generated through the factor loading parameter, λ. If λ is estimated to be nonzero, u_i influences a person's decision to participate in the treatment and the likelihood of affirming the j^{th} item, rendering the BRM biased and inconsistent. The magnitude and sign of the factor loading parameter indicates the size and direction of the selection, allowing researchers to test additional hypotheses about selection behavior.

Substituting Eq. (5) into Eq. (1) yields

$$P(Y_{ij} = 1|T_i,X_i,u_i,e_i^*) = \frac{\exp\left(\beta_T T_i + \beta_X' X_i + \lambda u_i + e_i^* - \sum_{k=1}^{J} \delta_k X_{kj}\right)}{1 + \exp\left(\beta_T T_i + \beta_X' X_i + \lambda u_i + e_i^* - \sum_{k=1}^{J} \delta_k X_{kj}\right)}. \tag{6}$$

Numerical methods have been devised for the BRSM described in Eq. (6). A detailed description of these methods is available in Rabbitt [22]. In this article, I use a specialized Stata program developed for the BRSM. The program is available from the author upon request.

Alternative methods have been developed for estimating the BRM parameters when one of the variables is endogenous using longitudinal data. These methods generally assume the source of endogeneity, u_i, is time invariant. Wilde and Nord [37] estimate a fixed-effects BRM with only time varying variables, while Moffitt and Ribar [21] relax some of these assumptions by estimating a correlated-random effects BRM. A similar approach was employed by Depolt, Moffitt, and Ribar [14] for the multiple indicator multiple cause model (MIMIC).

The latent regression model, with an endogenous treatment variable (Eqs. (4) and (5)), can be estimated using methods that are similar to those used when the treatment variable is exogenous. The parallel to the linear regression model in this case is 2SLS. Under 2SLS, Eq. (4) is estimated in the first stage using a linear probability model of the endogenous treatment variable on the explanatory and instrumental variables. The predicted probability of the treatment is then calculated based on the estimated parameters from the first stage and used in place of the endogenous treatment variable in the second stage regression of the person's ability estimate on the predicted probability of the treatment and the explanatory variables to estimate Eq. (5).

The TSM can also be used to estimate the latent regression model with an endogenous treatment variable. Unlike 2SLS, the TSM accounts for potential censoring that may result from the estimation of person ability (described above). Under the TSM, Eq. (4) is modeled using a probit model like the BRSM, but Eq. (5) is estimated using the person's ability estimate as the dependent variable. Estimation of the model is completed by jointly estimating the TSM versions of Eqs. (4) and (5) using the methods described in Terza [33].

The primary advantage of the BRSM is that it estimates the measurement (Eq. (1)) and latent regression (Eq. (5)) models jointly, while 2SLS and TSM regressions of the person ability estimates are estimated in two steps. Estimating the measurement model in the BRSM also creates new research opportunities in the measurement sciences, such as new insights into DIF analysis when the group indicator is endogenous. Such analyses could not be conducted using 2SLS or the TSM. Comparisons of these methods are performed using simulated data and then actual data is used to provides examples of the usefulness of the BRSM.

3. Simulation analysis

Simulations were performed to compare the different methods for estimating the parameters of the latent regression model with an endogenous treatment variable. Similar simulation studies have been performed by Christensen [9] and Zwinderman [38], but under the assumption that the explanatory variables are exogenous. Rabbitt [23] simulates item response data with an endogenous treatment variable under the BRSM assumptions, but does not consider other modeling strategies than the BRSM. I expect the BRSM will perform better than the 2SLS and TSM methods because the measurement and latent regression models (Eqs. (1) and (5) are jointly estimated, rather than estimated in two steps. Responses to an instrument with eight items were simulated and the item-difficulty parameters were set to [0, 2.5, 3.5, 4.5, 5.5, 6.5, 7.5, 8.5]. Samples of 500, 1000, 5000, and 10,000 persons were drawn to illustrate the relative differences in the methods with varying sample sizes. Each simulation was repeated 1000 times.

In all simulations, an explanatory variable (X) and instrumental variable (Z) were sampled from independent standard uniform distributions. The endogenous treatment variable (T) was generated by $T_i^* = -1.5 \cdot X_i - 0.75 \cdot Z_i + u_i > 0$, where the values of u_i are draws from an i.i.d. standard normal distribution. The person ability parameters were calculated as $\theta_i^* = -1.8 \cdot T_i + 1 \cdot X_i + 1.9 \cdot u_i + e_i^*$. Values of e_i^* were generated by drawing from an i.i.d. normal distribution with mean zero and variance exp(2.6), where exp(•) is defined above. Responses to the items were simulated by comparing the item response probabilities calculated according to Eq. (6) with random numbers from the standard uniform distribution. The item responses were assigned a value of one if the probability was greater than the random number, and zero otherwise.

Item-difficulty and person-ability parameters were estimated using the MML and Bayesian EAP methods, respectively, so that 2SLS and TSM methods could be estimated for comparison purposes. The dependent variable in these models is the estimated person-ability parameter. The BRSM was estimated with the behavioral component's constant constrained to zero so that all the item-difficulty parameters were identified. The model can also be estimated under different

Table 1
Estimates and standard errors of the latent regression model parameters from the BRSM, TSM, and 2SLS methods.

Parameter	True value	BRSM		TSM		2SLS	
		Coef.	SE	Coef.	SE	Coef.	SE
				N = 500			
$\beta_{Treatment}$	−1.800	−0.981	(1.530)	−2.065	(4.172)	−1.276	(2.099)
β_X	1.000	1.475	(1.214)	2.069	(2.786)	0.529	(1.281)
$\beta_{Constant}$	−1.500	−2.145	(1.255)	−2.152	(3.204)	2.442	(1.566)
λ	1.900	1.429	(0.939)	2.592	(2.589)	–	–
				N = 1,000			
$\beta_{Treatment}$	−1.800	−1.541	(1.233)	−2.210	(2.992)	−1.209	(1.370)
β_X	1.000	1.136	(0.915)	1.987	(1.964)	0.562	(0.836)
$\beta_{Constant}$	−1.500	−1.707	(0.978)	−2.008	(2.290)	2.367	(1.023)
λ	1.900	1.745	(0.762)	2.633	(1.857)	–	–
				N = 5,000			
$\beta_{Treatment}$	−1.800	−1.970	(0.666)	−2.306	(1.296)	−1.126	(0.578)
β_X	1.000	0.901	(0.461)	1.926	(0.847)	0.606	(0.353)
$\beta_{Constant}$	−1.500	−1.366	(0.515)	−1.877	(0.991)	2.296	(0.432)
λ	1.900	2.004	(0.412)	2.672	(0.804)	–	–
				N = 10,000			
$\beta_{Treatment}$	−1.800	−1.948	(0.482)	−2.348	(0.905)	−1.113	(0.405)
β_X	1.000	0.926	(0.331)	1.923	(0.593)	0.620	(0.248)
$\beta_{Constant}$	−1.500	−1.396	(0.372)	−1.869	(0.693)	2.280	(0.303)
λ	1.900	1.998	(0.298)	2.711	(0.562)	–	–

Notes: Models estimated using 1,000 replications of simulated data with varying number of persons and a constant instrument length of eight items. BRSM = behavioral Rasch selection model; TSM = Tobit selection model; and 2SLS = two stage least squares.

normalizing assumptions. For example, one of the item-difficulty parameters could be set to zero, identifying the behavioral constant. Only estimates of the latent regression model parameters are described in the following discussion. Simulation results for the other parameters are available from the author upon request.

Estimates of the latent regression model parameters and their standard errors from the BRSM, TSM, and 2SLS methods are given in Table 1. The parameters are generally underestimated by the 2SLS and BRSM methods (in samples under 1000 persons), and overestimated by the TSM method. In small samples (500 persons), no method is closer to the true value of the parameters than the other methods. As the sample size increases, the BRSM parameter estimates become closer to the true values, while the 2SLS and TSM parameter values continue to underestimate and overestimate the true values, respectively. In all simulations, the standard errors were smaller for the BRSM and 2SLS methods than the standard errors of the TSM method.

Efficiency gains of the BRSM method relative to the TSM and 2SLS methods are listed in Table 2. I calculate the efficiency gains as the ratio of the root mean square error (RMSE) from the TSM or 2SLS methods relative to the BRSM method. RMSE is used as a measure of an estimator's precision in this article, and is calculated by taking the square root of the average squared differences between the parameter estimates and their true values. The ratio is always greater than one for the parameters when comparing the TSM method to the BRSM method, indicating the BRSM method is preferred in terms of RMSE criteria. When comparing the 2SLS method to the BRSM method, the ratio is less than one for samples of 1,000 or fewer persons, but greater than one in samples of 5,000 or more persons. This is not surprising since both methods have similar negative biases and the 2SLS method has moderately smaller variances in small samples. The moderate preference of the 2SLS method, in terms of RMSE criteria, in smaller samples no longer holds as the sample size increases.

The relative efficiency gains of the BRSM method are not without cost. The computational requirements of this method increase with the sample size, length of the instrument, and the number of quadrature points required to evaluate the two-dimensional integral. In the simulations, the BRSM converged in approximately 36 s and 9.5 min, on average, in samples of 500 and 10,000 persons, respectively. Compared to the 2SLS method, where models were consistently estimated in roughly 5 s (on average), these gains in RMSE should be considered carefully. Convergence time for the TSM method were relatively close to those of the 2SLS method.

Table 2
Ratio of RMSE's for selected latent regression model parameters from the TSM and 2SLS methods relative to the BRSM method.

Parameter	Sample size			
	500	1,000	5,000	10,000
$\beta_{Treatment}$	1.6823 [0.7918]	1.4599 [0.7543]	2.0971 [1.1652]	2.4884 [1.4464]
β_X	1.7410 [0.7501]	1.6522 [0.7441]	2.6361 [1.0524]	3.3337 [1.2588]
$\beta_{Constant}$	1.7165 [2.0446]	1.4857 [2.6599]	2.0903 [6.5775]	2.4241 [9.0865]
λ	1.7256 [–]	1.5484 [–]	2.5765 [–]	3.3027 [–]

Notes: Models estimated using 1,000 replications of simulated data with varying number of persons and a constant instrument length of eight items. The ratio of RMSE's is calculated as the ratio of the TSM method to the BRSM method. The ratio of RMSE's is also calculated as the ratio of the 2SLS method to the BRSM method are in brackets. RMSE = root mean square error. 2SLS = two stage least squares.

4. Empirical examples

As an illustration of the BRSM method, I used food insecurity—defined as limited or uncertain access to adequate food because of a lack of money or other resources—and food and nutrition program receipt data from the 2001–2008 Current Population Survey Food Security Supplement (CPS-FSS). Only the necessary details are discussed here. See Coleman-Jensen et al. [11] and Rabbitt and Coleman-Jensen [24] for a detailed discussion of these data and food security measurement practices, respectively.

A sample of 35,571 households with incomes below 185 percent of the federal poverty line with at least one child (under age 18) responded to the CPS-FSS Household Food Security Survey Module

(HFSSM) and the other questions used in the empirical analyses in 2001–2008. The HFSSM consists of 18 food hardship questions designed to capture the conditions and behaviors of households having difficulty meeting their basic food needs at some point during the previous 12 months. In the empirical analyses, I examine the eight child HFSSM food hardship questions that are used to calibrate the U.S. child food insecurity scale. A listing of the HFSSM questions is available from Coleman-Jensen et al. [11], Bickel et al. [4], Rabbitt [22], or Rabbitt and Coleman-Jensen [24].

I primarily focus on the treatment variable indicating whether anyone in the household received Supplemental Nutrition Assistance Program (SNAP, formerly the Food Stamp Program) benefits within the previous 12 months. Although economic theory predicts SNAP receipt reduces food insecurity, there are many reasons why an empirical analysis might differ. The receipt of SNAP benefits is endogenous since it is a behavioral outcome that reflects household choices. An extensive discussion of the conceptual relationship between food insecurity and SNAP receipt, as well as a recent review of the literature and methodological replication study, can be found in Gregory, Rabbitt, and Ribar [15].

As additional explanatory variables, I also include controls for other factors commonly identified as determinants of food insecurity [15,16,32]. The log of real total household income, in 2001 dollars, and an indicator for whether someone in the household owns, or is currently purchasing, the home are included as controls for short- and long-term economic resources, respectively. I also control for the respondent's gender, age (and age-squared), race, ethnicity, and nativity; educational attainment of the most highly educated adult in the household; number of children and number of adults; presence of any elderly member (age 60 or older) and the presence of a disabled member; age of the youngest member of the household; household located in an urban area; state unemployment rate; percent of adults in the household who are employed; and state and year fixed effects. Sample descriptive statistics for these variables are included in Appendix Table B1 for the entire sample and by SNAP receipt.

Identification of the BRSM method in my empirical analyses is achieved using exclusion restrictions in the form of instrumental variables. Several instrumental variables have been proposed for SNAP. They are usually related to national- and state-SNAP polices hypothesized to affect the household's costs and benefits of participation. See Borjas [8], Gregory, Rabbitt, and Ribar [15], Gundersen and Oliveira [17], Rabbitt [22], Ratcliffe, McKernan, and Zhang [25], and Shaefer and Gutierrez [30] for examples of SNAP instruments and discussions of their validity. For the empirical analyses, I use two instrumental variables that have been employed by previous studies. The first instrument, an indicator for whether the household respondent is a non-citizen-immigrant, is related to changes in SNAP eligibility for non-citizens. The second instrument I consider is the state-level median recertification period for SNAP.

4.1. Child food insecurity and endogenous SNAP receipt revisited

Several studies of the relationship between SNAP receipt and food insecurity have led to findings that are statistically imprecise and indistinguishable from zero when employing instrumental variables methods with cross sectional data [15], [22]. Absent the ability to collect larger samples, the added efficiency of the BRSM method provides a promising way to obtain more precise estimates when using cross-sectional data. This is particularly important when examining children's food insecurity because it is a rare event. To obtain insights into the relationship between child food insecurity and SNAP receipt, I modeled the probability of the household's responses to the HFSSM child food hardship items using the BRSM method. The severity of child food insecurity is a latent variable denoted by θ, and it is assumed to be related to SNAP benefit receipt (the treatment, T) and a set of explanatory variables (X) according to Eq. (6).

To establish the implications of failing to correct for the endogeneity of SNAP receipt, I also estimated a baseline specification consistent with the conventional BRM, specified in Eq. (3). Under the BRM, SNAP receipt is assumed to be exogenous, which is anticipated to lead to biased estimates of the latent regression model parameters. While the BRM and BRSM can be estimated in two steps—the first step estimating the item-difficulty parameters, and the second step estimating the latent regression conditional on the values of the item-difficulty parameters [9,38]—I do not implement this approach since it would require adjusting the standard errors for estimation error in the first step. Rather, the measurement and latent regression models are jointly estimated.

Selected severity of child food insecurity parameters and their standard errors are given in Table 3. The complete set of parameter estimates is listed in Appendix Table A1. The columns report parameter estimates, separately, for the BRM and BRSM. The parameter estimates can be interpreted as estimated associations between changes in the explanatory variables and changes in the severity of child food insecurity.

Estimates of the latent regression of the severity of child food insecurity model parameters from the BRM suggest SNAP receipt is positively associated with the severity of child food insecurity (Table 3). Only after correcting for the endogeneity of SNAP receipt in the BRSM do the results suggest SNAP receipt is linked to reductions in the severity of child food insecurity. The parameter estimates also indicate that several other explanatory variables are associated with the severity of child food insecurity.

Endogeneity of SNAP receipt can be formally tested using the BRSM's selection parameter, λ, in Table 3. A test of the significance of this selection parameter indicates that it is statistically significantly different from zero at the one percent level or better, providing strong evidence of the endogeneity of SNAP receipt in this model. In addition, the direction and magnitude of the selection parameter shows the errors between the severity of child food insecurity and SNAP receipt equations are strongly positively correlated. Therefore, an increase in the unobservable variable affecting SNAP receipt, such as unobserved food needs or shocks to household resources, will also increase the severity of child food insecurity.

Causal inference on the SNAP receipt parameter from the BRSM requires appealing to a statistical framework for the estimation of causal parameters known as the Rubin Causal Model (RCM: [27,28]). The RCM is based on the potential outcomes model (POM), which assumes every household in the U.S. population of low-income households with children is potentially exposed to SNAP. Let θ_{i1} measure the severity of child food insecurity for household i receiving SNAP benefits, and θ_{i0} measure that when not receiving SNAP benefits. Since receipt and non-receipt of SNAP are mutually exclusive for household i, only one of two measures of the severity of child food insecurity are available for any given i. The unobservable measure is the counterfactual or hypothetical value of the severity of child food insecurity. The causal effect of SNAP on the severity of child food insecurity for household i is (θ_{i1}–θ_{i0}). Estimation of the average causal effect of SNAP is measured by the average treatment effect (ATE):

$$ATE = E(\theta|T = 1) - E(\theta|T = 0), \tag{7}$$

where E(•) is the expectation operator.

I estimate ATE's for child food insecurity and very low food insecurity among children based on the U.S. Department of Agriculture's food insecurity status classification system. These categories summarize meaningful ranges of the severity of child food insecurity based on the count of affirmed items (raw score) to the HFSSM child items. For more

Table 3
Selected estimates and standard errors of the latent severity of child food insecurity regression model parameters from behavioral Rasch models.

Variable	Behavioral Rasch model	Behavioral Rasch selection model
SNAP receipt, previous 12 months	1.356***	−1.827***
	(0.068)	(0.280)
Female respondent	0.520***	0.611***
	(0.064)	(0.068)
Married, spouse present in household	−0.297***	−0.744***
	(0.069)	(0.080)
Age of respondent	0.166***	0.134***
	(0.015)	(0.016)
Age of respondent, squared	−0.002***	−0.001***
	(0.000)	(0.000)
Black, non-Hispanic respondent	0.303***	0.444***
	(0.078)	(0.088)
Other race, non-Hispanic respondent	0.006	0.309**
	(0.134)	(0.132)
Hispanic respondent	0.400***	0.403***
	(0.094)	(0.099)
Highest educated adult in the household, less than a high school diploma or equivalent	0.316***	0.365***
	(0.076)	(0.083)
Highest educated adult in the household, some college	−0.036	−0.073
	(0.065)	(0.067)
Highest educated adult in the household, bachelor's degree or higher	−0.864***	−1.277***
	(0.104)	(0.111)
Immigrant respondent	0.095	−0.307***
	(0.089)	(0.098)
Number of adults in household	0.037	0.079**
	(0.036)	(0.039)
Number of children in household	0.351***	0.565***
	(0.026)	(0.033)
Elderly member in household	−0.256*	−0.305**
	(0.138)	(0.149)
Disabled member in household	0.842***	1.363***
	(0.099)	(0.112)
Age of youngest member in household	0.077***	0.064***
	(0.007)	(0.007)
Household in metropolitan area	0.366***	0.447***
	(0.072)	(0.072)
State unemployment rate	0.155***	0.247***
	(0.059)	(0.059)
Percent of adults in the household who are employed	−0.451***	−1.047***
	(0.082)	(0.100)
Natural log of real total household income $2001	−0.320***	−0.862***
	(0.043)	(0.068)
Home owned by a member of the household	−0.690***	−1.105***
	(0.064)	(0.077)
σ	3.880***	3.689***
	(0.041)	(0.056)
λ		1.913***
		(0.166)
State fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Log-likelihood function value	−47,945.41	−63,388.11
Number of households	35,571	35,571
Wald test, joint significance of instrumental variables		58.18
		[0.000]

Notes: Models estimated using weighted data on low-income households with children from the 2001–2008 CPS-FSS. Standard errors are in parenthesis and p-values are in brackets. The constant in the latent regression model was normalized to zero so that all the item-difficulty parameters could be estimated. For a listing of the explanatory variables used in the latent regression model, see Appendix Table B.1.

* Significant at 0.10 level.

** Significant at 0.05 level.

*** Significant at 0.01 level.

details on how these categories are defined, see Coleman-Jensen et al. [11]. The ATE's use parameter estimates from the BRSM to calculate the probability that household i's item response vector is greater than a raw score threshold, defined for each category, and assuming all households received and did not receive SNAP [22]. The probabilities are differenced and averaged based on Eq. (7) to obtain measures of the ATE.

Table 4 lists parameter estimates, standard errors, and ATE's for SNAP receipt from the BRM and BRSM. ATE's from the BRM should be

Table 4
Average treatment effect estimates of SNAP receipt within the previous 12 months on child food insecurity and very low food insecurity among children.

Variable	Behavioral Rasch model	Behavioral Rasch selection model
SNAP receipt, previous 12 months	1.356***	−1.827***
	(0.068)	(0.280)
	[0.091]	[−0.106]
	[0.010]	*[−0.014]*
Observable control variables	Yes	Yes
State fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Log-likelihood	−47,945.41	−63,388.11
Number of households	35,571	35,571

Notes: Models estimated using weighted data on low-income households with children from the 2001–2008 CPS-FSS. Standard errors are in parenthesis. Estimated average treatment effects on child food insecurity and very low food security among children are in brackets and italicized brackets, respectively. The constant in the latent regression model was normalized to zero so that all the item-difficulty parameters could be estimated. For a listing of the explanatory variables used in the latent regression model, see Appendix Table B.1.
*Significant at 0.10 level.
**Significant at 0.05 level.
*** Significant at 0.01 level.

interpreted with caution since this model does not correct for the endogeneity of SNAP receipt, while ATEs from the BRSM may be interpreted as causal estimates of the relationship between SNAP and the severity of child food insecurity since the model parameters are unbiased and consistently estimated under the assumption that SNAP receipt is endogenous.

The ATEs for child food insecurity and very low food security among children in Table 4 measure the average causal effect of SNAP receipt on child food insecurity and very low food security among children, respectively. Based on the ATEs, SNAP receipt among low-income households with children reduces the probability of a household experiencing child food insecurity and very low food security among children by 10.6 and 1.4 percent, respectively.

4.2. Child food insecurity and endogenous SNAP DIF

As demonstrated above, the BRSM is particularly useful for conducting causal inference on the relationship between latent ability and a treatment variable, assuming certain assumptions are satisfied. The model may also be used in analyses of DIF when the group indicator variable is endogenous. Addressing endogenous DIF requires re-specifying the BRSM as follows:

$$P(Y_{ij}=1|T_i,X_i,u_i,e_i^*) = \frac{\exp\left(\beta_T T_i + \beta_X' X_i + \lambda u_i + e_i^* - \sum_{k=1}^{J}\gamma_k (T_i \cdot X_{kj}) - \sum_{k=1}^{J}\delta_k X_{kj}\right)}{1+\exp\left(\beta_T T_i + \beta_X' X_i + \lambda u_i + e_i^* - \sum_{k=1}^{J}\gamma_k (T_i \cdot X_{kj}) - \sum_{k=1}^{J}\delta_k X_{kj}\right)}. \tag{8}$$

The item-difficulty parameter for the reference group for item k equals δ_k, and in the focal group it equals $(\gamma_k + \delta_k)$. Because the treatment variable (T_i) accounts for overall group differences in ability in the model, any deviation of γ_k from zero indicates uniform DIF for item k in the focal group (treated) compared to the reference group (untreated). The model, as specified in Eq. (8), also controls for observable differences in ability based on the other explanatory variables contained in X_i. DIF of one or more of the items may be explored using the interaction terms in the model. For a review of the methods using this approach to detect DIF, see Van den Noortgate and De Boeck [35].

Selected estimates from the BRM and BRSM, which are parameterized to detect uniform DIF related to SNAP receipt, are given in Table 5. A complete set of estimates is available in Appendix Table A2. The BRM estimates the model parameters under the assumption that DIF between SNAP recipients and non-recipients is exogenous, while the BRSM assumes SNAP related DIF is endogenous. Differences in the severity of child food insecurity between SNAP recipients and non-recipients are affected by the endogeneity of SNAP receipt. After correcting for the endogeneity of SNAP receipt, the direction of the relationship between the severity of child food insecurity and SNAP receipt changes from positive to negative, indicating SNAP receipt reduces the severity of child food insecurity. The selection parameter is also highly statistically significant (at the one percent level or better), indicating the BRM is rejected in favor of the BRSM.

While the endogeneity of SNAP receipt has an impact on estimates of the differences in the severity of child food insecurity, there does not appear to be much of an effect on SNAP related DIF. Under the BRM specification, no uniform DIF related to SNAP is detected, but the BRSM shows signs of moderate SNAP-DIF on the item “child(ren) skipped meals.” In this case, correcting for the endogeneity of SNAP receipt has resulted in the uniform-DIF parameter on this item increasing by approximately 80 percent and it is now marginally statistically significant.

5. Discussion

Several methods have been proposed for estimating the parameters from a latent regression of ability on a set of explanatory variables. These methods have been extensively tested and compared using actual and simulated data when the explanatory variables are exogenous [9,38]. The general consensus from these studies is that the BRM offers greater accuracy and precision than linear regression models of the person ability estimates. The growing need for studies with a causal focus in the social sciences suggests the need for methods that are similar, but address the endogeneity of the latent regression model’s explanatory variables. Causal inference based on model parameters estimated using observational data is particularly challenging because one or more explanatory variables may be endogenous.

Fortunately, methods have been developed for addressing the endogeneity of explanatory variables in the latent regression model context for IRT models. In this article, I used simulated and actual data on food insecurity from the 2001–2008 CPS-FSS to illustrate the current methodology. Simulations were conducted under the assumption that the latent regression model contained a endogenous binary variable. Future work should consider relaxing this to include continuous endogenous variables or more than one endogenous variable. Several instrumental variable methods were applied to the simulated data and compared. The simulation analysis suggests the newly developed BRSM method is more accurate and precise than the 2SLS and TSM methods in most cases. An exception to this occurs in smaller samples, where the 2SLS and BRSM methods appear to be on equal footing. When combined with the fact that the 2SLS method is more computationally efficient than the BRSM method, the 2SLS method may be preferred in small samples. Yet, it is important to note that the 2SLS method consistently underestimated the latent regression model parameters.

Empirical examples were presented in this article using actual data to illustrate the usefulness of the BRSM. The first example revisited the

Table 5
Selected estimates of uniform DIF parameters from behavioral Rasch models.

Variable	Behavioral Rasch model	Behavioral Rasch selection model
	Ability difference (latent regression) parameters	
SNAP receipt, previous 12 months	1.214***	−1.679***
	(0.285)	(0.406)
Constant	−14.433***	−7.770***
	(0.883)	(0.920)
	Uniform DIF parameters	
SNAP12 • "relied on few kinds of low-cost food"	−0.406	−0.178
	(0.285)	(0.284)
SNAP12 • "could feed child(ren) balanced meal"	0.160	0.325
	(0.283)	(0.283)
SNAP12 • "child(ren) were not eating enough"	−0.005	0.157
	(0.283)	(0.282)
SNAP12 • "cut size of child(ren)'s meals"	−0.071	0.153
	(0.284)	(0.285)
SNAP12 • "child(ren) were hungry"	−0.057	0.061
	(0.286)	(0.286)
SNAP12 • "child(ren) skipped meals"	0.281	0.518*
	(0.291)	(0.291)
SNAP12 • "child(ren) skipped meals, 3+ months"	0.087	0.305
	(0.297)	(0.297)
SNAP12 • "child(ren) did not eat for whole day"		
	Error components	
λ		1.854***
		(0.171)
Observable control variables	Yes	Yes
State fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Log-likelihood function value	−47,900.42	−63,349.59
Number of households	35,571	35,571
Wald test, joint significance of DIF parameters	88.78	74.07
	[0.000]	[0.000]

Notes: Models estimated using weighted data on low-income households with children from the 2001–2008 CPS-FSS. Standard errors are in parenthesis and p-values are in brackets. For a listing of the explanatory variables in the latent regression model, see Appendix Table B.1. SNAP = Supplemental Nutrition Assistance Program.
*Significant at 0.10 level.
**Significant at 0.05 level.
***Significant at 0.01 level.

relationship between SNAP receipt and child food insecurity. Findings from this example demonstrate the importance of addressing the endogeneity of explanatory variables in the BRM. The BRM, which assumes SNAP receipt is exogenous, produced a counterintuitive estimate for SNAP receipt that has been well-documented in the empirical literature [15]. After correcting for the endogeneity of SNAP receipt using the BRSM, I found SNAP receipt reduced the probability of low-income households with children experiencing child food insecurity and very low food security among children by 10.6 and 1.4 percent, respectively. These results can be interpreted as causal since the endogeneity of SNAP receipt was addressed and ATEs were constructed using the model parameters.

The second empirical example demonstrates how the BRSM can be used to estimate DIF parameters in the presence of an endogenous group indicator. My findings suggest the BRSM is useful for studies of DIF when the group indicator is endogenous. Prior to correcting for the endogeneity of the SNAP group variable, no uniform DIF between SNAP recipient and non-recipient households was detected. After controlling for the endogeneity of the SNAP group variable, the sign of the group variable, which was included to account for overall differences in the severity of children's food insecurity between the reference and focal groups, switched sign from positive to negative, and moderate SNAP related uniform DIF was detected for the item, "child(ren) skipped meals."

The discussion of the BRSM in this article was intentionally general to emphasize its utility in many settings. The model can be modified to include more flexible IRT models and polytomous item responses. The primary limitation of the model is the need for specialized software and the computational burden associated it. In some cases, it may also be difficult to select valid instruments necessary for model identification. This limitation is true, however, of any instrumental variables model. Regardless of these limitations, the model has the potential to open new areas of research in the fields of economics, psychometrics, and the other social sciences.

Acknowledgements

The views expressed in this article are those of the author and do not necessarily reflect those of the Economic Research Service or the U.S. Department of Agriculture. The author would like to thank David C. Ribar at the University of Melbourne, M. Taylor Rhodes at the U.S. Department of Agriculture's Economic Research Service, and participants at the 2016 U.C. Berkeley IMEKO Joint Symposium for helpful comments and suggestions.

Appendix A. Parameter estimates and standard errors from behavioral Rasch models of the severity of child food insecurity

Table A1
Estimates and standard errors from behavioral Rasch models parameterized to estimate the causal effect of the SNAP receipt on the severity of child food insecurity.

Variables	Behavioral Rasch model	Behavioral Rasch selection model	
	Severity of latent child food insecurity	SNAP participation	Severity of latent child food insecurity
SNAP receipt, previous 12 months	1.356***		−1.827***
	(0.068)		(0.280)
State SNAP median recertification period		0.018***	
		(0.005)	
Respondent is a non-citizen immigrant		−0.249***	
		(0.037)	
Female respondent	0.520***	0.146***	0.611***
	(0.064)	(0.019)	(0.068)
Married, spouse present in household	−0.297***	−0.392***	−0.744***
	(0.069)	(0.020)	(0.080)
Age of respondent	0.166***	−0.017***	0.134***
	(0.015)	(0.004)	(0.016)
Age of respondent, squared	−0.002***	0.000***	−0.001***
	(0.000)	(0.000)	(0.000)
Black, non-Hispanic respondent	0.303***	0.151***	0.444***
	(0.078)	(0.024)	(0.088)
Other race, non-Hispanic respondent	0.006	0.126***	0.309**
	(0.134)	(0.037)	(0.132)
Hispanic respondent	0.400***	0.009	0.403***
	(0.094)	(0.029)	(0.099)
Highest educated adult less than high school diploma or equivalent	0.316***	0.079***	0.365***
	(0.076)	(0.024)	(0.083)
Highest educated adult some college	−0.036	−0.040**	−0.073
	(0.065)	(0.019)	(0.067)
Highest educated adult bachelor's degree or higher	−0.864***	−0.329***	−1.277***
	(0.104)	(0.031)	(0.111)
Immigrant respondent	0.095	−0.187***	−0.307***
	(0.089)	(0.036)	(0.098)
Number of adults in household	0.037	0.075***	0.079**
	(0.036)	(0.011)	(0.039)
Number of children in household	0.351***	0.198***	0.565***
	(0.026)	(0.008)	(0.033)
Elderly member in household	−0.256*	−0.022	−0.305**
	(0.138)	(0.042)	(0.149)
Disabled member in household	0.842***	0.337***	1.363***
	(0.099)	(0.030)	(0.112)
Age of youngest member in household	0.077***	−0.026***	0.064***
	(0.007)	(0.002)	(0.007)
Household in metropolitan area	0.366***	−0.027	0.447***
	(0.072)	(0.020)	(0.072)
State unemployment rate	0.155***	0.054***	0.247***
	(0.059)	(0.017)	(0.059)
Percent of adults in the household who are employed	−0.451***	−0.595***	−1.047***
	(0.082)	(0.024)	(0.100)
Natural log of real total household income $2001	−0.320***	−0.538***	−0.862***
	(0.043)	(0.012)	(0.068)
Home owned by a member of the household	−0.690***	−0.448***	−1.105***
	(0.064)	(0.019)	(0.077)
Constant		5.236***	
		(0.169)	
"relied on few kinds of low-cost food"	3.697***		−3.235***
	(0.862)		(0.881)
"could feed child(ren) balanced meal"	5.628***		−1.289
	(0.863)		(0.881)
"child(ren) were not eating enough"	7.581***		0.721
	(0.864)		(0.881)
"cut size of child(ren)'s meals"	10.202***		3.291***
	(0.865)		(0.881)
"child(ren) were hungry"	10.794***		3.790***
	(0.865)		(0.882)
"child(ren) skipped meals"	11.800***		4.783***
	(0.867)		(0.883)
"child(ren) skipped meals, 3+ months	12.402***		5.480***
	(0.868)		(0.884)

(*continued on next page*)

Table A1 (*continued*)

Variables	Behavioral Rasch model	Behavioral Rasch selection model	
	Severity of latent child food insecurity	SNAP participation	Severity of latent child food insecurity
"child(ren) did not eat for whole day"	14.514***		7.487***
	(0.877)		(0.891)
σ	3.880***		3.689***
	(0.041)		(0.056)
λ			1.913***
			(0.166)
State fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Log-likelihood function value	−47,945.41	−63,388.11	
Number of households	35,571	35,571	
Wald test, joint significance of instrumental variables		58.18	
		[0.000]	

Notes: Models estimated using weighted data on low-income households with children from the 2001–2008 CPS-FSS. Standard errors are in parenthesis and p-values are in brackets. SNAP = Supplemental Nutrition Assistance Program.

* Significant at 0.10 level.

** Significant at 0.05 level.

*** Significant at 0.01 level.

Table A2
Estimates and standard errors from behavioral Rasch models parameterized to examine uniform DIF related to endogenous SNAP receipt.

Variable	Behavioral Rasch model	Behavioral Rasch selection model	
	Severity of latent child food insecurity	SNAP participation	Severity of latent child food insecurity
Ability difference (latent regression) parameters			
SNAP receipt, previous 12 months	1.214***		−1.679***
	(0.285)		(0.406)
State SNAP median recertification period		0.018***	
		(0.005)	
Respondent is a non-citizen immigrant		−0.250***	
		(0.037)	
Female respondent	0.515***	0.146***	0.606***
	(0.063)	(0.019)	(0.068)
Married, spouse present in household	−0.293***	−0.392***	−0.727***
	(0.069)	(0.020)	(0.081)
Age of respondent	0.164***	−0.017***	0.135***
	(0.015)	(0.004)	(0.016)
Age of respondent, squared	−0.002***	0.000***	−0.001***
	(0.000)	(0.000)	(0.000)
Black, non-Hispanic respondent	0.290***	0.151***	0.435***
	(0.078)	(0.024)	(0.088)
Other race, non-Hispanic respondent	0.023	0.126***	0.306**
	(0.134)	(0.037)	(0.131)
Hispanic respondent	0.408***	0.009	0.402***
	(0.094)	(0.029)	(0.099)
Highest educated adult in the household less than high school diploma or equivalent	0.313***	0.079***	0.363***
	(0.076)	(0.024)	(0.083)
Highest educated adult in the household some college	−0.031	−0.040**	−0.070
	(0.065)	(0.019)	(0.067)
Highest educated adult in the household bachelor's degree or higher	−0.860***	−0.329***	−1.264***
	(0.104)	(0.031)	(0.111)
Immigrant respondent	0.091	−0.186***	−0.296***
	(0.089)	(0.036)	(0.098)
Number of adults in household	0.032	0.075***	0.075*
	(0.036)	(0.011)	(0.039)
Number of children in household	0.351***	0.198***	0.559***
	(0.026)	(0.008)	(0.033)
Elderly member in household	−0.250*	−0.022	−0.303**
	(0.138)	(0.042)	(0.148)
Disabled member in household	0.850***	0.338***	1.357***
	(0.099)	(0.030)	(0.112)
Age of youngest member in household	0.078***	−0.026***	0.065***
	(0.007)	(0.002)	(0.007)

(*continued on next page*)

Table A2 (*continued*)

Variable	Behavioral Rasch model	Behavioral Rasch selection model	
	Severity of latent child food insecurity	SNAP participation	Severity of latent child food insecurity
Household in metropolitan area	0.363***	−0.027	0.449***
	(0.072)	(0.020)	(0.072)
State unemployment rate	0.152***	0.054***	0.244***
	(0.059)	(0.017)	(0.059)
Percent of adults in the household who are employed	−0.444***	−0.596***	−1.029***
	(0.082)	(0.024)	(0.101)
Natural log of real total household income $2001	−0.316***	−0.538***	−0.843***
	(0.043)	(0.012)	(0.069)
Home owned by a member of the household	−0.691***	−0.448***	−1.094***
	(0.064)	(0.019)	(0.078)
Constant	−14.433***	5.239***	−7.770***
	(0.883)	(0.169)	(0.920)
Uniform DIF and reference group parameters (measurement component)			
SNAP12 • "relied on few kinds of low-cost food"	−0.406		−0.178
	(0.285)		(0.284)
SNAP12 • "could feed child(ren) balanced meal"	0.160		0.325
	(0.283)		(0.283)
SNAP12 • "child(ren) were not eating enough"	−0.005		0.157
	(0.283)		(0.282)
SNAP12 • "cut size of child(ren)'s meals"	−0.071		0.153
	(0.284)		(0.285)
SNAP12 • "child(ren) were hungry"	−0.057		0.061
	(0.286)		(0.286)
SNAP12 • "child(ren) skipped meals"	0.281		0.518*
	(0.291)		(0.291)
SNAP12 • "child(ren) skipped meals, 3 + months"	0.087		0.305
	(0.297)		(0.297)
SNAP12 • "child(ren) did not eat for whole day"			
"relied on few kinds of low-cost food"	−10.679***		−10.688***
	(0.214)		(0.216)
"could feed child(ren) balanced meal"	−8.953***		−8.926***
	(0.211)		(0.213)
"child(ren) were not eating enough"	−6.935***		−6.849***
	(0.208)		(0.210)
"cut size of child(ren)'s meals"	−4.278***		−4.275***
	(0.208)		(0.211)
"child(ren) were hungry"	−3.691***		−3.732***
	(0.209)		(0.212)
"child(ren) skipped meals"	−2.845***		−2.910***
	(0.211)		(0.214)
"child(ren) skipped meals, 3 + months"	−2.151***		−2.209***
	(0.215)		(0.219)
"child(ren) did not eat for whole day"			
Error Components			
σ	3.885***		3.705***
	(0.041)		(0.056)
λ			1.854***
			(0.171)
Observable control variables	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Log-likelihood function value	−47,900.42	−63,349.59	
Number of households	35,571	35,571	
Wald test, joint significance of DIF parameters	88.78	74.07	
	[0.000]	[0.000]	
Wald test, joint significance of instrumental variables		58.29	
		[0.000]	

Notes: Models estimated using weighted data on low-income households with children from the 2001–2008 CPS-FSS. Standard errors are in parenthesis and p-values are in brackets. DIF = differential item functioning, and SNAP = Supplemental Nutrition Assistance Program.

* Significant at 0.10 level.

** Significant at 0.05 level.

*** Significant at 0.01 level.

Appendix B. Means and standard deviations of the analysis sample

Table B1
Means and standard deviations of the empirical example data.

Variable	All households	No SNAP last year	Received SNAP last year
	SNAP participation		
SNAP receipt, previous 12 months	0.307	0.000	1.000
	(0.461)	(0.000)	(0.000)
	Food insecurity		
"relied on few kinds of low-cost food"	0.346	0.286	0.482
	(0.476)	(0.452)	(0.500)
"could feed child(ren) balanced meal"	0.208	0.173	0.287
	(0.406)	(0.378)	(0.452)
"child(ren) were not eating enough"	0.107	0.083	0.160
	(0.308)	(0.275)	(0.367)
"cut size of child(ren)'s meals"	0.033	0.023	0.054
	(0.178)	(0.150)	(0.226)
"child(ren) were hungry"	0.024	0.017	0.040
	(0.153)	(0.129)	(0.196)
"child(ren) skipped meals"	0.014	0.011	0.021
	(0.117)	(0.102)	(0.145)
"child(ren) skipped meals, 3+ months"	0.010	0.007	0.016
	(0.099)	(0.083)	(0.127)
"child(ren) did not eat for whole day"	0.003	0.002	0.005
	(0.052)	(0.043)	(0.068)
Food insecurity among children	0.205	0.167	0.291
	(0.404)	(0.373)	(0.454)
Very low food security among children	0.017	0.012	0.029
	(0.130)	(0.110)	(0.167)
	Standard explanatory variables		
Female respondent	0.626	0.567	0.760
	(0.484)	(0.496)	(0.427)
Married, spouse present in household	0.487	0.574	0.290
	(0.500)	(0.494)	(0.454)
Age of respondent	37.399	38.260	35.456
	(11.483)	(11.289)	(11.677)
White, non-Hispanic (reference)	0.439	0.455	0.403
	(0.496)	(0.498)	(0.491)
Black, non-Hispanic respondent	0.218	0.170	0.327
	(0.413)	(0.376)	(0.469)
Other race, non-Hispanic respondent	0.050	0.053	0.043
	(0.219)	(0.225)	(0.204)
Hispanic respondent	0.292	0.322	0.226
	(0.455)	(0.467)	(0.418)
Highest educated adult in the household less than high school diploma or equivalent	0.190	0.166	0.245
	(0.392)	(0.372)	(0.430)
Highest educated adult in the household high school diploma or equivalent (reference)	0.385	0.373	0.413
	(0.487)	(0.484)	(0.492)
Highest educated adult in the household some college	0.325	0.336	0.299
	(0.468)	(0.472)	(0.458)
Highest educated adult in the household bachelor's degree or higher	0.100	0.125	0.044
	(0.300)	(0.331)	(0.205)
Immigrant respondent	0.275	0.323	0.167
	(0.447)	(0.468)	(0.373)
Number of adults in household	1.984	2.094	1.737
	(0.910)	(0.906)	(0.868)
Number of children in household	2.131	2.062	2.285
	(1.140)	(1.080)	(1.250)
Elderly member in household	0.082	0.087	0.069
	(0.274)	(0.282)	(0.254)
Disabled member in household	0.082	0.054	0.144
	(0.274)	(0.226)	(0.352)
Age of youngest member in household	6.247	6.623	5.398
	(5.076)	(5.137)	(4.829)
Household in metropolitan area	0.780	0.787	0.766
	(0.414)	(0.410)	(0.423)
State unemployment rate	5.374	5.360	5.405
	(0.980)	(0.979)	(0.981)
Percent of adults in the household who are employed	0.606	0.671	0.458
	(0.376)	(0.336)	(0.419)

(*continued on next page*)

Table B1 (*continued*)

Variable	All households	No SNAP last year	Received SNAP last year
Natural log of real total household income income $2001	9.620	9.829	9.149
	(0.764)	(0.635)	(0.818)
Home owned by a member of the household	0.421	0.502	0.236
	(0.494)	(0.500)	(0.425)
Instrumental variables			
State SNAP median recertification period	8.032	8.041	8.011
	(3.186)	(3.225)	(3.098)
Respondent is a non-citizen immigrant	0.184	0.218	0.109
	(0.388)	(0.413)	(0.311)
Number of households	35,571	24,715	10,856

Notes: Means and standard deviations (in parenthesis) calculated using weighted household data from the 2001–2008 CPS-FSS.

References

[1] E.B. Andersen, Sufficient statistics and latent trait models, Psychometrika 42 (1) (1977) 69–81.
[2] E. Andersen, M. Madsen, Estimating the parameters of the latent population distribution, Psychometrika 42 (3) (1977) 357–374.
[3] D. Andrich, A rating formulation for ordered response categories, Psychometrika 43 (4) (1978) 561–573.
[4] G. Bickel, M. Nord, C. Price, W. Hamilton, J. Cook, Guide to Measuring Household Food Security: Revised 2000, U.S. Department of Agriculture, Food and Nutrition Service, 2000.
[5] A. Birnbaum, Some latent trait models and their use in inferring an examinee's ability, in: F.M. Lord, M.R. Novick (Eds.), Statistical Theories of Mental Test Scores, Addison-Wesley, Reading, MA, 1968.
[6] R.D. Bock, M. Aitkin, Marginal maximum likelihood estimation of item parameters: application of an EM algorithm, Psychometrika 46 (4) (1981) 443–459.
[7] R.D. Bock, R.J. Mislevy, Adaptive EAP estimation of ability in a microcomputer environment, Appl. Psycholog. Measur. 6 (4) (1982) 431–444.
[8] G.J. Borjas, Food insecurity and public assistance, J. Public Econom. 88 (7) (2004) 1421–1443.
[9] K.B. Christensen, From rasch scores to regression, J. Appl Measur. 7 (2) (2006) 184–191.
[10] K.B. Christensen, J.B. Bjorner, S. Kreiner, J.H. Petersen, Latent regression in log-linear Rasch models, Commun. Statist.-Theory and Methods 33 (6) (2004) 1295–1313.
[11] A. Coleman-Jensen, M.P. Rabbitt, C.A. Gregory, A. Singh, Household Food Security in the United States in 2016. U.S. Department of Agriculture, Economic Research Service, 2017.
[12] R.J. de Ayala, The Theory and Practice of Item Response Theory, Guilford Publications, New York, NY, 2013.
[13] P. De Boeck, M. Wilson, Descriptive and explanatory item response models, *Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach* (2004) 43–74.
[14] R.A. DePolt, R.A. Moffitt, D.C. Ribar, Food stamps, temporary assistance for needy families and food hardships in three American cities, Pacif. Econom. Rev. 14 (4) (2009) 445–473.
[15] C.A. Gregory, M.P. Rabbitt, D.C. Ribar, The supplemental nutrition assistance program and food insecurity, in: J.P. Ziliak, J. Bartfield, T. Smeeding, C. Gundersen (Eds.), The New Safety Net: Why SNAP Matters, Stanford University Press, Stanford, CA, 2016.
[16] C. Gundersen, B. Kreider, J. Pepper, The economics of food insecurity in the United States, Appl. Econom. Perspect. Policy 33 (3) (2011) 281–303.
[17] C. Gundersen, V. Oliveira, The food stamp program and food insufficiency, Am. J. Agri. Econom. 83 (4) (2001) 875–887.
[18] H. Hoijtink, Linear and repeated measures models for the person parameters, Rasch Models—Foundations, Recent Developments, and Applications, Springer, New York, 1995, pp. 203–214.
[19] A. Kamata, Item analysis by the hierarchical generalized linear model, J. Edu. Measur. 38 (1) (2001) 79–93.
[20] K.S. Maier, A Rasch hierarchical measurement model, J. Edu. Behav. Statist. 26 (3) (2001) 307–330.
[21] R.A. Moffitt, D.C. Ribar, Rasch analyses of very low food security among households and children in the three city study, South. Econom. J. 82 (4) (2016) 1123–1146.
[22] M.P. Rabbitt, Measuring the Effect of Supplemental Nutrition Assistance Program Participation on Food Insecurity Using a Behavioral Rasch Selection Model, Unpublished Manuscript, University of North Carolina, Greensboro, 2014.
[23] M.P. Rabbitt, A note on the usefulness of the behavioral Rasch selection model for causal inference in the social sciences, J. Phys.: Conf. Ser. 772 (1) (2016) 1–6.
[24] M.P. Rabbitt, A. Coleman-Jensen, Rasch analyses of the Spanish translation of the U.S. household food security survey module, J. Econom. Social Measur. 42 (2) (2017) 171–187.
[25] C. Ratcliffe, S.M. McKernan, S. Zhang, How much does the supplemental nutrition assistance program reduce food insecurity? Am. J. Agri. Econom. 93 (4) (2011) 1082–1098.
[26] G. Rasch (1960/1980), Probabilistic Models for Some Intelligence and Attainment Tests, Danish Institute for Educational Research, Copenhagen, (Expanded edition, Chicago: University of Chicago Press, 1980).
[27] D.B. Rubin, Estimating causal effects of treatments in randomized and non-randomized studies, J. Edu. Psychol. 66 (5) (1974) 688–701.
[28] D.B. Rubin, Bayesian inference for causal effects: the role of randomization, Ann. Statist. 6 (1) (1978) 34–58.
[29] L. Sanathanan, S. Blumenthal, The logistic model and estimation of latent structure, J. Am. Statist. Assoc. 73 (364) (1978) 794–799.
[30] H.L. Shaefer, I.A. Gutierrez, The supplemental nutrition assistance program and material hardships among low-income households with children, Social Serv. Rev. 87 (4) (2013) 753–779.
[31] T.J. Seong, Sensitivity of marginal maximum likelihood estimation of item and ability parameters to the characteristics of the prior ability distributions, Appl. Psycholog. Measur. 14 (3) (1990) 299–311.
[32] M.D. Smith, M.P. Rabbitt, A. Coleman-Jensen, Who are the World's food insecure? New evidence from the food and agriculture organization's food insecurity experience scale, World Develop. 93 (2017) 402–412.
[33] J.V. Terza, Parametric nonlinear regression with endogenous switching, Econom. Rev. 28 (6) (2009) 555–580.
[34] J. Tobin, Estimation of relationships for limited dependent variables, Econometrica 26 (1) (1958) 24–36.
[35] W. Van den Noortgate, P. De Boeck, Assessing and explaining differential item functioning using logistic mixed models, J. Edu. Behav. Statist. 30 (4) (2005) 443–464.
[36] W.J. Van der Linden, Handbook of item response theory, Volume One: Models, CRC Press, Boca Raton, FL, 2016.
[37] P. Wilde, M. Nord, The effect of food stamps on food security: a panel data approach, Rev. Agri. Econom. 27 (3) (2005) 425–432.
[38] A.H. Zwinderman, A generalized Rasch model for manifest predictors, Psychometrika 56 (4) (1991) 589–600.
[39] A.H. Zwinderman, Response Models with Manifest Predictors, Handbook of Modern Item Response Theory, Spring-Verlag, 1997, pp. 245–256.