Trakya Eğitim Dergisi
Cilt 11, Sayı 3
Eylül 2021, 1227-1238

Geliş Tarihi: 03.09.2020

ISSN: 2630-6301
Doi: 10.24315/tred.789832
Araştırma Makalesi/
Research Article

Trakya Journal of Education
Volume 11, Issue 3
September 2021, 1227-1238

Yayına Kabul Tarihi:19.02.2021

EFFECTS OF DIFFERENT MULTIPLE IMPUTATION TECHNIQUES ON THE MODEL FIT OF CONFIRMATORY FACTOR ANALYSIS

FARKLI ÇOKLU VERİ ATAMA TEKNİKLERİNİN DOĞRULAYICI FAKTÖR ANALİZİ MODEL UYUMU ÜZERİNDEKİ ETKİSİ

Akif AVCU[1]

ÖZ: Bugüne kadar kayıp verilerin istatistiksel analizler üzerindeki etkilerini incelemek için birçok araştırma gerçekleştirilmiştir ve bu durumla başa çıkabilmek için farklı yöntemler geliştirilmiştir. Kayıp verinin silinmesini içeren yöntemler örneklem büyüklüğünün önemli miktarda azalmasına sebep olmakta ve analizlerin istatistiksel gücünü düşürmektedir. Bu duruma bir alternatif olarak önerilen kayıp veri kestirimine dayalı yöntemler araştırmacıların yoğun ilgisini çekmektedir. Bu yöntemler içerisinde çoklu veri atama teknikleri göreceli olarak daha yakın bir geçmişe sahiptir ve daha iyi kestirimler sağlamaktadır. Çoklu veri atama tekniklerinin üstünlüğü düşünüldüğünde, gerçekleştirilen bu çalışmanın amacı farklı çoklu veri atama tekniklerinin doğrulayıcı faktör analizi model uyumu üzerideki etkisinin değerlendirilmesidir. Bu amaç doğrultusunda örneklem büyüklüğü, kayıp veri mekanizması, kayıp veri yüzdesi, madde sayısı ve kayıp veri atama tekniğini kontrol edilerek tek boyutlu yapıya sahip veri setleri üretilmiştir. Kayıp veri tekniklerinin etkileri tam veri setleri ve veri ataması gerçekleştirilmiş veri setleri için elde edilmiş χ^2 model uyum istatistikleri arasındaki fark ile değerlendirilmiştir. Elde edilen sonuçlar çoklu veri atama tekniklerinin geleneksel regresyon temelli veri atama tekniğine kıyasla daha iyi sonuçlar sağladığını göstermiştir. Bu bulgular daha sonrasında tartışılarak daha iyi test uygulmaları için bir takım önerilerde bulunulmuştur.

Anahtar sözcükler: Kayıp veri, çoklu veri atama, benzetim.

ABSTRACT: So far, many types of research have been conducted to investigate the impact of missing data on statistical analysis and various methods have been developed to deal with the problem. The methods based on removing observations with missing values from the dataset cause the sample size to drop dramatically and the statistical power of the analyzes to be decreased. Therefore, as an alternative solution, the estimation of missing values seized intensive attention from researchers. Among these methods, multiple imputation techniques are relatively more recent and provide better estimations. Considering the superiority of multiple imputation techniques, the aim of the current study is to investigate the effects of different multiple imputation techniques on the model fit of confirmatory factor analysis. For this aim, datasets with the unidimensional structure were simulated to manipulate sample size, missing data mechanism, percentage of missing data, number of items, and missing data imputation technique. The effect of multiple imputation techniques was evaluated based on the difference of χ^2 model fit statistics for complete datasets and imputed datasets. The results showed that multiple imputation techniques provided better results than conventional regression-based imputation. Those finding were discussed later and some recommendations were given for better testing applications.

Keywords: Missing data, multiple imputation, simulation.

Bu makaleye atıf vermek için:
Avcu, A. (2021). Farklı Çoklu Veri Atama Tekniklerinin Doğrulayıcı Faktör Analizi Model Uyumu Üzerindeki Etkisi. *Trakya Eğitim Dergisi, 11*(3), 1227-1238

Cite this article as:
Avcu, A. (2021). Effects of different multiple imptutation techniques on the model fit of confirmatory factor analysis. *Trakya Journal of Education, 11*(3), 1227-1238

UZUN ÖZET

Giriş

Eğitim veya psikoloji alanında kullanılan ölçüm araçlarının nihai amacı, bir kişinin bir ya da daha fazla gizli değişken üzerindeki puanını tahmin etmektir. Ancak, bazı katılımcıların en az bir maddeye cevap vermemesi oldukça yaygındır. Yanıt veri matrisinin bu boş kısımları kayıp olarak adlandırılır.

[1] Arş. Gör. Dr., Marmara Üniversitesi, Eğitim Fakültesi, Psikolojik Danışmanlık ve Rehberlik Anabilim Dalı,, avcuakif@gmail.com, ORCID: 0000-0003-1977-7592

Kayıp veriler ile başa çıkmada geleneksel olarak veri silme ya da regresyon temelli veri atama yöntemleri tercih edilmektedir.

Kayıp verilerin üç farklı mekanizmaya sahip olabileceğini belirtilmiştir: tamamen rastgele (TR), rastgele eksik (R) ve rastgele olmayan (RO). Bir değişkendeki kayıp verilerin olasılığı, değişkenin kendisi veya veri setindeki diğer değişkenlerle ilgili değilse, kayıp veriler TR olarak kabul edilir. Diğer taraftan, bir değişkendeki kayıp veri olasılığı değişkenin kendisine değil de veri kümesindeki diğer değişkenlere bağlıysa, bu R mekanizmasıdır (Allison, 2002). Son olarak, bir değişkendeki kayıp veri olasılığı değişkenin kendisine bağlıysa, bu kayıp veri mekanizma RO olarak adlandırılır. TR test edilebilir tek kayıp veri mekanizmasıdır ve kayıp verilerle başa çıkmak için kullanılan tekniklerin çoğu, TR mekanizması için uygundur (Schafer and Graham, 2002).

Geleneksel olarak kullanılan veri silme yöntemlerine alternatif yaklaşım olan çoklu veri atama (MI) kayıp verilerin atanmasına yönelik günümüzde en gelişmiş çözüm olarak kabul edilir. Bu yaklaşımda tek bir veri atama yöntemi ile belirlenen değeri kullanmak yerine, iki veya daha fazla yöntemden elde edilen değerler dikkate alınarak atama gerçekleştirilir ve daha güvenilir sonuçlar verir (Doove, Van Buuren & Dusseldorp, 2014).

Kayıp veri sorunu yapısal eşitlik modelleme tekniklerini kullanan araştırmacılar için de geçerli bir sorundur (Jöreskog, 1977). Bu modeller test edildiğinde kayıp verilerin neden olduğu en ciddi sorun, parametre tahminlerine yanlılık karışması ve model uyumunun bozulmasıdır. Buradan hareketle farklı çoklu veri atama teknikleri olan Yordayıcı ortalama eşleştirme (PMM: predictive mean matching), Ağırlıklı yordayıcı ortalama eşleştirme WPMM: Weighted predictive mean matching), Sınıflandırma ve regresyon ağaçları (CART: Classsification and regression trees), Bayesian doğrusal regresyon (BLR: Bayesian linear regression) ve Bootstrap kullanarak doğrusal regresyon (LRUB: Linear regression using bootstrap) ve geleneksel doğrusal regresyon tahmini değerleri (PVLR: predictive value for linear regression) kullanılarak veri ataması gerçekleştirildiğinde bunun doğrulayıcı faktör analizi model uyumu üzerindeki etkileri açısından incelenmesi bu araştırmanın amacını oluşturmaktadır.

Yöntem

Verilerin Türetilmesi

Verilerin türetilmesi sürecinde R istatistik ortamında kullanılan "*mirt*" programı kullanılmıştır. Veriler türetilirken kayıp veri mekanizması (TR, R ve RO), kayıp veri yüzdesi (%5 ve %10), örneklem büyüklüğü (100, 250, 500 ve 1000) ve madde sayısı (5 ve 10) koşulları dikkate alınarak farklı özelliklere ait veri setleri elde edilmiştir. Verilerin tamamı tek boyutlu ve 1-5 aralığında puanlanan maddeler içeren test özellikleri dikkate alınarak türetilmiştir. Sonrasında ise R istatistik ortamında "*mice*" paketi (van Buuren & Groothuis-Oudshoorn, 2011) kullanılarak veriler belirli mekanizmalara göre silinmiş ve kayıp veriye sahip veri setleri elde edilmiştir. Daha sonra, aynı paket kullalarak farklı teknikler ile veri ataması gerçekleştirilmiş ve tam veri setleri elde edilmiştir. İlk türetilen veri setleri "gerçek" olarak kabul edilmiş ve veri ataması ile tamamlanan veri setleri ile karşılaştırılmıştır. Karşılaştırma işlemi için her iki veri seti ile gerçekleştirilen doğrulayıcı faktör analizinden elde edilen χ^2 istatistiklerinin ($\Delta\chi^2$ olarak ifade edilen) farkına bakılmıştır.

Verilerin Analizi

Doğrulayıcı faktör analizi R istatistik ortamında bulunan "*lavaan*" paketi kullanılarak gerçekleştirilmiştir. Sonuçların genellenebilirliğini artırmak için analizler 50 kez tekrarlanmıştır. Bulgular kısmında Tablo 1'de yer alan değerler bu 50 iterasyonun aritmetik ortalamasına karşılık gelmektedir.

Daha sonra, kontrol edilen koşulların, 50 iterasyondan sonra elde edilen $\Delta\chi^2$ ortalama değerler üzerindeki tekil etkileri ve etkileşimleri incelenmiştir. Bu analiz faktöryel ANOVA ile yapılmıştır. Faktöriyel ANOVA 2 (eksik veri yüzdesi) X 4 (örneklem büyüklüğü) X 2 (madde sayısı) X 6 (veri atama teknikleri) olarak tasarlanmıştır. Faktöriyel ANOVA'da eksik veri mekanizması bir koşul olarak alınmamış ve analizler her bir mekanizma için ayrı ayrı gerçekleştirilmiştir. Bu analizler SPSS for Windows 21.0 ile gerçekleştirilmiştir.

Değerlendirme Ölçütü

Bilindiği üzere daha küçük χ^2, daha iyi model uyumu anlamına gelir. Doğrulayıcı faktör analiz, tam veri seti ve veri ataması ile elde edilen veri seti için gerçekleştirildikten sonra ki kare değerleri arasındaki fark olan $\Delta\chi^2$ değeri hesaplanmıştır. Bu değer teorik olarak ± sonsuz aralığı arasında olabilir. Pozitif değerler, atama teknikleri ile elde edilen veri setiyle gerçekleştirilen analiz için elde edilen model uyumunun, tam veri setiyle elde edilenden daha kötü olduğunu gösterir. Bu nedenle sıfıra yakın değerler, kullanılan veri atama tekniğinin model uyumunu etkilemediğini gösterirken, değerlerin artması, atanan verilerle gerçekleştirilen analizin oldukça daha kötü uyum değerleri sağladığını gösterir.

Bulgular

Eksik veri yüzdesi n = 100 veri setleri için % 5'ten % 10'a çıktığında $\Delta\chi^2$ değerlerinin istisnasız olarak arttığı görülmüştür. Bu bulguya göre, eksik veri yüzdesi arttıkça model uyumunun kötüleştiği sonucuna varılabilir. Aynı zamanda başka bir bulguya göre en yüksek $\Delta\chi^2$ değerlerinin geleneksel regresyon tabanlı veri atama tekniği olan PVLR tekniği için elde edildiği görülmüştür. Benzer bulgular n = 250 veri setleri için de geçerlidir. Bu iki veri büyüklüğü için en ideal sonuçlar CART yöntemi için elde edilmiştir. Örneklem büyüklüğü n = 500 koşulu için BLR ve LRUB tekniklerinin CART ile karşılaştırılabilir sonuçlar verdiği görülmüştür. Son olarak, örneklem büyüklüğü n = 1000 olduğu veri setleri için en ideal sonuçları BLR ve LRUB teknikleri vermiştir. Her koşul altında model uyumunu en fazla olumsuz etkileyen teknik ise geleneksel PVLR olarak belirlenmiştir.

Daha sonrasında gerçekleştirilen faktöriyel ANOVA madde sayısı, örneklem büyüklüğü, kayıp veri yüzdesi ve kullanılan veri atama tekniğinin $p <0.01$ düzeyinde anlamlı tekil etkilere sahip olduğunu göstermiştir. Ayrıca, bu koşullar için test edilen ikili etkileşimlerin anlamlı olduğu bulunmuştur.

Sonuç ve Tartışma

Bulgular, örneklem büyüklüğünün 100 ve 250 olduğu veri kümeleri için, CART tekniğinin en düşük $\Delta\chi^2$ değerlerini sağladığını ortaya koymaktadır. Yani, CART tekniği uygulandığında model uyumu daha az etkilenmektedir. Hali hazırdaki alan yazın da bu bulguyu desteklemektedir. Elde edilen bir diğer bulguya göre ise eksik verilerin yüzdesi arttıkça, genellikle tüm yöntemler için daha yüksek $\Delta\chi^2$ değerleri elde edilmektedir. Benzer şekilde, örneklem büyüklüğünün de χ^2 değerleri üzerinde etkili olduğu bulunmuştur. Buna göre örneklem büyüklüğü arttıkça teknikler kayıp veri atamalarında modelle daha uyumlu değerler atamaktadırlar. Son olarak, eksik veri mekanizması açısından değerlendirildiğinde ise TR mekanizmasına sahip kayıp verilerle gerçekleştirilen atamaların model uyumu açısından daha avantajlı olduğu belirlenmiştir. Ek olarak, manipüle edilen koşulların iki yönlü ve üç yönlü etkileşimleri faktöriyel ANOVA ile incelenmiştir ve anlamlı tekil ve etkileşim etkilerinin olduğu belirlenmiştir. Bulgular, kullanılan tekniğin madde numarası, örneklem büyüklüğü ve eksik veri yüzdesi ile ikili etkileşimlere sahip olduğunu desteklemiştir. Bundan sonra gerçekleştirilecek olan çalışmalarda ise beklenti maksimizasyon yöntemi ile çoklu veri atama tekniklerinin karşılaştırılması önerilmiştir.

INTRODUCTION

The eventual aim of the educational or psychological tests is to estimate a person's score on one or more latent variables. The estimation is made based on a person's responses to a bunch of items. However, it is common to witness some participants do not respond to at least one item. These unavailable parts of the response the data matrix are called missing. In essence, missing data is a phenomenon that occurs as a result of complex interactions between the person's characteristics and the properties of the item (Rubin, 1976). Data loss may be due to many identified or unspecified reasons. For example, the test takers may leave one or more items unanswered by mistake, they do not know the answer or are afraid to guess. In the process of measurement and evaluation, the problem of missing item responses is a common problem. The best possible approach to deal with missing data is to plan the research well and to collect the data carefully (Wisniewski, Leon, Otto & Trivedi, 2006) but even the best testing conditions are satisfied, there is no guarantee for obtaining a full dataset with no missing part.

Even if the problem of nonresponse is as old as the history of measurement and researchers have made efforts to prevent it, missing data is still a problem in current studies and makes it difficult to make inferences about the latent trait intended to be measured (Hohensinn & Kubinger, 2011). An increasing number of researches have been conducted on the impact of missing data on statistical analysis and various methods have been developed to deal with the problem. The interested reader is encouraged to see Schafer and Graham (2002) for a comprehensive review of ways to deal with missing data. Even so, researchers still don't pay enough attention to missing data problems (Özberk, Kabasakal & Öztürk, 2017). In the following part, some basic concepts were provided to the readers about the phenomenon of the missing data imputation process.

Missing Data Mechanisms

Studies on missing data began to increase after Rubin's (1976) classification for missing data mechanisms. Accordingly, missing data are stated to occurs through three different mechanisms: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). These mechanisms define the relationship between the probability of missing data and the variables being measured.

If the probability of missing data in a variable is not related to the variable itself or to other variables in the data set, the missing data are accepted as having the MCAR mechanism. To put it in more concrete terms, in the MCAR mechanism, missing values in an X variable are independent of the other variables and the values of X itself. If the missing values follow an MCAR pattern, unavailable observations from the list could give a random sample of the complete data set. MCAR put more limitation on the nature of missingness than other mechanisms and is a strict assumption that is not easy to meet in practice (Muthén, Kaplan & Hollis, 1987) because it is based on the assumption that missingness is not related to data itself (Allison, 2002). On the other hand, if the probability of missing data in a variable does not depend on the variable itself but depends on other variables in the data set, the missingness in such data is called MAR. In other words, the probability of missingness in variable X is not related to the value of X after other variables are statistically controlled. This assumption is weaker compared to MCAR (Allison, 2002) and is less restrictive as the observed values need not be a simple random sample of the complete data set. Although the MAR mechanism is called random, there is a systematic mechanism in the missingness pattern (Baraldi and Enders, 2010). Finally, if the probability of missing data in a variable depends on the variable itself, this missingness mechanism is called MNAR. MNAR mechanism implies that the probability of missing data in the X variable is related to itself even after other variables in the data set are statistically controlled. If the missing data has MNAR mechanism, it will have no random pattern. Among the other ones, MCAR is the only testable missing data mechanism (Peugh & Enders, 2004) and most of the techniques to deal with missing data assume the MCAR mechanism. On the other hand, if the missing data has a MAR or MNAR mechanism, it is not possible to prove this statistically.

Strategies to Deal With Missing Data

For many years, researchers have used different methods (such as deletion, ignoring, or replacing the missing value with an appropriate one) to overcome the missingness problem. However, although ignoring missing data is cited as an alternative, it is not preferred as it causes a significant loss of information (Guan & Yusoff, 2011). For this reason, it is possible to gather methods of dealing with missing data under two main broad subcategories. These are namely deletion and imputation of missing data. The deletion of missing data could be conducted as pairwise and listwise deletion (Marshwise 1998; Roth, 1994). In both cases, analyzes are only performed with the remaining complete dataset.

In listwise deletion, the researcher only uses cases to respond to all variables; observations with missing values are removed. In other words, cases with at least one missing value are deleted from the dataset, and analyzes are performed only with individuals whose data is complete. Although the method has desirable properties such as ease of implementation and comparability of univariate statistics, it can result in inefficient parameter estimates by possible deletion of large amounts of data. Listwise deletion is the most common method for handling missing data, but its implementation depends on the assumption of the MCAR. When the MCAR assumption doesn't hold, listwise deletion can cause bias in parameter estimates (Rahman & Davis, 2013).

The method of pairwise deletion uses all possible data at hand. In this method, analysis is made with the available part of the data, disregarding the missing part. Cases in the data set are not deleted but missing data are not taken into account (Durant, 2005). The difference of this method from listwise deletion is that the unit that contains missing data on some variable is not eliminated from the dataset when it has full data for variables in concern while this unit is excluded when variables are used for analysis where this unit has missing data for this specific variable(s) (Howell, 2007).

Researches have shown that listwise deletion could give unbiased parameter estimates under MCAR (Wothke, 2000) and biased estimates under MAR (Muthén et al., 1987). In addition, it has been reported that listwise deletion provides less efficient parameter estimates than other methods (Arbuckle, 1996). Like listwise deletion, pairwise deletion also requires the assumption of MCAR and yields biased parameter estimates under MAR (Wothke, 2000).

Deletion methods can be preferable when the data contains small amounts of missing data. However, since it causes data loss, it can lead to serious errors and is therefore not usually a very efficient method (Nie et al., 1975). In fact, if the amount of missing data is too large, the problem of singularity can be encountered after removing missing subjects. It means that the number of items becomes higher than the number of observations, and the estimated variance-covariance matrices may be non-positive-definite (Nassiri et al, 2018).

Removing observations with missing values from the dataset causes the sample size to drop dramatically and the statistical power of the analyzes to be decreased. Therefore, as an alternative solution to these methods, estimation of missing values was proposed. These methods are based on estimating missing data with some data imputation methods and analyzing the completed dataset. These methods are based on assigning a value to the cell where the data is missing. These methods are preferred because they do not affect the sample size and prevent parameter estimates from being biased (Kim & Curry, 1977).

Among these methods, the most commonly preferred ones are mean imputation, regression imputation, expectation-maximization (EM) algorithm, and multiple imputation. The mean imputation approach is based on the assignment of the average value of a variable to missing values for this variable. Although mean imputation is widely used due to its easiness, assigning the average value to all of the missing data will decrease the variance and will cause the covariance values to decrease for multivariate analysis.

The regression method is based on the estimation or prediction of variables containing missing data using information from variables that do not contain. The superiority of the approach on mean imputation is that it preserves sample variation and covariation between variables. However, it was reported that mean imputation and regression imputation may disrupt the marginal and joint distributions of variables (Little and Rubin, 2002). One further disadvantage of the regression method is that the existing relationship in the dataset may become stronger and the variability decreases (Kros & Brown, 2003).

Another widely preferred alternative method to overcome these limitations is the EM method which was proposed by Dempster, Laird and Rubin (1977). The EM method is an iterative and two-step method, where the E step corresponds to the best possible estimates for missing data, while the M corresponds to the estimates of the mean, standard deviation, or correlation when missing data is imputed. This process continues until the change in the estimated values decreases considerably (Alpar, 2003).

The most important advantage of the EM approach over the regression method is that it uses more information in estimating missing data (Oğuzlar, 2001). On the other hand, the EM algorithm can be a very time-consuming process when datasets containing too much missing data (Bennett, 2001). In addition, the biggest disadvantage of EM is that it does not calculate values such as standard error and confidence interval during parameter estimation (Xu, Baines, & Wang, 2014). Therefore, even if the parameter estimates are very strong, it is not possible to perform hypothesis testing with the estimates obtained with EM.

A more recent alternative approach to the traditionally used Regression analysis and EM approach is multiple imputation (MI) and is considered the state-of-the-art solution to the imputation of missing data (Doove, Van Buuren & Dusseldorp, 2014). This method was firstly proposed by Rubin in 1987. Instead of using the value determined by a single imputation method, it is based on the use of the imputed value obtained from two or more methods. Therefore, the method of multiple imputation aims to obtain a combined estimated value. This combined estimation value is usually the average of the estimated values obtained by two or more methods. Multiple estimation method yields more reliable results than the values obtained by a single method. In addition, it gives robust results when the variables in the analysis violate the normality (Alpar, 2003).

Contrary to what its name implies, the MI is composed of a set of different techniques. For detailed theoretical explanation and formulations of these different techniques readers are encouraged to look at Doove, van Buuren, Dusseldorp, 2014), Van Buuren (2007), Breiman et al. (1984), Rubin

(1986), and Little (1988). In a recent paper, twenty-four different multiple imputation techniques were reported (van Buuren & Groothuis-Oudshoorn, 2011).

Missing Data in the Context of Confirmatory Factor Analysis

Missing data problems are also a concern for researchers who use structural equation modeling (SEM) techniques where it is also highly probable to face missing data for one or more variables. Even missing data problem is the case, confirmatory factor analysis, and structural equation modeling (see, for example, Jöreskog, 1977) still need to be applied. In a study by Fabrigar, Wegener, MacCallum, and Strahan (1999), it was found that approximately 50% of factor analysis studies had sample sizes below 200. In these studies, deleting the data will cause significant data loss, and analysis will lose generalizability power, biased factor loading may be estimated, some problems may occur related to model convergence and even extra factors could be extracted. Deleting cases with missing data is a greater concern with smaller examples (van Praag, Dijkstra & Van Velzen, 1985). For this reason, it is more convenient to maintain the sample size by using appropriate data imputation techniques instead of deleting data in factor analysis studies.

According to Peng, Harwell, Liou, Ehman (2006) when a statistical model is applied, the most serious problem caused by the missing data is the problem of bias to the parameter estimations. The effect of imputation techniques on Item Response Theory-based model estimations was already proved (Kalkan, Kara & Kelecioğlu, 2018). Statistical analysis results in biased parameter estimates, reduced accuracy, and deterioration of model fit when missing data is not handled properly. As a result, it can cause the drawing of erroneous conclusions.

Aim of the Study

Based on the fact that multiple imputation methods give better results compared to traditional methods, multiple imputation methods should be preferred when imputing data (Allison, 2003; Misztal, 2012). The effectiveness of multiple imputation techniques overestimation of reliability of scores was already proved (Akın Arıkan & Soysal, 2018) while few findings are available on its effect on validity (eg. Çüm & Gelbal, 2015). In line with this fact, the aim of this study was to compare different multiple imputation methods in terms of their effectiveness of model fit of confirmatory factor analysis. For this aim, five different multiple imputation techniques (Predictive Mean Matching; PMM, Weighted Predictive Mean Matching; WPMM, Classification and Regression Trees; CART, Bayesian Linear Regression; BLR and Linear Regression using Bootstrap; LRUB) and classical Linear Regression predicted values (PVLR) were compared in terms of their effect on model fit when confirmatory factor analysis was applied.

METHOD

Data Simulation

A comprehensive data simulation process and CFA analyses were carried out for the purpose of the study. The simulation process consists of three main steps: (1) creation of data sets, (2) creation of datasets with missing values, and (3) imputation of missing values. Accordingly, datasets having unidimensional structure are simulated in the first step. Later, by using one of the two missing data mechanisms (MAR or MCAR) some cases were deleted from the dataset. Later, datasets were imputed with multiple imputation techniques and conventional regression-based imputation.

Taking into account the study (Henson and Roberts, 2006) that states that most studies have fewer than 200 observations, the sample size conditions of the current study were manipulated to be in a wider range. Further, as Schafer (1999) stated, missing data percentage of %5 is desirable while %10 represents the rate where biased estimates are possibly observed (Bennett, 2001). In addition, the item number of 5 and 10 represent typical numbers of items for scales with unidimensional structure. Accordingly, simulated item numbers were selected based on this reality. All data sets were simulated as having a unidimensional structure.

In general, five different variables were manipulated in this study: sample size (100, 250, 500, and 1000), missing data mechanism (MAR and MCAR), percentage of missing data (5% and 10%), and the number of items (5 and 10) and missing data imputation technique (PMM, WPMM, CART, BLR, LRUB, and PVLR). Considering these conditions, complete datasets are simulated using the “simdata” function in “mirt” (Chalmers, 2012) package available R statistical environment (R Core Team, 2020). The data simulation process was repeated 50 times to increase the generalizability of the results.

Analysis

The data deletion process was performed to obtain datasets with missing values by using the "mice" (van Buuren & Groothuis-Oudshoorn, 2011) package in the R statistical environment. Later, the same package was used for the data imputation process. Investigation of the effect of data the imputation techniques on CFA results was achieved by comparing the model fit results obtained from both the complete dataset before creating missing datasets and imputed datasets. More clearly, unidimensional CFA was performed for both the complete data set and the imputed data set. The comparison was made based on the difference of χ^2 statistics of these analyzes (as denoted $\Delta\chi^2$). CFA was carried out using the "lavaan" (Rosseel, 2012) package in the R statistical environment.

Later, the unique and interaction effects of conditions being manipulated on $\Delta\chi^2$ values of 50 iterations were examined. These analyses were performed with factorial analysis of variance (ANOVA). The design of factorial ANOVA was 2 (missing percentage) X 4 (sample size) X 2 (number of items) X 6 (imputation techniques). Missing data mechanism was not taken as a condition in factorial ANOVA and analyses were repeated for both MAR and MCAR conditions. These analyses were performed in SPSS for Windows 21.0.

Evaluation Criteria

For the current study, χ^2 model fit statistics were used to compare the imputation techniques. As known, smaller χ^2 implies a better model fit. CFA was performed with the complete dataset and imputed dataset. The chi-square value of CFA with the complete dataset was regarded as expected value or "true" value. Later, the same CFA procedure was conducted with the imputed dataset and the χ^2 value obtained from this analysis was regarded as an observed value. Later, the difference of χ^2 value was calculated by simply subtracting the expected chi-square value from the observed value, and $\Delta\chi^2$ values were obtained for each iteration. This value can theoretically be between the range of ± infinity. Positive values indicate that the model fit of the CFA performed with the imputed dataset is worse than the one obtained with the complete dataset while negative values indicate that the model fit is better for the imputed dataset. The value of zero indicates that the model fit of CFA for both datasets is exactly the same. For this reason, values close to zero imply that the data imputation technique used does not affect the model fit and this is the desired result while the deviation of the positive values from zero indicates that the CFA performed with imputed data provide a quite worse model fit values which are not desirable.

FINDINGS

The aim of this study is to examine the effectiveness of different data imputation techniques on CFA goodness of fit results across different conditions (sample size, the number of items, the percentage of missing data, and missing data mechanism). Table 1 shows that the values of $\Delta\chi^2$ increase without exception when the missing data percentage increases from 5% to 10% for n=100 data sets. Based on this finding, it could be inferred that the model fit worsened as the missing data percentage increases. At the same time, according to another finding, it was observed that the highest $\Delta\chi^2$ values were obtained for the PVLR technique which represents traditional regression-based data imputation. This result implies that when data imputation was carried out with the PVLR technique, imputation yields a dataset incompatible with the underlying model and worsens the model fit. On the other hand, the lowest values of $\Delta\chi^2$ were obtained for the CART technique.

As the sample size in datasets was set at n = 250, the analysis revealed that $\Delta\chi^2$ values obtained for data sets with the missing percentage of 5% were similarly lower than those obtained for data sets with a 10% missing percentage. This result implies that as the percentage of imputed data increases, the predicted model fit deteriorates more. In addition, in line with the previous findings, the highest $\Delta\chi^2$ values were obtained for the PVLR method while the lowest values were obtained for the CART technique. According to this finding, the model fit worsens relatively less when using the CART technique. In addition, when the number of items was set to 5, the lowest $\Delta\chi^2$ values were observed for the MAR mechanism and the missing percentage is 5%, while the highest values are obtained for the data sets with the MAR missing mechanism and the missing percentage is 10%. On the other hand, when there are 10 items in the dataset and the missing values were 10%, higher $\Delta\chi^2$ values were observed whereas the lowest values were found for data sets with the MCAR missing data mechanism and the missing percentage is 5%.

For the datasets with a sample size of 500, the CART technique was still the best-performing one while BLR and LRUB techniques were able to give more comparable results compared to n = 100 and n = 250 conditions. On the other hand, the PVLR is still the worst-performing technique because it gave the highest $\Delta\chi^2$ values. That is to say, when the sample size was increased to 500, the model fit of CFA after implementing CART, BLR, and LRUB techniques was relatively less affected, while the PVLR technique deteriorated the model fit results. Compared across the conditions, the lowest $\Delta\chi^2$ values were obtained for data sets with the MAR missing data mechanism, where the missing data percentage was 5% and the highest $\Delta\chi^2$ values were obtained for the data sets with the MAR missing data mechanism where the missing data percentage was 10%. For data sets with 10 items, it was observed that $\Delta\chi^2$ values did not change much according to the missing data mechanism, but differences were observed based on the missing data percentage.

Finally, as the sample size rose to 1000, it was observed that the lowest $\Delta\chi^2$ values were obtained for LRUB and BLR techniques, respectively. The PVLR, on the other hand, yields the highest $\Delta\chi^2$ values. When evaluated in terms of tested conditions, for data sets where the number of items is 5, the lowest $\Delta\chi^2$ values were observed for the MAR missing data mechanism and for 5% missing data percentage, while the highest values are observed for the data sets with the MAR mechanism for 10% missing data percentage. On the other hand, in data sets where the number of items is 10, changes due to missing data mechanisms were smaller, while $\Delta\chi^2$ changes due to missing data percentage were notable and the lowest values were observed for data sets where the missing data percentage is 5%. Finally, when evaluated across different sample size conditions, it was found that $\Delta\chi^2$ values did not show significant changes in terms of different sample conditions.

Table 1.

$\Delta\chi^2$ values for different imputation techniques accross different conditions

n	Item #	Type	% of missing	PMM	WPMM	CART	BLR	LRUB	PVLR
100	5 items	MAR	5%	3.85	3.65	3.09	2.60	3.30	3.99
			10%	6.97	6.31	6.83	6.58	7.69	12.03
		MCAR	5%	1.50	0.97	1.43	1.10	1.81	2.44
			10%	6.00	6.06	6.50	6.81	6.75	10.63
	10 items	MAR	5%	14.39	9.41	8.66	11.83	10.71	16.48
			10%	25.76	19.11	17.68	27.02	25.43	44.29
		MCAR	5%	13.10	9.41	8.65	11.96	11.07	16.14
			10%	24.70	18.44	15.79	22.14	24.64	50.89
250	5 items	MAR	5%	1.64	1.52	2.40	1.17	1.41	2.02
			10%	7.50	4.76	5.72	6.38	6.34	10.13
		MCAR	5%	2.84	2.37	2.83	2.21	2.24	3.32
			10%	3.72	3.45	3.17	2.04	3.66	6.22
	10 items	MAR	5%	16.71	12.77	11.04	13.37	13.20	17.97
			10%	23.48	22.67	20.01	21.99	25.13	39.48
		MCAR	5%	10.23	7.47	6.65	7.56	8.21	12.91
			10%	27.06	23.19	19.02	22.38	23.41	39.80
500	5 items	MAR	5%	2.57	2.04	2.52	2.41	2.67	3.23
			10%	6.54	6.03	5.33	4.54	6.09	7.79
		MCAR	5%	2.99	3.34	2.88	2.70	2.97	3.46
			10%	4.21	3.41	5.03	3.41	3.51	6.99
	10 items	MAR	5%	12.84	11.99	12.37	11.47	11.56	16.15
			10%	24.53	24.19	19.09	23.76	20.89	26.78
		MCAR	5%	13.47	10.14	11.97	10.64	10.04	14.77
			10%	22.30	23.61	18.79	19.56	22.23	36.37
1000	5 items	MAR	5%	0.98	1.36	1.25	1.31	1.21	1.84
			10%	6.79	5.73	6.07	4.94	4.18	8.26
		MCAR	5%	2.14	2.19	2.40	1.89	1.91	2.63
			10%	6.26	4.32	5.13	3.93	4.20	6.61

10 items	MAR	5%	11.44	12.24	11.40	10.26	9.51	13.46
		10%	27.32	23.36	21.56	20.68	21.20	37.38
	MCAR	5%	12.03	12.21	10.10	10.47	10.47	13.73
		10%	29.08	26.18	22.66	23.32	21.94	38.05

The findings of factorial ANOVA were given in Table 2. The findings in the table were interpreted only in terms of missing imputation techniques because the focus of this study is to compare the effectiveness of imputation techniques and specific results for other variables are beyond the scope of this research. The analyses were conducted separately for datasets where MAR and MCAR missing data mechanisms exist. In addition, when necessary, multiple comparisons were performed with Tukey post hoc analysis.

The results revealed that there were significant unique effects for each item number (IN), sample size (SS), missing percentage (MP), and imputation technique (T) conditions at $p<0.01$ significance level. When this finding is evaluated together with the findings in Table 1, it can be inferred that decreasing the missing percentage in the data set also caused a significant decrease in $\Delta\chi^2$ values and provided model fit results closer to values obtained from the complete dataset. When evaluated in terms of the effects of the techniques, the HSD test shows that PMM and PVLR give relatively higher $\Delta\chi^2$ values ($p<0.01$) while CART provides significantly the lowest $\Delta\chi^2$ values than PMM, PVLR, and LRUB techniques. These observed unique effects are valid for both MAR and MCAR mechanisms. These findings support that the preferred missing data imputation technique changes the model fit of further CFA considerably.

When two-way interactions are examined, the results showed that there are significant two-way interactions between T condition and IN, SS, and MP conditions for both missing data mechanisms ($p <0.01$). Similarly, when evaluated in terms of three-way interactions, it was determined that the T condition was found to be in interaction with IN - SS ($p <0.05$), IN - MP ($p <0.01$), and SS - MP ($p <0.01$) conditions. These findings showed that these observed variations across techniques can further differ as a result of their interaction with sample size, percentage of missing data, and sample size conditions. Finally, results suggested that no four-way interaction was observed ($p>0.05$).

Table 2.

Unique and interaction effects of conditions being manipulated and imputation techniques.

	MAR		MCAR	
Conditions and Interactions	*F*	*p*	*F*	*p*
Item Number (IN)	3026.71	.001**	3101.38	0.001**
Sample Size (SS)	5.27	.001**	5.56	0.001**
Missing Percentage (MP)	1081.11	.001**	1098.95	0.001**
Techique (T)	68.84	.001**	77.74	0.001**
IN * SS	2.22	.084	1.67	0.171
IN * MP	260.12	.001**	520.41	0.001**
IN * T	30.89	.001**	42.34	0.001**
SS * MP	2.18	.089	7.24	0.001**
SS * T	2.08	.009**	3.57	0.001**
MP * T	25.49	.001**	30.29	0.001**
IN * SS * MP	2.73	.042*	8.24	0.001**
IN * SS * T	1.41	.132	2.02	0.011*
IN * MP * T	9.55	.001**	16.67	0.001**
SS * MP * T	1.16	.294	1.80	0.029*
IN * SS * MP * T	0.68	.807	1.41	0.133

Note: **p<0.01; *p<0.05

DISCUSSION AND CONCLUSION

The aim of this study is to investigate the effects of PMM, WPMM, CART, BLR, LRUB, and PVLR multiple imputation techniques on CFA model fit. The findings revealed that for datasets where the sample size is 100 and 250, the CART technique provides the lowest $\Delta\chi^2$ values which imply that when the imputation was conducted with this technique model fit of the data was less negatively affected. On the other hand, as the sample size was increased to 500, BLR and LRUB techniques yielded comparable χ^2 values to the CART technique, and for data sets with a sample size of 1000, BLR and LRUB techniques provided even better results. A recent study carried out by Chhabra Vashisht & Ranjan, 2017) also reported that CART, LRUB, and BLR performed better performances compared to other imputation techniques in terms of the standard errors they yield for imputing missing values. Similarly, the superiority of the CART was also reported in one another study (Hayes, Usami, Jacobucci & McArdle, 2015). On the other hand, the PVLR technique has been was found to be the worst-performing technique for all conditions without an exception. Since this method is based on traditional linear regression, the superiority of multiple imputations over the traditional method was supported once again in this study context (i.e. Allison, 2003).

According to another finding, as the percentage of missing data increases, it was generally found that all methods yield more χ^2 values. This result implies that, the reproducibility of model fit results after imputation is highly related to the missing data percentage. Similarly, the sample size was also found to be effective on χ^2 values. Accordingly, as the sample size increases, the techniques impute missing values more compatible with the model. Finally, if we evaluate in terms of missing data mechanism, results revealed that when the dataset has missing values with MCAR mechanism, further model fit results were found to be less affected after imputation. Similarly, in the study conducted by Yuan and Bentler (2000), it was stated that the analyzes performed in datasets with MCAR structure were less biased. Those findings could be interpreted as indirectly supporting the findings obtained in the current study.

In addition, two-way and three-way interactions of manipulated conditions were examined by factorial ANOVA. The findings suggested that the technique used has binary interactions with item number, sample size, and percent of missing data. For example, depending on the number of items, the CART method yields more similar $\Delta\chi^2$ values, while the values obtained for the PVLR showed more variability. Similar findings were observed for the percentage of missing data. When the three interactions were examined, it was determined that the technique variable has significant interactions with item number & sample size, item number & missing percentage, and sample size & missing percentage pairs. These findings showed that the effectiveness of multiple imputation techniques may change according to different conditions, and it is very important for researchers to choose the right techniques considering these complex interactions.

This study provided valuable information for researchers. On the other hand, as in every study, it is not free of some limitations. For example, only the regression method was preferred when comparing multiple imputation methods with conventional techniques while the EM method was not included. Even the multiple imputation methods were already been reported as superior (Misztal, 2012), the comparison of multiple imputation with EM could provide further insights into the literature. Similarly, only five multiple imputation was given a place in this study. Further research is strongly recommended which includes more multiple imputation techniques to compare. In addition, model-based item response theory techniques should also be utilized when imputing missing data. Future research is clearly needed to compare the item response theory-based method with multiple imputation techniques on item response theory-based calibration results. Finally, this study was conducted with simulated unidimensional datasets. A future study should be conducted when real datasets are used and/or when datasets have a multidimensional structure.

REFERENCES

Akın Arıkan, C., & Soysal, S. (2018). Investigation of Reliability Coefficients According to Missing Data Imputation Methods. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *33*(2), 316-336.

Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: Sage Publications.

Alpar, R. (2003). *Uygulamalı Çok Değişkenli İstatistiksel Yöntemlere Giriş 1*, 2. Basım, Ankara: Nobel Yayın Evi.

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling* (pp. 243-277). Mahwah, NJ: Lawrence Erlbaum Associates, Inc

Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of school psychology*, *48*(1), 5 37.

Bennett D. A. (2001). How can I deal with missing data in my study?. *Australian and New Zealand journal of public health*, *25*(5), 464-469.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., (1984). *Classification and Regression Trees*. Wadsworth, Pacific Grove, CA.

Chalmers, R.P. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software*, *48*(6), 1-29.

Chhabra, G., Vashisht, V., & Ranjan, J. (2017). A comparison of multiple imputation methods for data with missing values. *Indian Journal of Science and Technology*, *10*(19), 1-7.

Çüm, S, Gelbal, S. (2015). Kayıp Veriler Yerine Yaklaşık Değer Atamada Kullanılan Farklı Yöntemlerin Model Veri Uyumu Üzerindeki Etkisi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, *1*(35) , 87-111.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society*, Series B, *39*, 1–38.

Doove, L. L., Van Buuren, S., & Dusseldorp, E. (2014). Recursive partitioning for missing data imputation in the presence of interaction effects. *Computational Statistics & Data Analysis*, *72*, 92-104.

Durant, G.B. (2005). *Imputation Techniques to handling item-nonresponse in the social sciences: A methodical Review*, National Centre for Research Methods Working Paper Series, ESRC National Centre for Research Methods and Southampton Statistical Sciences Research Institute (SRI), University of Southampton.

Enders, C.K. (2010). *Applied missing data analysis*. Guilford press.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272–299.

Guan, N. C., & Yusoff, M. S. B. (2011). Missing values in data analysis: Ignore or Impute?. *Education in Medicine Journal*, *3*(1): e6-e11

Hayes, T., Usami, S., Jacobucci, R., & McArdle, J. J. (2015). Using Classification and Regression Trees (CART) and random forests to analyze attrition: Results from two simulations. *Psychology and aging*, *30*(4), 911-929.

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

Howell, D. C. (2007) The analysis of missing data. In: Outhwaite, W. & Turner, S. (Eds.) *Handbook of Social Science Methodology*. London: Sage.

Jöreskog, K.G. (1977), Structural Equation Models in the Social Sciences: Specification, Estimation and Testing. In: Krishnaiah, P. R. (Eds), *Applications of Statistics*,. Amsterdam: North-Holland Publishing Co., pp. 265-387.

Kalkan, Ö., Kara, Y., Kelecioğlu, H. (2018). Evaluating Performance of Missing Data Imputation Methods in IRT Analyses. *International Journal of Assessment Tools in Education*, *5*(3) , 403-416 .

Kim, J., & Curry, J. (1977). The Treatment of Missing Data in Multivariate Analysis. *Sociological Methods & Research*, *6*, 215-240.

Kros, J. F., & Brown, M. L. (2003). Data mining and the impact of missing data. *Industrial Management & Data System*, *103*(8), 611-621.

Little R.J.A. (1988). Missing data adjustments in large surveys (with discussion). *Journal of Business Economics and Statistics*, *6*, 287-301

Little, R.J.A. & Rubin, D.B. (2002). *Statistical analysis with missing data*. (Second Edition.) New York: Wiley.

Marsh, H. W. (1998). Pairwise deletion for missing data in structural equation models: Nonpositive definite matrices, parameter estimates, goodness of fit, and adjusted sample sizes. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(1), 22-36.

Misztal, M. (2012). Imputation of Missing Data Using R Package. *Acta Universitatis Lodziensis Folia Oeconomica*, *269*, 132-144.

Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika*, *52*(3), 431-462.

Nassiri, V., Lovik, A., Molenberghs, G. et al. (2018). On using multiple imputation for exploratory factor analysis of incomplete data. *Behavioral Research Methods*, *50*, 501-517.

Nie, W. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K. and Bent, D. H. (1975). *SPSS Statistical package for the social sciences*. McGraw-Hill, New York.

Oğuzlar, A. (2001). *Alan Araştırmalarında Kayıp Değer Problemi ve Çözüm Önerileri*, 5. Ulusal Ekonometri ve İstatistik Sempozyumu, Adana: Çukurova Üniversitesi, 20-22 Eylül 2001, s.1-28.

Özberk, E.H., Kabasakal, K.A., & Öztürk, N.B. (2017). Investigating the Factors Affecting Turkish Students' PISA 2012 Mathematics Achievement Using Hierarchical Linear. *Hacettepe University Journal of Education*. *32*, 554-559.

Peng, C., Harwell, M., Liou, S., Ehman, L. (2006). Advances in missing data methods and implications for educational research. In Sawilowsky, S. S. (Eds.), *Real data analysis* (pp. 31–78). Charlotte, NC: New Information Age.

Peugh, J.L. and Enders, C.K. (2004). Missing data in educational research. *Review of educational research*. *74*(4), 525-556.

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rahman M.M., Davis D.N. (2013) Machine Learning-Based Missing Value Imputation Method for Clinical Datasets. In: Yang GC., Ao S., Gelman L. (Eds.), *IAENG Transactions on Engineering Technologies. Lecture Notes in Electrical Engineering*, vol 229. Springer, Dordrecht

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1-36.

Roth, P. L. (1994). Missing data: A conceptual review for applied psychologists. *Personnel Psychology*, *47*, 537-560.

Rubin D.B. (1986). Statistical matching using file concatenation with adjusted weights and multiple imputations. *Journal of Business Economics and Statistics*, *4*, 87-94

Rubin, D.B. (1976). Inference and missing data. *Biometrika*, *63*, 581-592.

Rubin, D.B. (1987). *Multiple Imputationfor Nonresponse in Surveys*, New York: John Wiley.

Schafer, J. L. (1999). Multiple imputation: a primer. Statistical methods in medical research, 8(1), 3-15.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147-177.

van Buuren S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical methods in medical research*, *16*(3), 219-242.

van Buuren, S, Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3), 1-67.

van Praag, B.M.S., T. K. Dijkstra & J. van Velzen (1985) Least squares theory based on general distributional assumptions with an application to the incomplete observations problem. *Psychometrika*, *50*, 25-36.

Wisniewski S.R., Leon A.C., Otto M.W. ve Trivedi M.H., (2006). Prevention of missing data in clinical research studies. *Biological Psychiatry*, *59*(11), 997-1000

Wothke, W. (2000). Longitudinal and multi-group modeling with missing data. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multiple group data: Practical issues, applied approaches and specific examples* (pp. 219–240). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Xu, C., Baines, P. D., & Wang, J. L. (2014). Standard error estimation using the EM algorithm for the joint modeling of survival and longitudinal data. *Biostatistics (Oxford, England)*, *15*(4), 731-744.

Yuan, K.-H., & Bentler, P. M. (2000). 5. Three Likelihood-Based Methods for Mean and Covariance Structure Analysis with Nonnormal Missing Data. *Sociological Methodology*, *30*(1), 165–200.