Impact Of Missing Data On Rasch Model Estimations

Sümeyra SOYSAL
Hacettepe University, Turkey
sumeyrasoysal@hotmail.com

Çiğdem Akın ARIKAN
Hacettepe University, Turkey
akincgdm@gmail.com

Hatice INAL
Hacettepe University, Turkey
haticeinall@hotmail.com

ABSTRACT
This study aims to investigate the effect of methods to deal with missing data on item difficulty estimations under different test length conditions and sampling sizes. In this line, a data set including 10, 20 and 40 items with 100 and 5000 sampling size was prepared. Deletion process was applied at the rates of 5%, 10% and 20% under conditions of completely missing at random (MCAR) and missing at random (MAR) data structures in these full data sets. Pursuant to deletion process, values were assigned through regression and mean imputation methods being among missing data methods. The method of leaving the missing data blank (non-imputation) was also examined. In Rasch model, CML, JML and Pairwise estimations of item difficulty parameter were evaluated in comparison with parameters estimated from full data sets. To this end, RMSE was used as an evaluation criterion. At the end of the research, the least amount of estimation errors was obtained in JML method of Rash model than CML and Pairwise methods and it was found that Pairwise method had similar performance as CML method. It was found that errors in the estimations obtained through three different estimation methods among the methods to deal with missing data increased as missing data rate increased; and decreased as sample size increased. In a big sampling, non-imputation of missing data and regression-based estimations offered good results under many conditions as missing data rate increased. It was found that test length affected CML and Pairwise estimations among missing data imputation methods under many conditions

INTRODUCTION
Missing data is a very common problem for researchers in many studies in the field of education and psychology. De Ayala, Plake and Impara (2001) reported that there are three reasons for missing data in their research. The first reason is that individuals respond to only one subgroup of test in test applications such as individualized tests, multi-staged tests or common item design for nonequivalent groups and thus missing data emerges in sub-test sections which are not submitted to individuals (Eggen and Verhelst, 2011; Misley and Wu, 1988; Shin,2009). However, this situation does not cause any problem in ability estimation of individuals. As a second reason, missing data may emerge in lack of sufficient time to complete test application of an individual. In this case, non-responded items do not give information about abilities of individuals; therefore these items can be excluded for ability estimations. As a third reason, individuals do not want to respond to some test items even though they have time. Misley and Wu (1988), on the other hand, report that the presence of missing data may be dependent on ability level besides characteristics and demographic features of individuals.

Missing data is a remarkable problem since many model and theories such as item response theories (IRT) are based on the expectation that data are complete. However, amount and distribution of missing data in data matrix obtained especially at the end of research are more important for statistical estimations. Missing data were collected in three classes being completely at random, at random and not at-random by Little and Rubin (2002). Data missing completely at random (MCAR) are independent from observed and non-observed data and are random sampling of the observed data. Data missing at random (MAR) are dependent on observed variable and not dependent on non-observed variable. MCAR and MAR missing structures do not lead to systematic error in statistical estimations. Missing data, which are not MCAR and MAR, are data missing not at-random (MNAR) and dependent on variable which is observed to be missing or not. According to Mislevy and Wu (1988), a suitable algorithm must be used considering missing data estimation in parameter estimation of IRT-based data in presence of missing data.

The effect of missing data on the estimation of IRT -based individual and ability parameters has been discussed for years in psychometrics. In this field, studies were conducted on the effect of missing data on parameter estimations under the conditions of different missing data structures and rate; the effect of missing data on parameter estimations; comparison of methods to deal with missing data or methods of parameter estimations on missing data.

In their study, De Ayala, Plake and Impara (2001) investigated ability parameter estimations of dealing with missing data for 3-parameter logistic (3 PL) model. Accordingly, they reported that considering missing data as incorrect response leads more biased results than considering them as blank. In the study conducted by Finch (2008) on the effect of missing data imputation methods on the estimation of item difficulty and discrimination parameters based on 3PL model, it was found that efficiency of missing data imputation methods differs by missing data structure. In their study, Andreis and Ferrari (2012) investigated item parameter estimations in different missing rates and missing data structures in terms of methods to deal with missing data according to multi-dimensional two-parameter logistic (M2PL) model. In the present study, it was found that estimation bias increased as the rate of missing data increased; miss forest, forward imputation and multivariate imputation by chained equations gave similar results under almost every condition. Zhang and Walker (2008) investigated the effect of methods to deal with missing data in data sets having different missing data rates, sample size and item number on individual-model fit and individual ability estimations. At the end of this research, it was found that individual-model fit decreased and ability estimation bias increased as missing data rate increased; considering missing data as incorrect response gave the worst result while pairwise deletion method gave the best.

DeMars (2002) investigated the effect of missing data in MCAR and MAR structures on difficulty parameter through joint maximum likelihood (JML) and marginal maximum likelihood (MML) estimation methods. De Mars (2003) researched the correlation between individual abilities rate of not-reached items in terms of JML and MML estimations of 1PL model. In this study, it was found that two estimation methods estimated item parameters unbiasedly when rate of not-reached items was not dependent on abilities of individuals, JML estimated unbiasedly and MML estimated lower than real situations when not-reached rate was correlated with abilities of individuals.

Heine and Tarnai (2015) investigated estimations based on MML, CML and Pairwise methods at different rates of data missing completely at random in item difficulty parameter in the context of Rasch model. According to this research, it was found that three methods gave similar results when there was no missing data and pairwise-based parameter estimations even at %35 rate of missing data were as stable as those obtained from MML method. Custer, Sharairi and Swift (2012) discussed the effect of missing data on parameter estimations on the basis of JML method while Mislevy et al. (2005) investigated it on the basis of MML method.

Studies (Hohensinn & Kubinger, 2011) reported that missing data lead to bias in statistical results. The effect of missing data methods to deal with them on statistical estimations could be more clearly examined through IRT models that can estimate with missing data matrix. At this point, to research this case through Rasch model estimation methods was considered important. Apart from maximum likelihood estimation methods which are frequently used in the literature, it was aimed to attract attention to less known and less used pairwise-based estimations. In this content, this study aims to investigate the effect of methods to deal with missing data on item difficulty estimations under different test length conditions and sampling sizes. To this end, following three questions were asked.

1) What is the effect of data structures being missing at random or completely at random and missing data rate on CML estimations of item difficulty parameter?
2) What is the effect of data structures being missing at random or completely at random and missing data rate on JML estimations of item difficulty parameter?
3) What is the effect of data structures being missing at random or completely at random and missing data rate on Pairwise estimations of item difficulty parameter?

Rasch Model and Estimation Methods

Rasch Model: In item response theory, models are named by the number of parameters included. Rasch model known as 1 PL model in the literature is mathematically same with 1 PLM. However, discrimination parameter in Rasch model equals to 1 when it is equal for all items in 1 PL model (de Ayala, 2009). Rasch model is a special form of 3-parameter logistic model in which opportunity parameter is minimum and discrimination parameters are the same for all items (Hambleton and Swaminathan, 1985). Mathematical formula of Rasch model is presented as follows.

$$P_i(\theta) = \frac{\exp(\theta - b_i)}{1 + \exp(\theta - b_i)}$$

$P_i(\theta)$: Likelihood of an individual with θ to give correct response to item i.

bi : difficulty of item i

Parameter b in Rasch model is item difficulty parameter and generally gets a value between -2 and +2. In this model, it is assumed that performance of an individual is solely affected by item difficulty.
In this research, CML, JML and pairwise estimation methods were used for the estimation of item difficulty parameter based on Rasch model.

Joint maximum likelihood (JML): JML 1, 2 and 3 parameters used in estimations of both ability and item parameters can also be used in models (Hambleton, Swaminathan and Rogers, 1991). Unknown ability levels in JML are considered as known values by using estimated ability levels. Afterwards, unknown ability levels are used to estimate item parameter and arrangements are made in the estimations of ability levels through estimated item parameters. In other words, JML estimation is iterative and includes estimation of item and ability parameters. In the first stage, individual parameters are estimated and in the second stage, item parameters are estimated. The first iteration of these two stages includes initial values for item parameters; in this way, ML estimations of individual parameters are obtained. Afterwards, item parameters are estimated by using first individual parameters and this process maintains until the change of item parameter is very little among iterations (Hambleton, Swaminathan and Rogers, 1991; Embretson and Reise, 2000; de Ayala, 2009). JML has some advantages and disadvantages.

Constrained Maximum Likelihood (CML): CML can be applied only to Rasch model and other models that are extension of Rasch model (Hambleton, Swaminathan and Rogers, 1991; Embretson and Reise, 2000). In CML estimation; unknown ability levels examine response pattern likelihood without the parameters of ability level. Since total score of individual is a sufficient statistical information for the estimation of ability level, there is no need for more information for parameter estimations (Embretson and Reise, 2000). Therefore, 2 PL and 3 PL are not used in the model.

Pairwise Method: It is method that ensures the calculation of parameters based on constrained item category frequencies obtained from pairwise comparisons of items. Chopin (1968, 1985) showed the practical application of this method which was considered as an alternative to the calculation of item difficulty parameter by Rasch (1966). Chopin (1985) developed two methods being iteractive based on maximum likelihood for the estimations based on these pairwise comparisons (e.g. Andrich and Luo, 2003) and non-iteractive based on the estimation of the least squares of item parameters as well. In the present study, pairwise method based on the estimation of the least squares of item parameters was used (e.g. Heine, 2015). Pairwise method is based on the comparison of responses given to two items taking stand from the likelihood of an individual for giving correct or incorrect response to an item. An individual can give four different responses to two items; correct response to both items, correct response to at least one item and incorrect response to both items. These likely cases are presented with equations below by Rasch (1966) and Chopin (1985).

$$P(x_{vi}=0,\ x_{vj}=0)=\frac{1}{1+e^{(\Theta_v-\sigma_i)}}\times\frac{1}{1+e^{(\Theta_v-\sigma_j)}}$$

$$P(x_{vi}=1,\ x_{vj}=0)=\frac{e^{(\Theta_v-\sigma_i)}}{1+e^{(\Theta_v-\sigma_i)}}\times\frac{1}{1+e^{(\Theta_v-\sigma_j)}}$$

$$P(x_{vi}=0,\ x_{vj}=1)=\frac{1}{1+e^{(\Theta_v-\sigma_i)}}\times\frac{e^{(\Theta_v-\sigma_j)}}{1+e^{(\Theta_v-\sigma_j)}}$$

$$P(x_{vi}=1,\ x_{vj}=1)=\frac{e^{(\Theta_v-\sigma_i)}}{1+e^{(\Theta_v-\sigma_i)}}\times\frac{e^{(\Theta_v-\sigma_j)}}{1+e^{(\Theta_v-\sigma_j)}}$$

In these equations Θ_v, v presents ability level of an individual and σ_i and σ_j present item difficulty of items i and j. Taking stand from these four likelihood, constrained frequency matrix of items and the least square estimations of item parameters are calculated. Individual parameters are estimated through weighted likelihood approach under the assumption of stable item parameters. Choppin's (1985) studies can be examined for all calculations in Pairwise method and Heine's (2015) studies can be examined for a sample application matrix.

METHOD

Research Type

This study is a fundamental research since it investigates the effect of different ability estimation methods on estimation errors of item difficulty parameter in the presence of missing data.

Data Production and Analysis

R 3.5 program was used to produce full data sets within the context of this research. Full data sets were produced in accordance with Rasch model according to three test lengths being 10, 20 and 40 and two sampling sizes being 100 and 1000. Item difficulty parameter (β) and ability parameter (θ) N~(0.1) had normal distribution. Data sets with missing data were obtained by deleting data at the rates of 5%, 10% and 20% in missing completely at random (MCAR) and missing at random (MAR) structures through codes written in R program from full data set. Afterwards, mean and regression imputation methods were used for the missing data and these data sets were turned into full data sets. HotDeckImputation (Joenssen, 2015) package was used for mean imputation and mice (van Buuren and Groothuis-Oudshoorn, 2015) package was used for regression imputation being among missing data imputation methods.

In this research, 196 conditions; being 2 (100x1000) sample sizes, 2 missing data types (MCARxMAR), 3 missing data rates (5x%10x%20%), 3 test lengths (10x20x40), 3 estimation methods (JMLxCMLxPairwise) and 2 (mean and regression) methods to deal with missing data were investigated. 30 replication was applied for each condition.

Current R packages were used for the estimation methods used in this research. eRM (Mair et al., 2015) was used for CML estimation, sirt (Robitzsch, 2016) was used for JML and pairwise (Heine, 2015) package was used for pairwise method.

Research results were evaluated by comparing parameters estimated from full data sets which were obtained according to full data sets and missing data imputation methods based on these sets. To this end, RMSE was used as an evaluation criterion for the validity of the estimations of item difficulty parameter.

FINDINGS

In this chapter, the difference between estimations of item difficulty parameter obtained from data sets which were completed by the methods to deal with missing data and estimations obtained from full data sets under the conditions detected within the context of the research. The findings are presented based on the research questions.

What is the effect of data structures being missing at random and missing completely at random and missing data rates on CML estimations of item difficulty parameter under different sampling size and test length conditions?

According to sample size, test length and and missing data structures, RMSE values of conditional maximum likelihood (CML) estimations of item difficulty parameter are presented in Figure 1 and Figure 2.

According to Figure 1, the least erroneous estimations were obtained from leaving the missing data blank for all conditions of missing data rate and item length for N=100 under MCAR missing data structure. RMSE values of CML estimations obtained through regression imputation method were lower than mean imputation when test length was 10 and 20 while the results were reverse when test length was 40 items. It was found that estimation errors increased for each method to deal with missing data as missing data rate increased under all conditions. This finding is remarkable especially for mean and regression imputation methods. It was found that estimation errors obtained through the method of leaving missing data blank and mean imputation method did not vary much according to the variable of test length while estimation errors generally increased as item numbers increased in regression imputation method. Similar results were obtained for MAR missing data structure.

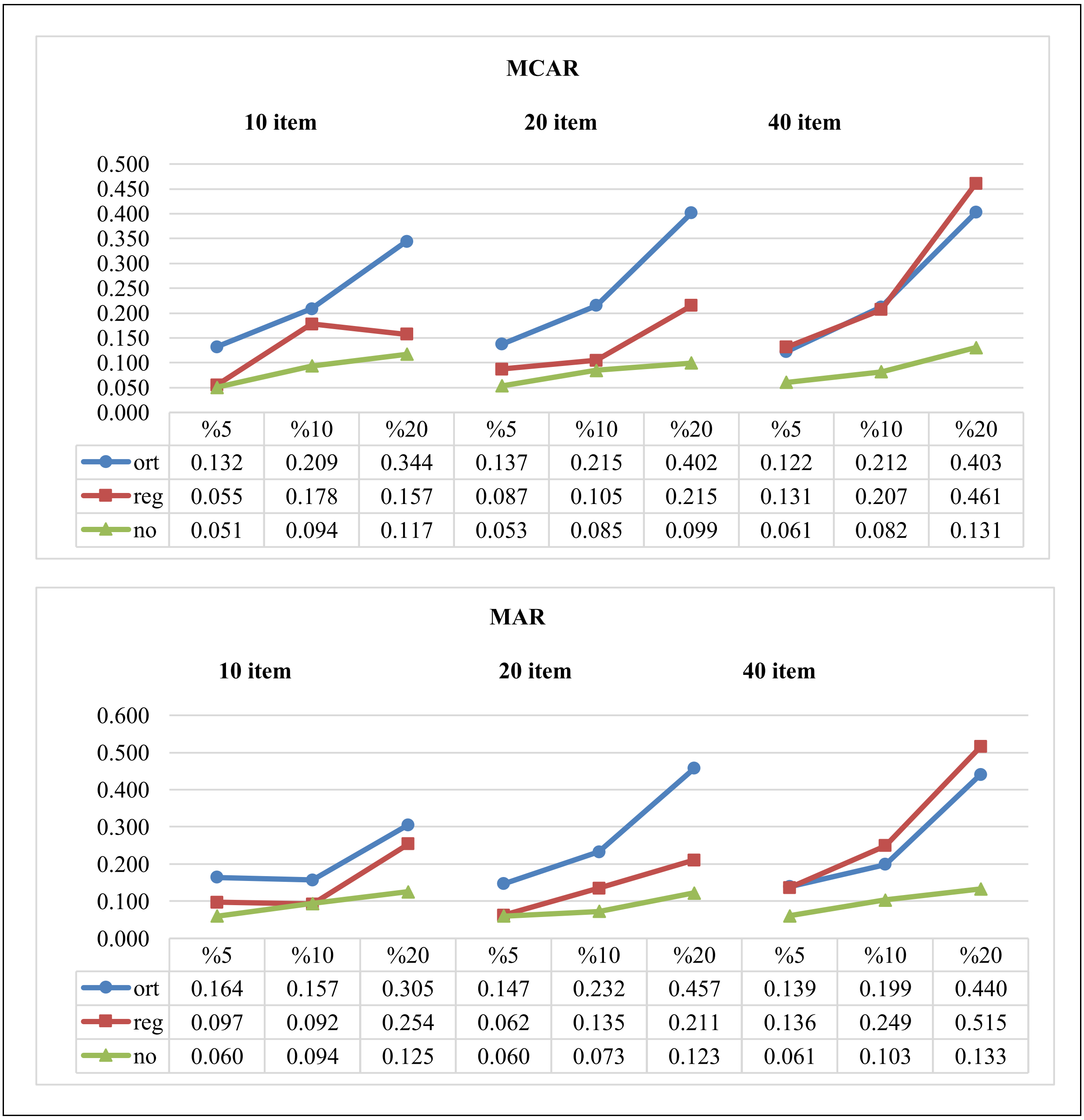
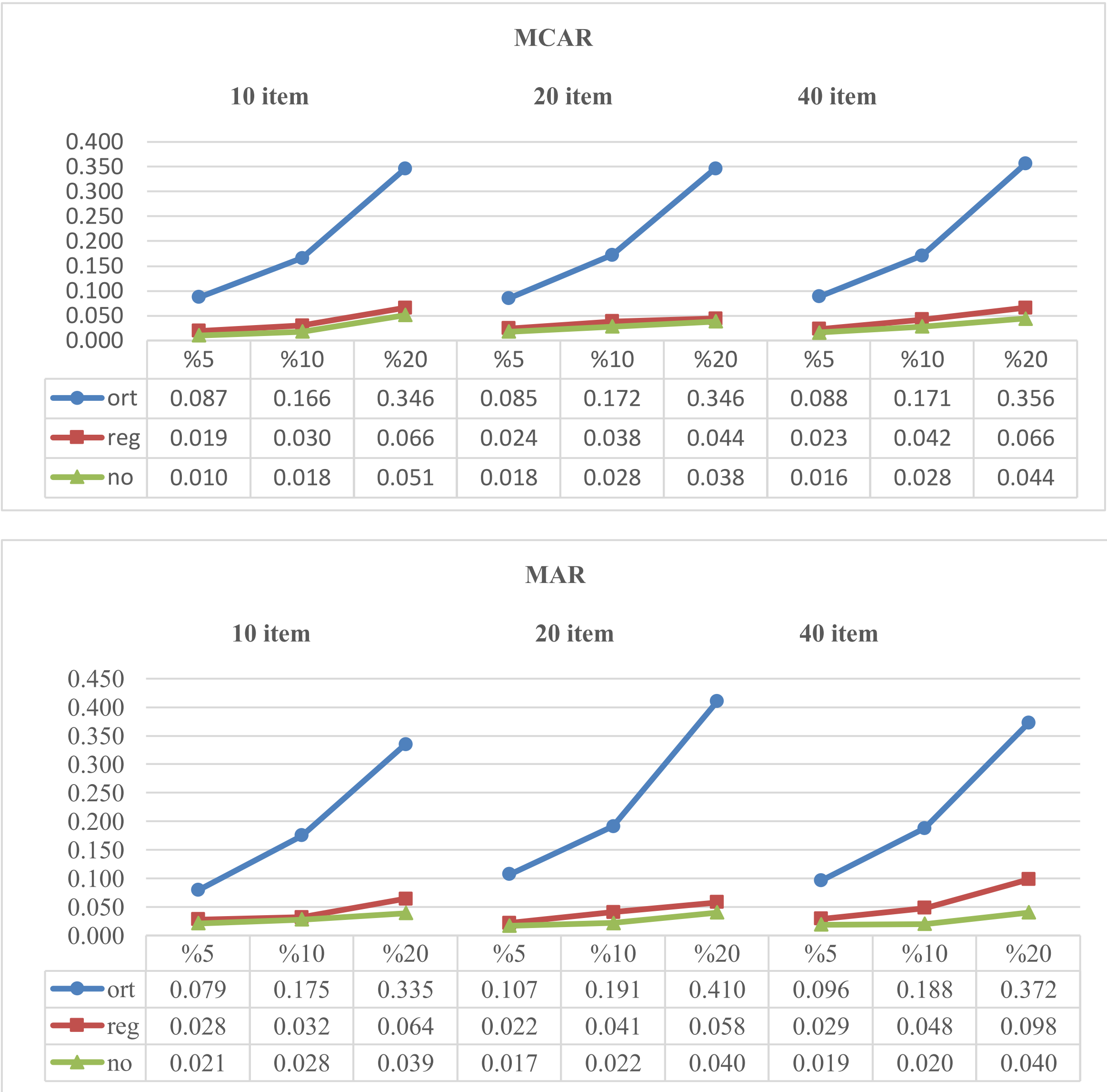


Figure 1. RMSE Values of CML Estimations of Item Difficulty Parameter (N=100)

According to Figure 2, RMSE values obtained through regression imputation methods and the method of leaving the missing data blank under all conditions of test length and missing data rates for N=1000 within the context of

MCAR missing data structure are very close to each other and the least erroneous estimations are obtained from these two methods. Estimation errors increased as missing data rate increased for each method to deal with missing data while this finding changed at minimum level by test length. Similar results were obtained in MAR missing data structure for N=1000 as well.



MCAR

	%5	%10	%20	%5	%10	%20	%5	%10	%20
ort	0.087	0.166	0.346	0.085	0.172	0.346	0.088	0.171	0.356
reg	0.019	0.030	0.066	0.024	0.038	0.044	0.023	0.042	0.066
no	0.010	0.018	0.051	0.018	0.028	0.038	0.016	0.028	0.044

MAR

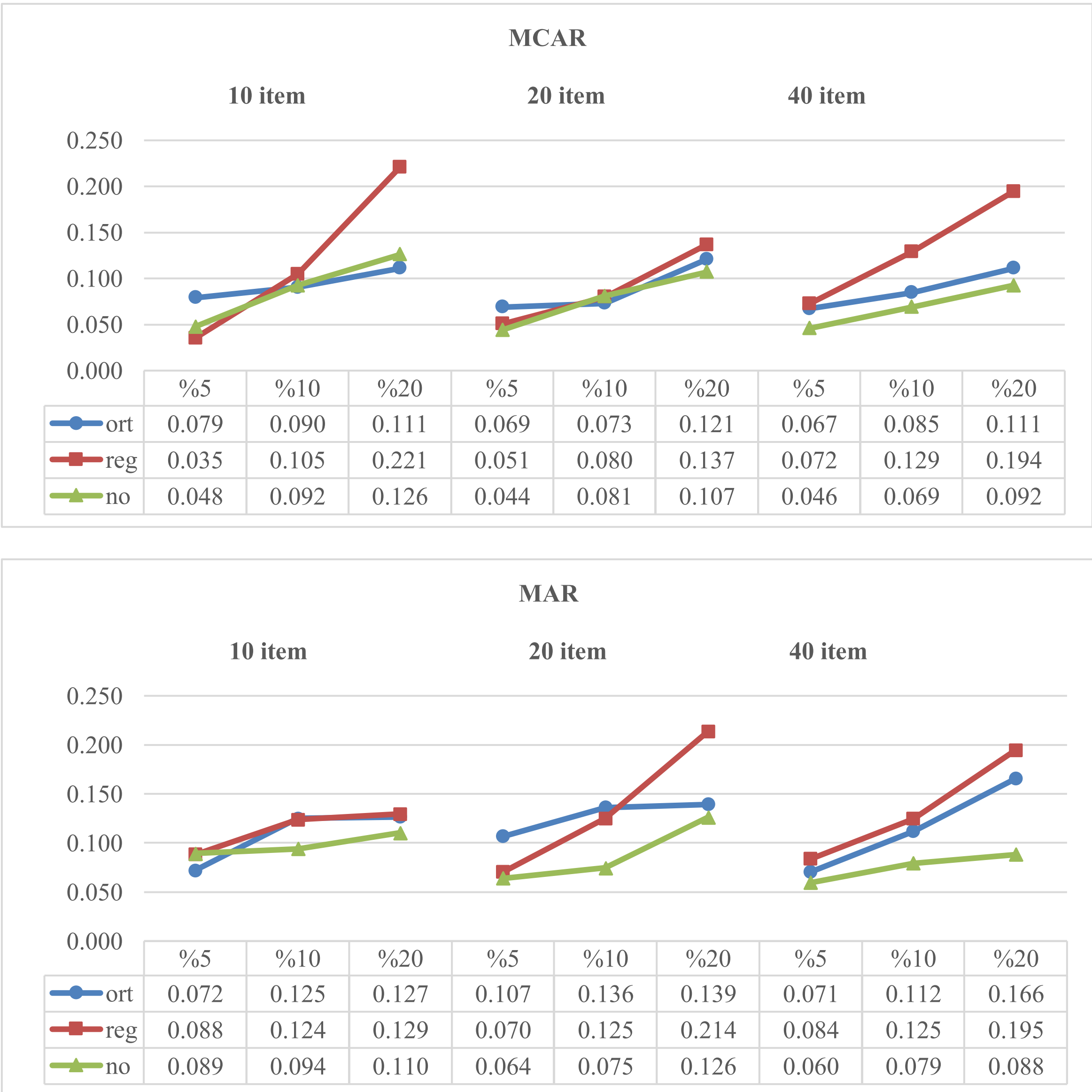
	%5	%10	%20	%5	%10	%20	%5	%10	%20
ort	0.079	0.175	0.335	0.107	0.191	0.410	0.096	0.188	0.372
reg	0.028	0.032	0.064	0.022	0.041	0.058	0.029	0.048	0.098
no	0.021	0.028	0.039	0.017	0.022	0.040	0.019	0.020	0.040

Figure 2. RMSE Values of CML Estimations of Item Difficulty Parameter (N=1000)

What is the effect of data structures being missing at random and missing completely at random and missing data rates on JML estimations of item difficulty parameter under different sampling size and test length conditions?

According to sample size, test length and and missing data structures, RMSE values of joint maximum likelihood (JML) estimations of item difficulty parameter are presented in Figure 3 and Figure 4.

According to Figure 3, the least erroneous JML estimations were obtained from leaving the missing data blank for all conditions of missing data rate and item length for N=100 under MCAR missing data structure. RMSE values obtained for mean imputation method and the method of leaving missing data blank were close to each other under all conditions. The most erroneous estimations were obtained from regression imputation method under many conditions. It can be seen that estimations obtained from three methods are similar to each other when the item number is 20. It can be seen that error increases as missing data rate increases for each method according to test length, yet regression imputation method is mostly affected by the change in item number. It can also be seen that similar results are obtained for MAR missing data structure for N=100 and the least erroneous estimations are obtained from the method of leaving the missing data blank. In contrast to results obtained from MCAR structure, three methods had similar results under MAR structure when the item number was 10.



MCAR	%5	%10	%20	%5	%10	%20	%5	%10	%20
ort	0.079	0.090	0.111	0.069	0.073	0.121	0.067	0.085	0.111
reg	0.035	0.105	0.221	0.051	0.080	0.137	0.072	0.129	0.194
no	0.048	0.092	0.126	0.044	0.081	0.107	0.046	0.069	0.092

MAR	%5	%10	%20	%5	%10	%20	%5	%10	%20
ort	0.072	0.125	0.127	0.107	0.136	0.139	0.071	0.112	0.166
reg	0.088	0.124	0.129	0.070	0.125	0.214	0.084	0.125	0.195
no	0.089	0.094	0.110	0.064	0.075	0.126	0.060	0.079	0.088

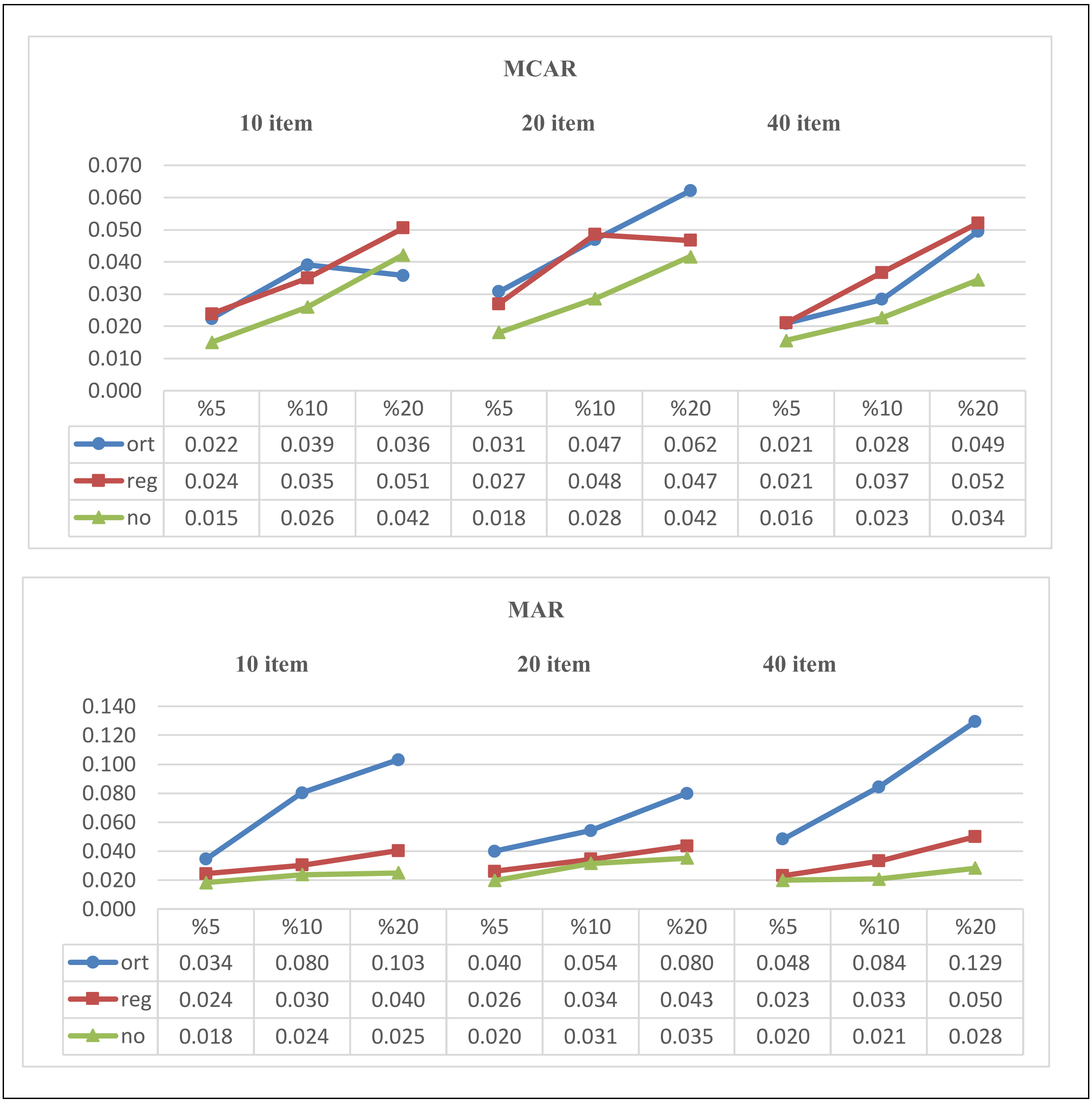
Figure 3. RMSE values of JML Estimations of Item Difficulty Parameter (N=100)

According to Figure 4, estimation errors obtained from three methods are similar to each other for all conditions of item number and missing data rate for N=1000 under MCAR missing data structure while the lowest RMSE values were obtained from the method of leaving the missing data blank. It can be seen that error obtained for every three method did not change much by item number and missing data rate. Although the results obtained for MAR missing data structure for N=1000 were similar, higher RMSE values were calculated by mean imputation method compared to two other methods.

What is the effect of data structures being missing at random and missing completely at random and missing data rates on Pairwise estimations of item difficulty parameter under different sampling size and test length conditions?

According to sample size and missing data structures, RMSE values of pairwise likelihood estimations of item difficulty parameter are presented in Figure 5 and Figure 6.

According to Figure 5, the least erroneous pairwise estimations were obtained from leaving the missing data blank for all conditions of missing data rate and item length for N=100 under MCAR missing data structure. The most erroneous estimation was obtained from mean imputation method under most of the conditions while it was obtained from regression imputation method when item number was 40. Estimation errors obtained from each method increase as missing data rate increases. It can be seen that estimation errors obtained from regression methods increase as item number increases and the least affected method from this case is the method of considering the missing data blank.



MCAR

	%5	%10	%20	%5	%10	%20	%5	%10	%20
ort	0.022	0.039	0.036	0.031	0.047	0.062	0.021	0.028	0.049
reg	0.024	0.035	0.051	0.027	0.048	0.047	0.021	0.037	0.052
no	0.015	0.026	0.042	0.018	0.028	0.042	0.016	0.023	0.034

MAR

	%5	%10	%20	%5	%10	%20	%5	%10	%20
ort	0.034	0.080	0.103	0.040	0.054	0.080	0.048	0.084	0.129
reg	0.024	0.030	0.040	0.026	0.034	0.043	0.023	0.033	0.050
no	0.018	0.024	0.025	0.020	0.031	0.035	0.020	0.021	0.028

Figure 4. RMSE Values of JML Estimations of Item Difficulty Parameter (N=1000)

It can also be seen the least erroneous estimations are obtained from the method of considering the missing data blank under MAR missing data structure for N=100 as well. Except for the condition having 40 items, the most erroneous estimations were obtained from mean imputation method. It can be seen that estimation errors obtained from regression imputation methods increase as item number increases.

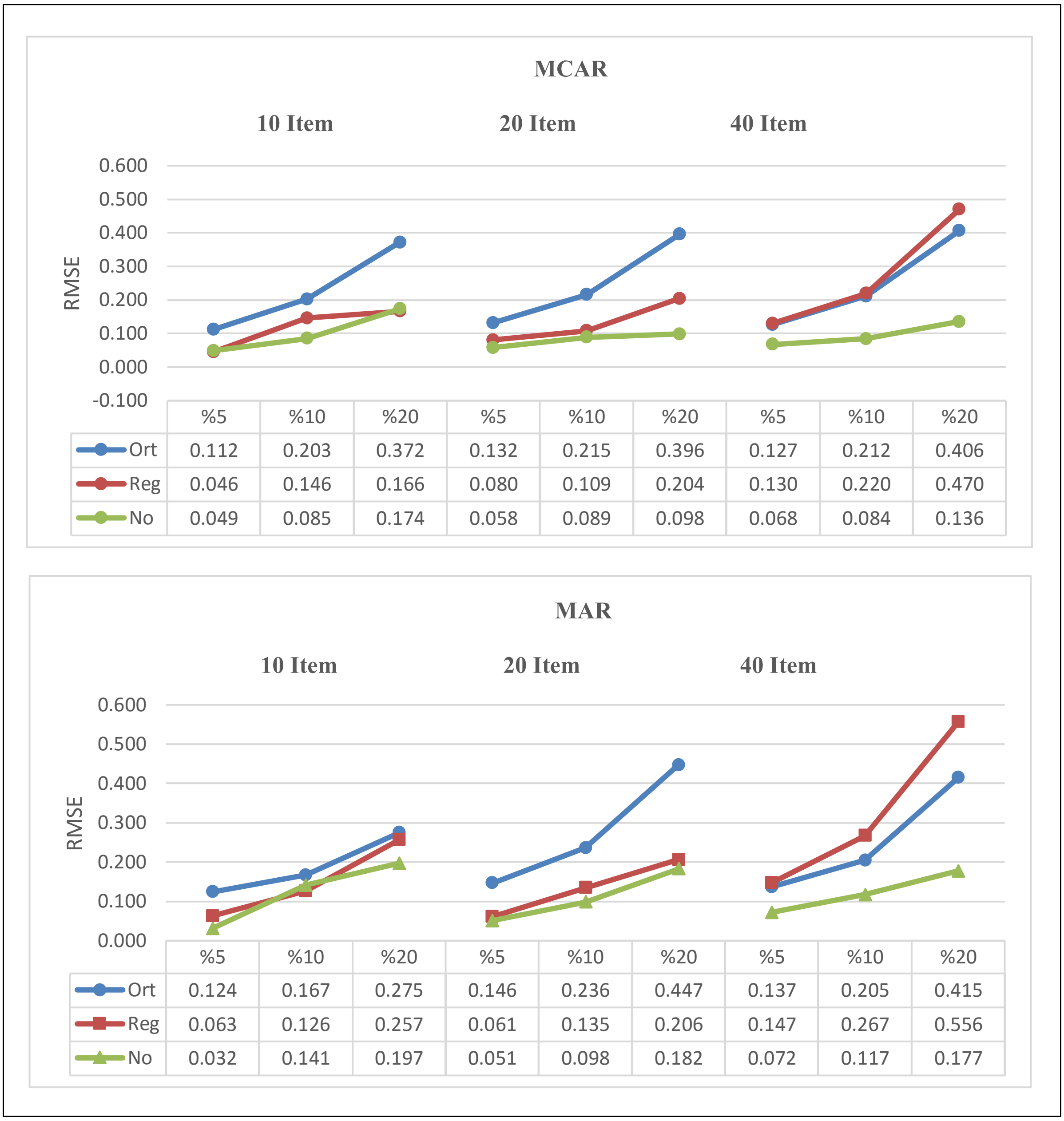
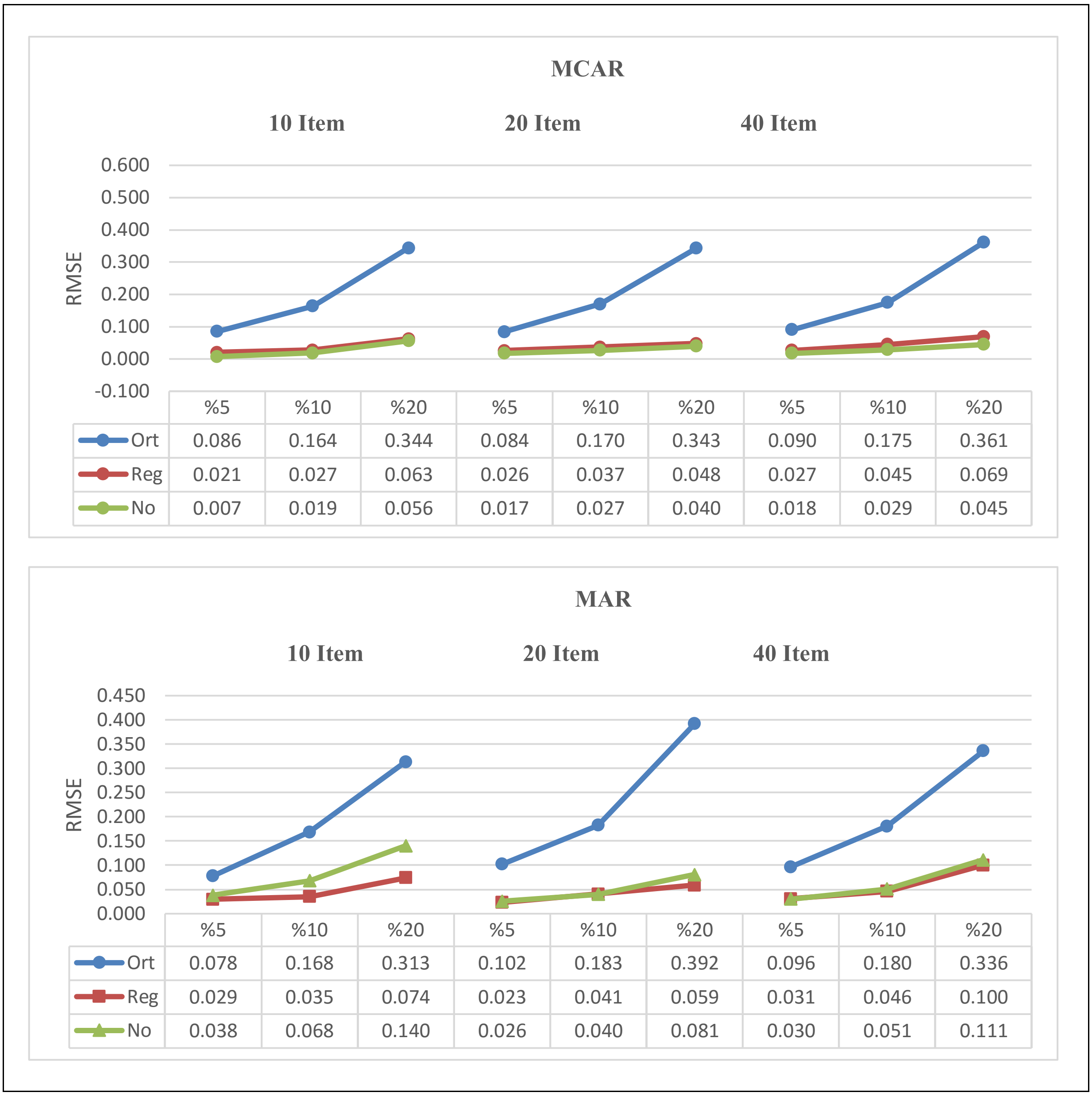


Figure 5. RMSE values of Pairwise Estimations of Item Difficulty Parameter (N=100)

According to Figure 6, the least erroneous pairwise estimations were obtained from leaving the missing data blank for all conditions of missing data rate and item length for N=1000 under MCAR missing data structure. Under all conditions, estimations obtained from regression imputation method and the method of considering the missing data blank were similar to each other while the most erroneous estimations were obtained from mean imputation method. Estimation errors obtained from each method increase as missing data rate increases while they are affected by item number at minimum level.

It can be seen that regression imputation method offers similar results with the method of leaving the missing data blank, but less erroneous results under many conditions within MAR missing data structure for N=1000. The most erroneous estimations were obtained from mean imputation method. Estimation errors obtained from each method increase as missing data rate increases while they are affected by item number at minimum level.



MCAR

	%5	%10	%20	%5	%10	%20	%5	%10	%20
Ort	0.086	0.164	0.344	0.084	0.170	0.343	0.090	0.175	0.361
Reg	0.021	0.027	0.063	0.026	0.037	0.048	0.027	0.045	0.069
No	0.007	0.019	0.056	0.017	0.027	0.040	0.018	0.029	0.045

MAR

	%5	%10	%20	%5	%10	%20	%5	%10	%20
Ort	0.078	0.168	0.313	0.102	0.183	0.392	0.096	0.180	0.336
Reg	0.029	0.035	0.074	0.023	0.041	0.059	0.031	0.046	0.100
No	0.038	0.068	0.140	0.026	0.040	0.081	0.030	0.051	0.111

Figure 6. RMSE Values of Pairwise Estimations of Item Difficulty Parameter (N=1000)

CONCLUSIONS

This study investigated the effect of missing data on Rasch model estimations under the conditions of missing data structure, missing data rate, methods of dealing with missing data, sample size and test length. In addition to likelihood estimations methods frequently used in the literature, pairwise method was also used as a different approach. It was aimed to attract attention to this method based on constrained pairwise for item categories.

The least erroneous constrained maximum likelihood (CML) estimations of item difficulty parameters were obtained from the method of considering the missing data blank under all conditions of sample size, test length and missing data rates under MCAR and MAR missing data structures. In the majority of the conditions, the most erroneous estimations were obtained from mean imputation method. In small sampling, estimation errors increased for each three methods to deal with missing data as missing data rate increased by item number. In big sampling, estimation errors increased as missing data rate increased in mean imputation method; estimation errors changed in minimal level according to missing data rates in the methods of regression imputation and considering the missing data blank. The increase in item number increased estimation errors in regression imputation and mean imputation methods, respectively only for small sampling; while the effect of item number on estimation errors was minimal under the rest of conditions.

The least erroneous joint maximum likelihood (JML) estimations of item difficulty parameters was obtained from the method of considering the missing data blank under all conditions of sample size, test length and missing data rates under MCAR and MAR missing data structures. In small sampling, it was observed that regression imputation method was most affected by the change in item number while estimation errors increased as item number increased under most of the conditions. In big sampling, on the other hand, regression imputation method was most affected by the increase in item number under MCAR structure while mean imputation method was most affected under MAR structure. Estimation errors increased as missing data rate increased in mean imputation method; estimation errors changed in minimal level according to missing data rates in the methods of regression imputation and considering the missing data blank.

The least erroneous pairwise estimations of item difficulty parameters was obtained from the method of considering the missing data blank under all conditions of test length and missing data rates for small samplings under MCAR and MAR missing data structures. In big samplings, it was found that regression imputation method gave similar, yet better results than the method of leaving the missing data blank under MAR structure and the most erroneous estimations were found from mean method. In small sampling, it was observed that estimation errors obtained from regression imputation method increased as item number increased.

Considering all simulation conditions, it was found that joint maximum likelihood (JML) method offered a better result compared to two other estimation methods. Constrained maximum likelihood (CML) and pairwise methods had similar performances under many conditions.

Estimation errors obtained from the method of leaving the missing data blank were lower than those obtained from mean and regression imputation methods under all conditions except for pairwise estimations under MAR structure and N=1000 condition. In general, the most erroneous estimations were obtained from mean imputation method. It was found that regression imputation method generally gave more biased results compared to JML-based estimations.

Errors of item difficulty parameter estimations obtained through three different estimation methods among three methods to deal with missing data increased as missing data rate increased (Andreis and Ferrari, 2012; Zhang and Walker,2008); and decreased as sample size increased. In big samplings, considering the missing data blank and regression-based estimations offered good results although missing data rate was high. It was observed that CML and Pairwise estimations based on mean and regression imputation were affected by test length under many conditions. However, Heine and Tarnai (2015) reported that Pairwise method gave better results than CML method under MCAR missing data structure. Considering that the study conducted by Heine and Tarnai (2015) was based on real data and depended on the result of a single application, performances CML and Pairwise methods can be investigated through different simulation studies.

According to the research, it can be suggested that mean imputation method should not be used where an imputation method is needed to deal with missing data during the parameter validation studies based on Rasch model and CML, JML and pairwise methods. Instead, regression imputation method which gives similar results as the ones obtained through missing data matrix can be preferred. JML method can give less erroneous results than CML and Pairwise methods in item parameter validation studies in Rasch model. This study showed that Pairwise method has a similar performance with CML method. Therefore, it can be suggested as an alternative estimation method for parameter estimation studies of researchers.

In further studies, the conditions studied in this research can be repeated for individual parameter estimations or for different levels of these conditions. Similarly, the effect of missing data can be examined in various item response theory models. It is assumed that the performance of pairwise method which is based on matrix with or without missing data during parameter validation studies must be investigated in similar or different studies.

REFERENCES

Andreis, F., & Ferrari, P. A. (2012). Missing Data And Parameters Estimates in Multidimensional Item Response Models. *Electronic Journal of Applied Statistical Analysis*, *5*(3), 431-437.

Andrich, D., & Luo, G. (2003). Conditional Pairwise Estimation in the Rasch Model for Ordered Response Categories using Principal Components. *Journal of Applied Measurement, 4(3),* 205–221.

Ayala, R. J., Plake, B. S., & Impara, J. C. (2001).(2001). The Impact Of Omitted Responses On The Accuracy of Ability Estimation in İtem Response Theory. *Journal of Educational Measurement*, *38*(3), 213-234.

Choppin, B. (1968). Item Bank Using Sample-free Calibration. *Nature*, 219(5156), 870–872.

Choppin, B. (1985). A fully Conditional Estimation Procedure for Rasch Model Parameters. *Evaluation in Education*, 9(1), 29–42.

Custer, M., Sharairi, S., & Swift, D. (2012). A Comparison of Scoring Options for Omitted and Not-Reached Items through the Recovery of IRT Parameters When Utilizing the Rasch Model and Joint Maximum Likelihood Estimation.*Online Submission*.

DeMars, C. (2002). Incomplete data and item parameter estimates under JMLE and MML estimation. *Applied Measurement in Education, 15*, 15–31.

DeMars, C. (2003) Missing Data and IRT Item Parameter Estimation. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.,
April , 2003. [ERIC Document Reproduction Service No. ED 476 175]

Eggen, T. J., & Verhelst, N. D. (2011). Item calibration in incomplete testing designs. *Psicológica: Revista de metodología y psicología experimental*, *32*(1), 107-132.

Embretson, S.E., & Reise, S.P. (2000). *Item response theory for psychologists.* Maheah, New Jersey: Lawrence Erlbaum Associates, Publishers

Finch, H. (2008). Estimation of Item Response Theory Parameters in the Presence of Missing Data. *Journal of Educational Measurement, 45(3),* 225-245.

Hambleton, R. K., & Swaminathan, H. (1985*). Item response theory: Principles and applications* (Vol. 7). Springer Science & Business Media.

Hambleton Ronald K, Swaminathan H, Rogers HJane. (1991).*Fundamentals of item response theory*. Newbury Park, CA: Sage Publications

Heine, J. H., & Tarnai, C. (2015). Pairwise Rasch model item parameter recovery under sparse data conditions. *Psychological Test and Assessment Modeling*, *57*(1), 3-36.

Heine, J., H. (2015). Package "pairwise", https://cran.r-project.org/web/packages/pairwise/pairwise.pdf.

Hohensinn, C., & Kubinger, K. D. (2011). On the impact of missing values on item fit and the model validness of the Rasch model. *Psychological Test and Assessment Modeling*, *53*(3), 380-393.

Joenssen, D.W. (2015). Package "HotDeckImputation", https://cran.r-project.org/web/packages/HotDeckImputation/HotDeckImputation.pdf.

Mair, P., Hatzinger, R., Maier, M. J., Rusch, T. (2015). Package "eRM", https://cran.r-project.org/web/packages/eRm/eRm.pdf.

Mislevy, R.J. & Wu, P-K (1996). Inferring examinee ability when some item responses are missing. Research Report RR-96-30-O R. Princeton: Educational Testing Service.

Rasch, G. (1966). *An informal report on the present state of a theory of objectivity in comparisons*. Universitetets Statistiske Institut.

Robitzsch, A. (2016). Package "sirt", https://cran.r-project.org/web/packages/sirt/sirt.pdf.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. Boca Raton, FL: Chapman
and Hall/CRC.

Shin, S. H. (2009). How to treat omitted responses in Rasch model-based equating. *Practical Assessment Research & Evaluation*, *14*(1).

Zhang, B., & Walker, C. M. (2008). Impact of missing data on person—model fit and person trait estimation. *Applied Psychological Measurement*, *32*(6), 466-479.