British Journal of Mathematics & Computer Science
4(10): 1437-1449, 2014

SCIENCEDOMAIN international
www.sciencedomain.org

An Item Response Model for Understanding Item Non-response in Ghanaian Surveys

N. K. Frempong[1*], I. Wahab Abdul[1] and E. Okyere[2]

[1]*Department of Mathematics, College of Science, KNUST, Kumasi, Ghana.*
[2]*Department of Computer Science, Catholic University College of Ghana, Sunyani – Brong Ahafo, Ghana.*

Original Research Article

Received: 24 July 2013
Accepted: 17 October 2013
Published: 25 March 2014

Abstract

Survey research has been widely used in public opinion research in Ghana. Ghanaian researchers are happy about data richness and they are also concerned about data quality. In this paper Item Response Theory (IRT) has been used to identify the most appropriate IRT model for understanding item. The techniques are appropriate and practical.

A questionnaire data on Ghana collected in the 5^{th} wave of the World Values Survey was used for the analysis. The five categories of survey questions that are most difficult to answer by respondents were Life Related Questions, Value Related Questions, Political Related Questions, Income Related Questions and Democracy Related Questions. Missing or 'don't know' responses were assigned a 0 score, and 1 was assigned to answered items. The data was analysed based on four IRT models namely, the constrained Rasch model, the unconstrained Rasch model, the two parameter logistic model, and the three parameter logistic model. These models were explored to determine the most appropriate model for the data. In this paper, the unconstrained Rasch model emerged as the best model for understanding item non-response. We found that, income related questions had the highest difficulty parameter, hence the most difficult category of survey questions to answer. It was also found that, if an individual does not answer a survey question or give a 'don't know' answer, it is not only because of the question's difficulty but also because the respondent doesn't want to answer.

Keywords: Item Non-response, Item Response Theory (IRT), Unconstrained Rasch Model, World Values Survey.

1 Introduction

As statisticians, we use data gathered from surveys to make informed decisions and give recommendations to clients on ways to improve their expected outcomes. In today's survey research industry, we do a lot to help ensure our data meet certain standards: screener questions

**Corresponding author: nkf.cos@gmail.com;*

target the specific audience we want, online panels take many steps to ensure their samples contain the target we need, we weight respondents to match specific population demographics, etc. However, one of the most over-looked problems is that of non-response bias. In data collection, there are two types of non-response: item and unit non-response. Item non-response occurs when certain questions in a survey are not answered by a respondent. Unit non-response takes place when a randomly sampled individual cannot be contacted or refuses to participate in a survey. Among all such concerns, item non-response has caught researchers' special attention. Comparative studies are hugely affected by this problem: but what should we do when confronting a large amount of item non-response while still interested in drawing valid inferences from the available data since it is obviously not appropriate to ignore it and discard all 'don't know' answers. Proper understanding of the missing data mechanism can be a huge step in dealing with item non-response. It has been argued that, when respondents fail to answer a survey question, there are three possible meanings: *Don't know, Don't care, or Don't want to tell.* Don't know, as an easy expression of no idea, no opinion, and hard to choose, is mainly because of ignorance, ambivalence, or idea conflicts. Don't care, discloses to what extent a respondent makes efforts to formulate an answer to a survey question. In this case, respondents' interest in the question or in the survey as a whole may play a role in item non-response. Don't want to tell, is usually associated with political context and prevalent social norms. Respondents may fail to answer questions because of political fear or social desirability. However, item non-response in Ghanaian surveys is often speculated as a problem but rarely researched. Even without sampling problems, this should be a great concern for those who do poll survey in Ghana. It is suspected that the prevalence of item non-response problems may be due to the fact that ordinary Ghanaian people may lack cognitive abilities to form concrete opinions to certain survey questions due to their low education level. According to [1], item non-response is the failure to obtain information for a question within an interview or questionnaire. Even though it results in missing values to particular questions, it does not mean that item non-response fails to contain any information. Rubin in 1976 differentiated among three kinds of item non-response according to the underlying missing data mechanism: *missing completely at random* (MCAR), *missing at random* (MAR), and *missing not at random* (MNAR). To define the three kinds of item non-response, [2] distinguished between the observed data Y_{obs} and the missing data Y_{miss}. These constitute the complete data matrix $Y = (Y_{obs}, Y_{miss})$. We adapted this notation to the latent variable framework. Y is the complete data matrix that consists of the observed item responses Y_{obs} and the omitted responses Y_{miss} of the k items Y_1 to Y_k, indexed by i. The values of a latent variable ξ can also be considered to be missing data. The MCAR is the case where the distribution of the item non-response data is independent of the item response data. The MAR holds if the distribution of the missing mechanism is only dependent on the observed data but not dependent on the unobserved values of the missing data. The third type, called MNAR is the opposite of MAR. This means the conditional distribution of the missing data given the observed data depends on the unobserved data and possibly the latent variable(s). Item response theory (IRT) relates characteristics of items (item parameters) and characteristics of individuals (latent traits) to the probability of a positive response developed for dichotomous and polytomous data. In each case, the probability of answering correctly or endorsing a particular response category can be represented graphically by an item (option) response function (IRF/ORF). These functions represent the nonlinear regression of a response probability on a latent trait, such as conscientiousness or verbal ability [3].

This paper was motivated by the need to use *Item Response Theory (*IRT) to identify the most appropriate IRT model for understanding item non-response; identify the categories of survey questions that are most difficult to answer by respondents: and, find out the reason behind 'don't know' responses and missing data.

2. Methodology

2.1 Data and World Value Survey Definition

Data for this study was obtained from the World Values Survey. The World Values Survey (WVS is a global research project that explores people's values and beliefs, how they change over time and what social and political impact they have. It is carried out by a worldwide network of social scientists who, since 1981, have conducted representative national surveys in almost 100 countries. The WVS originated from European Values Study (EVS) and extended to countries outside Europe in 1981, which constituted the first wave of the WVS. The surveys aim to be longitudinal as well as cross-cultural. The 2nd wave was conducted in 1990, ten years after the 1st and embraces 42 countries. The interval between the waves was shortened to 5 years for the third in 1995, fourth in 2000, and fifth in 2005 waves, which includes 52 and 64 countries separately. In total, the WVS covers 81 societies. The data on Ghana from the 5th wave [4] was selected for this study. The WVS was conducted by the Institute of Social Research at the University of Michigan (ISR) in collaboration with leading survey research organizations in each country. The Ghanaian survey in this wave was conducted by these principal investigators Markinor Thinking, Tracy Hammond and Mari Harris. The survey covers a variety of research topics, such as socio-cultural, moral, religious, and political values and attitudes. It employs detailed questionnaires and face-to-face interview techniques in methodology. Representative samples were drawn from each country and the number varies from 1000 to 3500 per country. The survey period for Ghana was from 19th February to 04th April 2007 which included a sample of about 1,534 individuals.

2.1.1 Sample selection

Respondents due to limitation of cognitive ability may be inclined not to answer or give 'don't know' answers to difficult questions, or to sensitive questions due to social desirability or political fear. The WVS covers a variety of topics that can be used to test the effects of respondents' cognitive ability on item non-response. We grouped all the questions into six categories: Life Related Questions (LRQ), Value Related Questions (VRQ), Political Related Questions (PRQ), Income Related Questions (IRQ), Democracy Related Questions (DRQ), and questions on socio-demographic features. Life related questions include those on attitudes to life, confidence, marriage, religion, and morality whereas value related questions consist of those reflecting personal values on environment, country priority, and future changes. Politics related questions include questions on institutional trust, political system, and international politics. Income related questions consist of those relating to family savings, and scale of income whiles democracy related questions consist of those on governance and democracy. The socio-demographic features include age, sex, educational level, and employment status. Questions in all six categories can be sensitive depending on social and political contexts. We finally selected one question randomly from each of the categories except those on socio-demographic features (since all questions in that category were all answered) to be used for the IRT modeling and construction of item characteristics curves. Based on literature study, manual coding scheme was used to code all items. All items were either assigned a value zero or one. Zero was used when an item was not answered or when a 'don't know' answer was provided, and one was assigned to all answered items. More on the sampling design is found in (*http://www.worldvaluessurvey.org*).

2.2 Modeling Approach

A variety of measurement models (Rasch Models) proposed by a Danish Mathematician, George Rasch, in 1960 has been applied to this study. The dependent variable in this study, item non-response, is measured as whether or not an item is answered by an individual. This is a dichotomously scored variable: 0 = no answer or 'don't know' answer to the item, 1 = answer to the item. The item non-response variable is extremely skewed; most people answered most items. Such data violate several assumptions of the usual regression methods. Item response theory (IRT) was employed to document the psychometric properties of item non-response and to derive the latent item non-response trait. The advantages of IRT, compared to other psychometric approaches, are well documented [5]. IRT provides error estimates that are specific to the trait level. Importantly, the latent trait level estimates are not scale-dependent and item characteristics are not group-dependent. Hence, as in this study, IRT methods demonstrate whether an individual's score at a particular latent trait level indicates that the probability of responding to an item is the same. In addition to the above advantages, IRT informs about the relationship between responses (in this study, answering or not answering a question) and the individual's latent trait (Question knowledge). Provided that the model fits the data, the information obtained from IRT analyses thus enables documenting question knowledge across the gradient of latent trait scores taking into account difference among items in discriminating between trait levels. The various IRT models explored in this study are described below;

$$P_i(\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}}, \tag{1}$$

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}, \tag{2}$$

$$P_i(\theta) = c_i + (1 - c_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}, \tag{3}$$

The model in equation (1) was first proposed by [6]. In 1968, [7] extended the Model in equation (1) to obtain the model in (2). Finally, (Lord, 1980) extended the model in equation (2) to obtain the model in (3). In the above models from (1) to (3), θ is a continuous variable (latent Question Knowledge trait) and for $i = 1,2,3, \ldots, n$, $P_i(\theta)$ is the probability of an individual with ability θ responding to item i, a_i is the item discrimination parameter for the i^{th} item, b_i is the difficulty parameter for the i^{th} item, c_i is the item pseudo-chance parameter for the i^{th} item, e is a transcendental number (natural log constant) whose value to three decimal places is 2.718, D is a scaling constant used to approximate the logistic model to the normal ogive model, and n is the size of the respondents. The analysis was done using an *R* package for latent trait modeling and item response theory analyses (*R Development Core Team*, 2010). This procedure utilizes the marginal maximum likelihood method to calibrate items and the Bayesian expected a posteriori method to estimate latent trait scores. The parameters are estimated by maximizing the approximate marginal log-likelihood under the conditional independence assumption, that is, conditionally on the latent structure the items are independent Bernoulli variates under the logit link. The probability of responding to an item is related to the question knowledge scale as a monotonically increasing S-shaped item response function (IRF). The trait value at which 50% of the sample responds is referred to as the item threshold parameter. The item discrimination (a) parameter is the slope of the item response function at this trait value. Higher item discrimination values are associated with steeper IRFs. In other words, higher discrimination parameters indicate a stronger relationship between question knowledge and observed responses. The item threshold parameter determines the position of the curve along the latent trait. A higher threshold parameter indicates that fewer individuals respond to a particular question. In other words, a higher trait score (higher score on the continuum of the question knowledge scale) is required for the person to respond to the particular question. The guessing parameter c is the probability of responding to an item i even if the person does not know the answer. When $c = 0$, the three-parameter model is equivalent to the two-parameter model. The degree to which these IRT models adequately fit the empirical data was indicated using various goodness-of-fit indices Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Likelihood ratio test was used to select the most suitable model for the data. Three assumptions needed to be satisfied when applying IRT models. The first assumption referred to as unidimensionality assumption which implies that the probability of responding to a question is a function of only one latent trait. The second assumption referred to as local independence is that, no relationship is present in an individual's responses to different items after taking into account the individual's latent trait level. Unidimensionality is a sufficient condition for satisfying the local independence assumption. Finally, the response of a person to an item can be modeled by a mathematical *item response function* (IRF). Given that the data adequately fit, one can make simple comparisons of the items and respondents since comparison of two items' difficulty parameters are assumed to be independent of any group of subjects being surveyed, and the comparison of two subjects' trait levels does not depend on any subset of items being administered.

3. Results and Discussion

At an initial step, descriptive statistics for the data are produced. We observe from Table 1 that the life related questions seems to have least difficult questions having the highest proportion of about 99% of responses, while the income related questions seems to be the most difficult one having the lowest proportion 90% of responses. The proportion of responses for the politics related question, democracy related question, and value related question were about 93%, 94%, and 95% of the respondents respectively. Frequencies of all possible total scores are provided from the preliminary analysis. The total score of a response pattern is simply its sum. For dichotomous items, this is the number of positive responses. In the analysis, 1217 out of the 1534 respondents responded to all questions on all 5 categories, 215 respondents responded to all questions on 4 categories, 78 respondents responded to all questions on 3 categories, 22 respondents responded to all questions on only 2 categories, and 2 respondents responded to all questions on only 1 category.

Table 1. Proportions for each level of response

Proportions for each level of response	**0(%)**	**1 (%)**
Life Related Questions (LRQ)	1.17	98.83
Politics Related Questions (PRQ)	6.98	93.02
Democracy Related Questions (DRQ)	5.67	94.33
Value Related Questions (VRQ)	4.76	95.24
Income Related Questions (IRQ)	10.43	89.57

We have the χ^2 p-values for pairwise associations between the five items, corresponding to the 2×2 contingency tables for all possible pairs. Before an analysis with latent variable models, it is useful to inspect the data for evidence of positive correlations. In this case, the ad hoc checks are performed by constructing the 2×2 contingency tables for all possible pairs of items and examine the chi-squared p-values. Inspection of non-significant results can be used to reveal 'problematic' items [8]. We observe from Table 2 that three pairs of items seem to have weak degree of association, and the life related item is included in all three pairs. The small number of non significant pairwise association poses the data for IRT modeling.

Table 2. Pairwise Associations

	Item i	*Item j*	*P. value*
1	LRQ	VRQ	1.000*
2	LRQ	PRQ	0.638*
3	LRQ	DRQ	0.071*
4	LRQ	IRQ	0.002
5	DRQ	VRQ	1e-03
6	VRQ	IRQ	1e-04
7	PRQ	VRQ	1e-06
8	PRQ	DRQ	1e-12
9	PRQ	IRQ	2e-13
10	DRQ	IRQ	2e-16

**not significant at 5%*

3.1 The Constrained Rasch Model

We start by fitting the original form of the Rasch model that assumes known discrimination parameter fixed at value one. In this Model, a respondent is characterized by a level on a latent trait (Question knowledge), and an item is characterized by a degree of difficulty. The larger the value of the difficulty parameter implies the more difficult the question. Table 3 presents results for the constrained Rasch model parameter estimates. The results of the descriptive analysis above are also validated by the model fit in Table 3, where the income related questions and the life related questions are the most difficult and the least easy, respectively. A transformation of the parameter estimates into probability estimates results is computed. The probability of responding to an item is seen as a function of the ratio of a respondent's level on the trait (Question Knowledge) to the item difficulty. The column $P(X = 1|\mathrm{Z} = 0)$ denotes the probability of responding to the i^{th} item for the average individual. These probabilities were sorted according to the difficulty estimates as shown in Table 3. we observe that the probability of the average individual responding to the life related question is higher than responding to the other related question.

Table 3. Difficulty and Probability estimates under the constrained Rasch model

	Difficulty	**Discrimination**	***P(x=1|z=0)***
LRQ	-4.9433	1	0.9929
PRQ	-3.4492	1	0.9692
DRQ	-3.2511	1	0.9627
VRQ	-3.0124	1	0.9531
IRQ	-2.5282	1	0.9261

3.2 The One-Parameter Logistic Model (1-PLM), Unconstrained Rasch Model

Both the constrained and the unconstrained Rasch models have similar features and are mathematically equivalent except that, where the constrained Rasch model had a fixed slope of unity for all items, the unconstrained Rasch model only requires the slope to be equal for all items. The discrimination parameter estimated from this model is 1.602 which is different from one. Comparing the difficulty parameters under this model, similar results shows that the income related question and the life related question are most difficult and easiest respectively. The probability of an average individual under the unconstrained Rasch model responding to the life related question is higher than responding to the value related question. Similarly, the probability of responding to the democracy related question is higher than responding to the politics related question for the average individual under the unconstrained Rasch model.

3.3 The Two-Parameter Logistic Model (2-PLM)

Here, we explore how the two-parameter logistic fits the data. Whereas the Rasch models constrain the discrimination parameter to be equal, the two-parameter logistic model allows the slope or discrimination parameter to vary across items. Discrimination is deemed high if its value is greater than 1.35 [9]. Table 4 presents results for the *2-PLM* estimates which shows that the discrimination parameter estimates is not the same for all items. Comparing the difficulty parameters under this model, we observe in Table 4 that the income related question and the life related question are most difficult and easiest respectively. In terms of discrimination, we observe that, all the questions have high discrimination, especially the democracy related question. The probability of an average individual under the two-parameter logistic model responding to the life related question is higher than responding to all other related questions. Similarly, the probability of responding to the democracy related question is higher than responding to the politics related question for the average individual under the two-parameter logistic model.

Table 4. Difficulty, Discrimination and Probability estimates under the 2-PLM

	Difficulty	***Discrimination***	***P(x=1|z=0)***
LRQ	-4.7096	1.0534	0.9930
VRQ	-2.7544	1.3655	0.9773
PRQ	-2.2738	1.5143	0.9690
DRQ	-2.0177	2.2710	0.9899
IRQ	-1.8684	1.5945	0.9516

3.4 The Three-Parameter Logistic Model (3-PLM)

Here, we explore how the three-parameter logistic model fits the data. Whereas the Rasch models constrain the discrimination parameter to be equal, the three-parameter logistic model allows the slope or discrimination parameter to vary across items and also incorporates a guessing parameter. This model is usually employed to handle the phenomenon of non-random guessing in the case of difficult items. Comparing the difficulty parameters under this model in Table 5, we observe that the value related question and the life related question are most difficult and easiest respectively. In terms of discrimination, we observe from Table 5 that, all the questions have very high discrimination. It is important to mention that under the three-parameter model, the values of the guessing parameter are not apparent since difficulty values are less than zero and discrimination values are greater than one [9].

Unlike the *1-PLM* and the *2-PLM*, it is shown in Table 5 that the probability of an average individual under the *3-PLM* responding to the life related question is lower than responding to the value related question. Also, the probabilities of responding to the democracy related question, the politics related question, and the income related question for the average individual under the three-parameter logistic model is certain.

Table 5. Guessing, Difficulty, Discrimination and Probability estimates under the 3-PLM

	Guessing	***Difficulty***	***Discrimination***	***P(x=1|z=0)***
LRQ	0.0547	-4.1906	1.2071	0.9940
IRQ	0.2908	-1.1578	56.9407	1.0000
DRQ	0.7933	-0.6688	42.6518	1.0000
PRQ	0.7869	-0.6422	47.5954	1.0000
VRQ	0.8678	-0.3507	36.9218	0.9999

3.5 Model Selection

To determine which of the four IRT models fitted above is the most appropriate for the data, the goodness of fit indicators which compares the unconstrained version of the Rasch model, the constrained Rasch model, the two-parameter logistic model, and the three parameter logistic models. The estimated goodness of fit indicators in appendix A1 shows that the unconstrained Rasch model has the smallest AIC value 3104.63 and BIC value 3136.64, hence the more suitable for the data. Adopting the unconstrained Rasch model as the most appropriate for our data, we produce results for the estimated Item Characteristic, the Item Information and the Test Information Curves. The Item Characteristic Curve (ICC) is the basic building block in IRT. The ICC models the relationship between a person's probability of responding to an item category and the level on the construct measured by the scale [9]. The properties of the ICC needed to describe the item's characteristics are its location and the steepness. The steepness of the ICC reflects the discrimination property of an item whereas the difficulty parameter which is represented by location is the point on the ability scale at which the probability of responding to the item is 0.5. We observe from Fig. 1 that the life related question and the income related question are the easiest and the most difficult respectively.

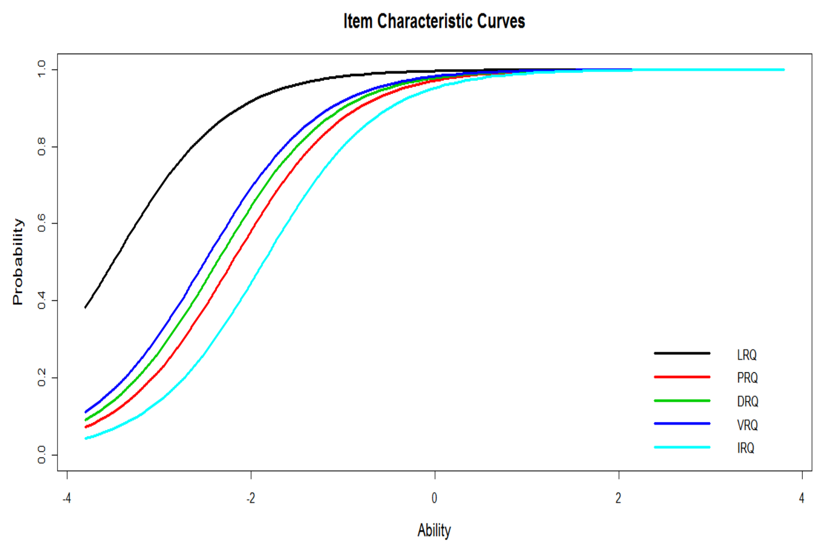

Fig. 1. Estimated Item Characteristic Curve obtained from the data

Item information is the amount of information based upon a single item. It can be computed at any ability level. Because only a single item is involved, the amount of information at any point on the ability scale is going to be rather small. An item measures ability with greatest precision at the ability level corresponding to the item's difficulty parameter [9]. We observe from Fig. 2 that the amount of item information for each item decreases as the ability level departs from the item difficulty and approaches zero at the extremes of the ability scale.

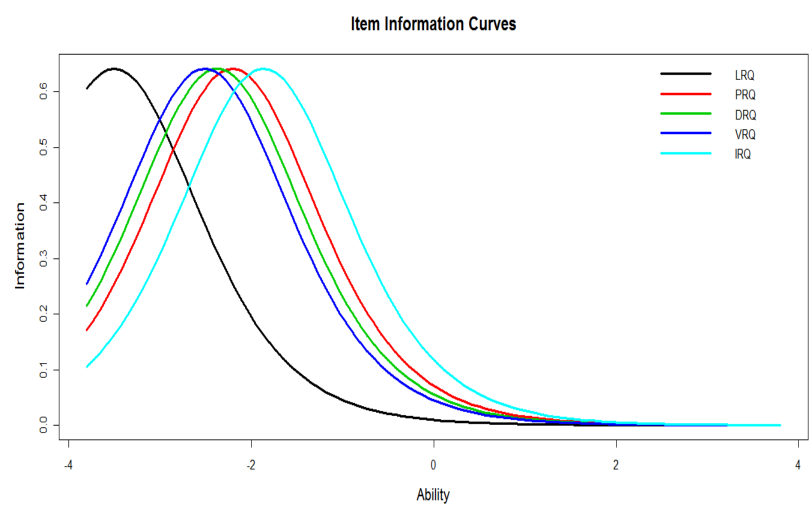

Fig. 2. Estimated Item Information Curve obtained from the data

Since a test is used to estimate the ability of an individual, the amount of information yielded by the test at any ability level can also be obtained. A test is a set of items; therefore, the test information at a given ability level is simply the sum of the item information at that level. The general level of the test information function will be much higher than that for a single item information function. Thus, a test measures ability more precisely than does a single item [9]. We observe from Fig. 3 that the maximum value of the test information function is at ability level -2. However, as the ability level increases, the amount of test information decreases significantly. This indicates that the items asked in our data mainly provide information for respondents with low ability. In particular, the amount of test information for ability levels in the interval (−4, 0) is almost 90%.

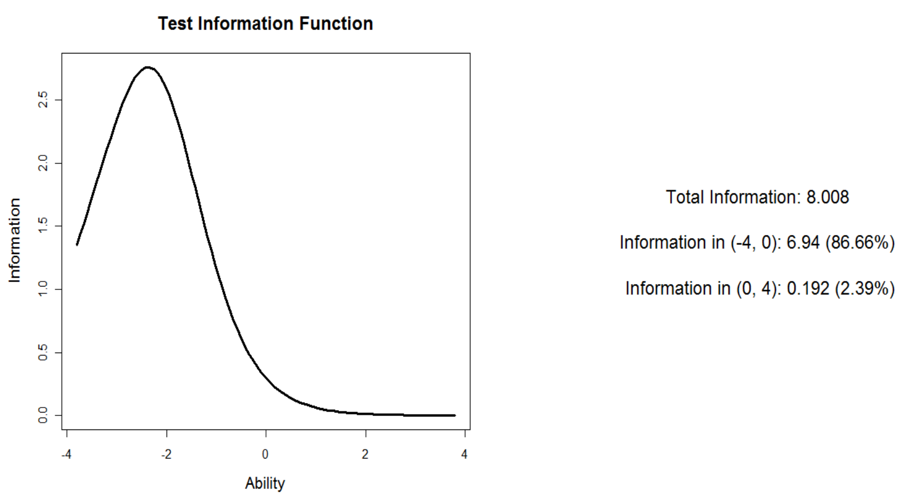

Fig. 3. Estimated Test Information Curve

3.6 Ability Estimates

Finally, the ability estimates for respondents are obtained. The primary purpose for using IRT in this study is to locate respondents on the ability scale. Since this will help us evaluate respondents in terms of how much underlying ability (Question knowledge) they possess. Factor scores or ability estimates are summary measures of the posterior distribution $P(Z/X)$, where Z denotes the vector of latent variables and X the vector of manifest variables. By default factor scores produces ability estimates for the observed response patterns. the items asked in the data mainly provide information for respondents with low ability (i.e., below 0). That is, most of the items in the dataset are relatively easy for the average respondent to answer. Fig. 4 is a Plot of a Kernel Density Estimation of the distribution of the factor scores (i.e., person parameters). Kernel density estimation is a non-parametric way of estimating the probability density function of a random variable. Kernel density estimation is a fundamental data smoothing problem where inferences about the population are made, based on a finite data sample [10]. It includes in the plot the item difficulty parameters (similar to the Item Person Maps). The plot confirms the fact that the data is extremely skewed.

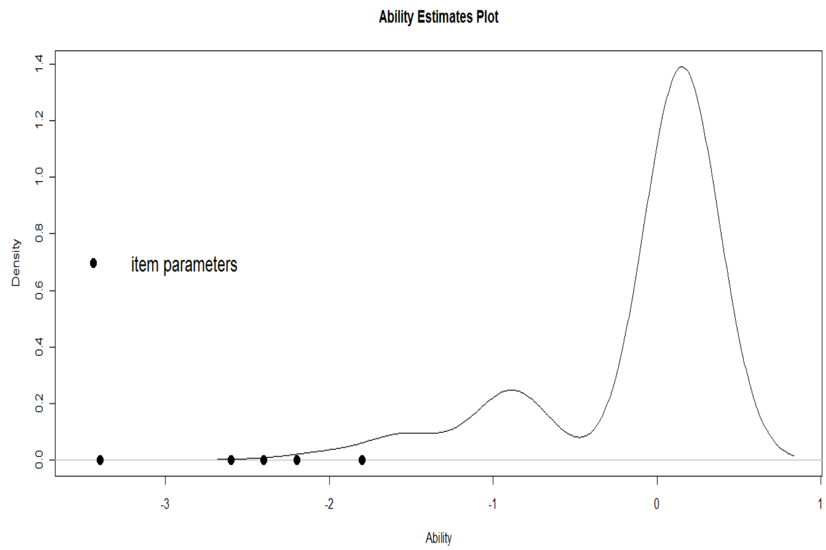

Fig. 4. Ability Estimates Plot

4. Conclusion

The implications from this paper are quite clear. First, we investigated to identify the most appropriate IRT model for understanding item non-response by exploring the four IRT models for dichotomous data which include the constrained Rasch model, the unconstrained Rasch model, the two-parameter logistic model, and the three-parameter logistic model. From the likelihood ratio test, we observed by looking at the AIC and BIC values that, the unconstrained Rasch model had the smallest AIC and BIC values. Hence, the most appropriate model for the data. Furthermore, we investigated to identify the categories of survey questions that are most difficult to answer by respondents. As indicated from the results of the unconstrained model, the income related question recorded the highest difficulty parameter. In terms of probability estimates, we observe that, the probability of responding to the income related question by the average individual as compared to the other categories of questions is the smallest. Therefore, the income related questions are the most difficult category of survey questions to answer by respondents. Finally, we analysed the reason behind don’t know responses and missing data; whether respondents don’t really know, don’t care, or don’t want to answer. From the selected model, the difficulty of a question explains whether or not an individual will respond to that question. We also observe from the ability estimates and the test information curve that, almost 90% of the total test information for ability levels lies in the interval (-4, 0). This means that most of the questions in the dataset were easy questions. Also, because the difficulty values are less than 0 and discrimination values are greater than 1, ‘don’t care’ which is usually associated with guessing is not apparent in the dataset. Therefore, if an individual does not answer a survey question or give a ‘don’t know’ answer, it is not only because of the question’s difficulty but also because the individual doesn’t want to answer.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] De Leeuw ED. Reducing Missing Data in Surveys: An Overview of Methods. Quality and Quantity. 2001;5(2):147-160.

[2] Rubin DB. Inference and missing data. Biometrika. 1976:63(3), 581–592.

[3] Hulin CL, Drasgow F, Parsons CK. Item Response Theory, application to Psychological measurements. Dow Jones-Irwin;1983

[4] World Values Survey Data. Accessed 25th October 2010) [online]. Available: http://www.worldvaluessurvey.org

[5] Embretson SE, Reise SP. Item Response Theory for Psychologists. Mahwah, New Jersey: Lawrence Erlbaum Associates; 2000.

[6] Rasch G. Probabilistic models for some intelligence and attainment tests. Chicago: MESA; 1960.

[7] Birnbaum A. Some latent trait models. In F. M. Lord and M. R. Novick ed. Reading: MA: Addison-Wesley; 1968.

[8] Rizopoulos D. An R Package for Latent Variable Modeling and Item Response Theory Analyses. Catholic University of Leuven: Journal of Statistical Software. 2006;17(5).

[9] Baker FB. The Basics of Item Response Theory. USA: ERIC Clearinghouse on Assessment and Evaluation; 2001.

[10] Scott D. On optimal and data-based histograms. Biometrika. 1979;66: 605–610.

APPENDIX

A1: Likelihood ratio test for the Constrained Rasch, 1-PLM, 2-PLM and the 3-PLM

Likelihood Ratio Table						
	AIC	**BIC**	**log.Lik**	**LRT**	**df**	**p. value**
Unconstrained Rasch	3104.63	3136.64	-1546.31			
Constrained Rasch	3136.83	3163.51	-1563.41	34.2	1	<0.001
2-PLM	3106.11	3159.46	-1543.05	6.52	4	0.163
3-PLM	3109.89	3189.92	-1539.95	12.74	9	0.175

Source: WVS (Ghana, 5th wave)

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
www.sciencedomain.org/review-history.php?iid=466&id=6&aid=4098