Statistica Sinica **31** (2021), 2309-2330
doi:https://doi.org/10.5705/ss.202020.0386

IDENTIFIABILITY OF BIFACTOR MODELS

Guanhua Fang[1], Jinxin Guo[2], Xin Xu[2], Zhiliang Ying[3] and Susu Zhang[4]

[1]*Baidu Research,* [2]*Northeast Normal University,*
[3]*Columbia University and* [4]*University of Illinois at Urbana-Champaign*

Abstract: The bifactor model and its extensions are multidimensional latent variable models, under which each item measures up to one subdimension on top of the primary dimension(s). Despite their wide applications to educational and psychological assessments, these multidimensional latent variable models may suffer from nonidentifiability, which can further lead to inconsistent parameter estimation and invalid inference. The current work provides a relatively complete characterization of identifiability for linear and dichotomous bifactor models and the linear extended bifactor model with correlated subdimensions. In addition, similar results for the two-tier models are developed. Illustrative examples on checking model identifiability by inspecting the factor loading structure are provided. Simulation studies examine the estimation consistency when the identifiability conditions are/are not satisfied.

Key words and phrases: Bifactor model, educational and psychological measurement, identifiability, item factor analysis, multidimensional item response theory, testlet, two-tier model.

1. Introduction

The bifactor method (Holzinger and Swineford (1937)) for factor analysis is a constrained factor analytic model that assumes the responses to a set of test items can be accounted for by $(G+1)$ uncorrelated latent dimensions, with one primary dimension, assessed throughout the test, and G secondary "group" dimensions. It further constrains each item to have a nonzero loading on only one of the G secondary dimensions. The bifactor method for factor analysis (henceforth referred to as the linear bifactor model) was originally developed for continuous indicators. However, it has been extended to bifactor item response models (e.g., Gibbons and Hedeker (1992); Gibbons et al. (2007); Cai, Yang and Hansen (2011)) for dichotomous, ordinal, and nominal item responses by introducing link functions, such as a probit link. The assumption of orthogonality among the secondary dimensions has also been relaxed in the extended bifactor

Corresponding author: Susu Zhang, Departments of Psychology and Statistics, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA. E-mail: szhan105@illinois.edu.

model (e.g., Jennrich and Bentler (2012); Jeon, Rijmen and Rabe-Hesketh (2013)) to allow for covariance between secondary dimensions unexplained by the primary dimension. The bifactor model with one primary dimension has been extended further to the two-tier model (Cai (2010)) with $L \geq 1$ primary dimensions and G secondary dimensions, with each item measuring up to one secondary dimension, and the secondary dimensions being independent of the primary ones.

The bifactor model and its extensions have demonstrated significant practical merits in educational and psychological assessments. In contrast to uni- or low-dimensional latent trait models, they can accommodate the local dependence among clusters of items that measure the same subdimensions, and can produce subdimension trait estimates. Compared with general multidimensional latent variable models, they not only allow for the production of overall score(s), but also remarkably reduce the computational burden of high-dimensional latent trait model estimation. The bifactor model and its extensions have hence been applied to hundreds of cognitive and psychological assessments. These include psychiatric screenings that cover various domains of clinical disorders (e.g., Gibbons, Rush and Immekus (2009)), personality instruments that examine multiple facets of the same trait (e.g., Chen et al. (2012)), intelligence batteries with multiple subscales (e.g., Gignac and Watkins (2013)), and patient-reported outcome measures with broad situational representations (e.g., Reise, Morizot and Hays (2007)). In educational testing, the bifactor model and its variants have seen wide applications to assessments that involve testlets, that is, multiple questions that originate from the same stem (e.g., passage Bradlow, Wainer and Wang (1999); DeMars (2006, 2012); Jeon, Rijmen and Rabe-Hesketh (2013); Rijmen (2010)). In longitudinal assessments with repeated administrations of the same item, bifactor and two-tier models can account for the within-person dependence of responses to the same item across time points (see, Cai et al. (2016)). Parameter estimation for the bifactor and two-tier models has been implemented in many commercial and open-source statistical software programs. The bifactor model is also robust in practice as it tends to fit any data set better than other confirmatory models, regardless of the population's true models (Caspi et al. (2014)). Thorough introductions to the bifactor model and its generalizations can be found in Reise (2012) and Cai, Yang and Hansen (2011).

Identifiability is a key issue in any type of latent variable modeling (Allman, Matias and Rhodes (2009); Xu and Zhang (2016); Xu (2017); Gu and Xu (2019); Chen et al. (2015); Chen, Li and Zhang (2019)). Intuitively, a model is identifiable if distinct parameter values produce unique probability distributions of observed responses. Model identifiability is a necessary condition for consistency of parameter estimation and valid statistical inference. Without any additional requirements, the bifactor model has been shown to be nonidentifiable: Under the linear bifactor model, Green and Yang (2018) showed that two distinct sets of model parameters could produce the same model-implied covariance matrix. Eid et al. (2018) showed that nonidentifiability can arise in structural equation models with the bifactor measurement model. The current study addresses the identifiability issue of the bifactor model and its extensions by providing a relatively complete theory. We obtain the sufficient and necessary conditions for the identifiability of the standard bifactor model with both continuous and binary responses. Furthermore, we give the necessary conditions for the extended bifactor model identifiability and sufficient conditions for the identifiability of the extended bifactor and two-tier models. For dichotomous responses, the discussion is limited to probit item response models, because the theoretical identifiability of logistic item response models with multivariate-normally distributed latent traits is a more complex issue. The identifiability of the aforementioned models can be achieved using the constraints on the loading structure of the items onto the general and specific dimensions. These results provide practitioners with a viable means of examining the identifiability of a certain test using a set of easily checkable conditions.

A number of simple rules for checking confirmatory factor model identification have been proposed in prior studies (e.g., Bollen (1989); Reilly (1995); Reilly and O'Brien (1996)). Among them, the t-rule, which requires that the number of unknown model parameters not exceed the number of unique covariance terms, provides a necessary, but not sufficient condition for model identification. Another set of well-known rules are the three-indicator and two-indicator rules (see Bollen (1989)), which are sufficient for identification, but require a simple factor loading structure, and thus are not applicable to bifactor models. Empirical tests for local identification based on the information or the Jacobian matrix have been implemented in factor model estimation programs such as LISREL (Jöreskog and Sörbom (1993)). However, the local identifiability of the parameters within the neighborhood of the estimates does not guarantee global identifiability in the entire parameter space. For many general factor models in which no simple rule is applicable, checking global identification requires solving a system of equations for observed and model-implied means and covariances to determine whether the solution to each parameter is unique, either manually or with the aid of computer algebra systems (Bollen and Bauldry (2010); Kenny and Milan (2012)). Owing to the special structure of the bifactor model, simple and checkable sufficient and necessary conditions can be developed. Readers are referred to Bollen (1989), Bollen and Bauldry (2010), and Jöreskog and Sörbom (1993) for introductions to common methods for checking general factor model identification.

The rest of the paper is organized as follows. Section 2 presents the identifiability results on the linear bifactor, extended bifactor, and two-tier models. Section 3 extends the theoretical results on the three models to dichotomous responses with the probit link. Section 4 discusses the connections between the new results and the existing literature on bifactor identifiability. A discussion of the findings is provided in Section 5. In the Supplementary Material, simulation studies are designed to verify the theoretical identifiability results by examining the estimation consistency under several identifiable and nonidentifiable loading structures. The Supplementary Material also contains proofs of the main theoretical results, as well as detailed examples.

2. Linear Bifactor Model and Extensions

This section presents the results on the identifiability of the linear standard bifactor, linear extended bifactor, and linear two-tier models, which assume that the response to each item is normally distributed, with mean equal to an intercept plus a linear combination of the latent factor scores. This class of models is thus suitable for continuous observed indicators.

It is worth introducing the concept of identifiability in general before moving on to specific models. In mathematical terms, a statistical model may be specified by a pair $(\mathcal{S}, \mathcal{P})$, where $\mathcal{S}$ is the set of possible observations or the sample space, and $\mathcal{P}$ is a set of probability distributions on $\mathcal{S}$, which is parameterized as $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$. The set Θ defines the parameter space. We say the model parameter θ^* is identifiable (or the model is identifiable at θ^*) if $\mathcal{P}_\theta(y) = \mathcal{P}_{\theta^*}(y)$, for all $y \in \mathcal{S}$, implies $\theta = \theta^*$. Essentially, identifiability implies that the underlying distribution of the observed data cannot admit two distinct sets of parameter values.

2.1. Standard bifactor model

The standard bifactor model (Holzinger and Swineford (1937)) assumes that the response to each item in a test can be explained by one general factor, which runs through the test, and up to one group factor, which runs through a subset of items. Without loss of generality, we refer to a subset of items that load on the same group factor as a testlet (e.g., Bradlow, Wainer and Wang (1999); DeMars (2006)), but the model formulation is equally applicable to psychological assessments with subdimensions and cognitive batteries with subtests. Specifically, consider a test with J items that can be partitioned into G testlets, where the gth testlet consists of J_g items and $\sum_g J_g = J$. Let $\mathcal{B}_g$ denote the set of items in the gth testlet. Under the standard linear bifactor model, the response to item $j \in \{1, \ldots, J\}$ in testlet $g_j \in \{1, \ldots, G\}$, Y_j, is given by

$$Y_j = d_j + a_{j0}\eta_0 + \sum_{g=1}^{G} a_{jg}\eta_g + \epsilon_j, \tag{2.1}$$

where η_0 is the respondent's latent score on the general factor, η_g is the latent score on the gth group factor, a_{jk} is item j's loading on the kth latent dimension, d_j is the item intercept, and ϵ_j is the random error unexplained by the latent factors. Across all items, ϵ_j is assumed to be independently distributed with mean zero and variance λ_j; that is, $\epsilon_j \sim N(0, \lambda_j)$. In addition, for item j in testlet g_j, it is assumed that $a_{jg} = 0$ for all $g \neq g_j$; in other words, the loadings of item j on all other group factors are restricted to zero. Thus, Equation (2.1) simplifies to

$$Y_j = d_j + a_{j0}\eta_0 + a_{jg_j}\eta_{g_j} + \epsilon_j. \tag{2.2}$$

Let $\boldsymbol{\eta} = (\eta_0, \eta_1, \ldots, \eta_G)^T$ denote a vector of the latent traits of a respondent, which is assumed to follow a multivariate normal distribution with zero mean and covariance $\boldsymbol{\Sigma}$; that is,

$$\boldsymbol{\eta} \sim \mathrm{MVN}(\mathbf{0}, \boldsymbol{\Sigma}). \tag{2.3}$$

The standard bifactor model further assumes that all general and group factors are independent; in other words, $\boldsymbol{\Sigma} = \boldsymbol{I}_{(1+G),(I+G)}$, where $\boldsymbol{I}_{(1+G),(I+G)}$ is the $(1+G) \times (1+G)$ identity matrix. The mean and standard deviation of each latent factor are fixed to zero and one, respectively, to resolve the location and scale indeterminacy of the latent dimensions.

Let $\boldsymbol{A} = [\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_G]$ denote the $J \times (1+G)$ matrix of factor loadings, where the first column $\mathbf{a}_0 = (a_{10}, \ldots, a_{J0})^T$ is the items' loadings on the general factor, and the subsequent G columns $(\mathbf{a}_1, \ldots, \mathbf{a}_G)$ are the loadings on each of the G testlets. Note that $\boldsymbol{A}$ is a sparse matrix, with most of the testlet-specific loadings restricted to zero by the single group factor loading assumption. Let $\mathbf{d} = (d_1, \ldots, d_J)^T$ denote a length-J vector of item intercepts, and let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_J)$ denote a vector of item unique variances. To resolve the sign indeterminacy of the latent factors, we further assume, without loss of generality, that for each factor, the first item that loads on the factor has a positive loading. Under the standard bifactor model, the parameter space $\mathfrak{P}$ is then given by $\mathfrak{P} = \{(A, \mathbf{d}, \boldsymbol{\lambda}) \mid \text{The sign of first nonzero element in every column of } A \text{ is positive}\}$. The identifiability of the standard bifactor model is defined as follows.

Definition 1. We say a linear bifactor model is identifiable at $(A, \mathbf{d}, \boldsymbol{\lambda})$ if, for any other set of parameters $(A^{'}, \mathbf{d}^{'}, \boldsymbol{\lambda}^{'})$ that defines the same probability distribution, it must hold that

$$(A, \mathbf{d}, \boldsymbol{\lambda}) = (A^{'}, \mathbf{d}^{'}, \boldsymbol{\lambda}^{'}). \tag{2.4}$$

Before stating the identifiability results on the standard bifactor model, we introduce some additional notation. Let $\bar{A}_g$ be the submatrix of A corresponding to the items in testlet g (i.e., B_g); that is, $\bar{A}_g = (\mathbf{a}_0[\mathcal{B}_g], \mathbf{a}_g[\mathcal{B}_g])$, where $x[\mathcal{B}_g]$ denotes a subvector with entries in the set $\mathcal{B}_g$. Furthermore, let $A[\mathcal{B}, :]$ denote the submatrix of A consisting only of rows in some generic set $\mathcal{B}$. We define the following subsets.

- $\mathcal{H}_1 = \{g \mid \mathbf{a}_0[\mathcal{B}_g] \neq \mathbf{0}\}$: the set of testlets in {1,..., G} with nonvanishing main factors; that is, at least one item in the testlet has a nonzero true loading on the main dimension, η_0;
- $\mathcal{Q}_g = \{j \mid \mathbf{a}_g[j] \neq 0\}$: the set of items in the gth testlet with nonzero true loadings on the testlet-specific factor;
- $\mathcal{H}_2 = \{g \mid$ there exists a partition of $\mathcal{B}_g, i.e., \mathcal{B}_g = \mathcal{B}_{g,1} \cup \mathcal{B}_{g,2},\ \mathcal{B}_{g,1} \cap \mathcal{B}_{g,2} = \emptyset,$ such that $\bar{A}_g[\mathcal{B}_{g,1}, :], \bar{A}_g[\mathcal{B}_{g,2}, :]$ are of full column rank.$\}$: the set of testlets that can be partitioned into two disjoint subsets of items, where each submatrix of the (main and testlet) factor loadings has full column rank.

The following theorem characterizes the sufficient and necessary conditions for the identifiability of the linear standard bifactor model.

Theorem 1. *Under the standard linear bifactor model, the model parameters are identifiable if and only if they satisfy one of the following conditions:*

$P1$ $|\mathcal{Q}_g| \geq 3$, *for all* $g = 1, \ldots, G$; $|\mathcal{H}_1| \geq 3$.

$P2$ $|\mathcal{Q}_g| \geq 3$, *for all* $g = 1, \ldots, G$; $|\mathcal{H}_1| = 2$; $|\mathcal{H}_2| \geq 1$.

Theorem 1 gives the minimum requirements for the identifiability of the standard bifactor model. Specifically, it requires that the test contains at least two testlets, each containing at least three items. In addition, if there are only two testlets, Theorem 1 requires that one of them can be partitioned into two disjoint subsets of items, such that both subsets have linearly independent primary and testlet-specific factor loadings.

2.2. Extended bifactor model

The extended bifactor model, also known as the oblique bifactor model (Jennrich and Bentler (2012)), relaxes the assumption of independence between the secondary dimensions. Instead of restricting the latent covariance matrix, $\mathbf{\Sigma}$, to be the identity matrix, the extended bifactor model allows the covariance between the latent dimensions to take the form of $\mathbf{\Sigma} = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{\Sigma}_G \end{pmatrix}$, where the covariance matrix for the testlet dimensions, $\mathbf{\Sigma}_G$, is positive definite, with all diagonal elements being one and no additional restriction on the off-diagonal elements. Note that the covariances between the primary dimension and each testlet dimension are still restricted to be zero.

Under the extended bifactor model, in addition to the item intercepts, loadings, and unique variances, the latent covariance matrix needs to be estimated. Let $\mathfrak{P} = \{(A, \mathbf{d}, \Sigma_G, \boldsymbol{\lambda}) \mid$ first nonzero element in every column of A is positive, $\mathrm{diag}(\Sigma_G) = 1, \Sigma_G$ positive definite$\}$ denote the parameter space of the extended bifactor model. Then, the identifiability of the extended bifactor model is defined as follows.

Definition 2. We say a linear extended bifactor model is identifiable at $(A, \mathbf{d}, \Sigma_G, \boldsymbol{\lambda})$ if, for any other set of parameters $(A', \mathbf{d}', \Sigma_G', \boldsymbol{\lambda}')$ that defines the same probability distribution,

$$(A, \mathbf{d}, \Sigma_G, \boldsymbol{\lambda}) = (A', \mathbf{d}', \Sigma_G', \boldsymbol{\lambda}') \tag{2.5}$$

must hold.

In addition to $\mathcal{H}_2$ and $\mathcal{Q}_g$ in the standard bifactor model identifiability results, we introduce another set that is key to the identifiability of the extended bifactor model:

- $\mathcal{H}_3 = \{g \mid \bar{A}_g$ has column rank $2\}$: the set of testlets with full main factor and testlet-specific factor information; that is, columns $\mathbf{a}_0[\mathcal{B}_g]$ and $\mathbf{a}_g[\mathcal{B}_g]$ are linearly independent.

Theorem 2 gives two sets of sufficient conditions for the identifiability of the extended linear bifactor model.

Theorem 2. *Under the linear extended bifactor model, the model parameters are identifiable if one of the following sets of requirements is satisfied:*

E1S $|\mathcal{Q}_g| \geq 3$, *for all* $g = 1, \ldots, G$; $|\mathcal{H}_3| \geq 3$.

E2S $|\mathcal{Q}_g| \geq 3$, *for all* $g = 1, \ldots, G$; $|\mathcal{H}_3| = 2$; $|\mathcal{H}_2| \geq 1$.

The sufficient conditions in Theorem 2 are very similar to the sufficient and necessary conditions for the standard bifactor model in Theorem 1, where $\mathcal{H}_3$ is the counterpart to $\mathcal{H}_1$. Unlike $\mathcal{H}_1$, which contains testlets with a nonzero main-factor loading vector, $\mathcal{H}_3$ further requires the main- and testlet-factor loading vectors to be linearly independent. Note that $E1S$ and $E2S$ are sufficient for the identifiability of the extended bifactor model, but they are not necessary. Theorem 3 provides the necessary, but not sufficient conditions, that is, the minimum conditions that need to be met.

Theorem 3. *Under the linear extended bifactor model, the model parameters are identifiable only if both conditions below are satisfied:*

$E1N$ $|\mathcal{Q}_g| \geq 2$, *for all* $g = 1, \ldots, G$; *and* $|\mathcal{Q}_g| \geq 3$, *for* g *such that* $\Sigma_G[g, -g] = \mathbf{0}$.

$E2N$ $|\mathcal{H}_3| \geq 2$.

Essentially, at least two testlets with linearly independent main- and testlet-factor loadings are required. In addition, each subdimension should be measured by at least two items, and if a subdimension is uncorrelated with others, at least three items are required, as before. The nonzero correlation between the testlet factors provides additional information on the testlet-specific loadings, reducing the number of required items per testlet to two for those testlets with nonzero correlations with others.

Note that $E1N$ and $E2N$ together are not enough for the identifiability of the extended bifactor model. Additional requirements are needed to render parameter identifiability. Theorem 2 gives one way to impose such additional requirements. The requirement of $|\mathcal{H}_2| \geq 1$ in $E2S$ may be replaced by other requirements. See the following proposition.

Proposition 1. *Under the linear extended bifactor model, the model parameters are identifiable if* $E1N$ *and* $E2N$ *are satisfied, and*

$E3S$ *There exist* $g_1, g_2 \in \mathcal{H}_3$, *such that* (1) $\Sigma_G[g_1, g_2] \neq 0$, (2) $|\mathcal{Q}_{g_1}| \geq 3, |\mathcal{Q}_{g_2}| \geq 3$, *and* (3) *the Kruskal rank of* $\bar{A}_{g_1}^T$ *is two. (A matrix* A *has Kruskal rank (Kruskal* (1977)) R *if any* R *columns of* A *are linearly independent).*

Remark 1. The gap between the necessary and sufficient conditions in Theorems 2 and 3 is that $|\mathcal{H}_3| \geq 2$ itself cannot guarantee identifiability. Either more testlets ($|\mathcal{H}_3| \geq 3$) are needed, or at least one of the testlets needs to be "strong" ($|\mathcal{H}_2| \geq 1$). Nonzero correlations (Σ) between testlet factors increase the complexity of the identifiability problem compared with the uncorrelated case.

Remark 2. The conditions in Theorems 1-3 are easy to check in practice. We need only to do simple algebras (i.e., counting the number of nonzero entries, computing the column rank, etc.) on the estimated loading matrix A and covariance matrix Σ. Additionally, $\mathcal{H}_2 \geq 1$ in Condition P2 generically holds when a testlet contains four or more items.

Example 1. Consider an extended bifactor model with three testlets, where testlet 1 has only two items. Suppose the true A and Σ_G are given by

$$A = \begin{pmatrix} a_{10} & a_{11} & 0 & 0 \\ a_{20} & a_{21} & 0 & 0 \\ a_{30} & 0 & a_{32} & 0 \\ a_{40} & 0 & a_{42} & 0 \\ a_{50} & 0 & a_{52} & 0 \\ a_{60} & 0 & a_{62} & 0 \\ a_{70} & 0 & 0 & a_{73} \\ a_{80} & 0 & 0 & a_{83} \\ a_{90} & 0 & 0 & a_{93} \end{pmatrix}, \qquad \Sigma_G = \begin{pmatrix} 1 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & 1 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & 1 \end{pmatrix}, \tag{2.6}$$

respectively, where $\sigma_{12}, \sigma_{13}, \sigma_{23} \neq 0$, and any A_g $(g = 1, 2, 3)$ has two linearly independent columns. According to Proposition 1, the model parameter is identifiable, even though $|\mathcal{Q}_1| = 2$.

2.3. Two-tier model

The two-tier model (Cai (2010)) extends the standard bifactor model by allowing for more than one primary dimension. Consider a test that measures L primary factors and G group factors. Under the two-tier model, denote the latent factors by $\boldsymbol{\eta} = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2)$; with $\boldsymbol{\eta}_1 = (\eta_1, \ldots, \eta_L)^T$ and $\boldsymbol{\eta}_2 = (\eta_{L+1}, \ldots, \eta_{L+G})^T$. The response to the jth item in testlet g_j is given by

$$Y_j = d_j + \sum_{l=1}^{L} a_{jl}\eta_l + \sum_{g=L+1}^{L+G} a_{jg}\eta_g + \epsilon_j. \tag{2.7}$$

Similarly to the bifactor model, ϵ_j is independent and normally distributed with $\epsilon_j \sim N(0, \lambda_j)$, and only the g_jth testlet factor loading is nonzero for item j; that is, $a_{jg} = 0$, for all $g \neq g_j$. The latent covariance matrix of $\boldsymbol{\eta}$ takes the form $\begin{pmatrix} \Sigma_L & \mathbf{0}^T \\ \mathbf{0}^T & I_{G\times G} \end{pmatrix}$, where Σ_L is an $L \times L$ positive-definite matrix with diagonal elements of one and no additional restriction on the off-diagonal elements.

Let $\mathfrak{P} = \{(A, \mathbf{d}, \Sigma_L, \boldsymbol{\lambda}) \mid$ first nonzero element in every column of A is positive, $\mathrm{diag}(\Sigma_L) = 1, \Sigma_L$ positive definite$\}$ represent the model space of the two-tier model. Furthermore, let $\mathbf{I}$ denote the configuration mapping function, with $\mathbf{I}(X) = \tilde{X}$, where $\tilde{X}_{ij} = \mathbf{1}\{X_{ij} \neq 0\}$ for an arbitrary matrix X. The two-tier model identifiability is defined as follows.

Definition 3. A linear two-tier model is identifiable at $(A, \mathbf{d}, \Sigma_L, \boldsymbol{\lambda})$ if, for any other set of parameters $(A^{'}, \mathbf{d}^{'}, \Sigma_L^{'}, \boldsymbol{\lambda}^{'})$ that defines the same distribution and satisfies $\mathbf{I}(A) = \mathbf{I}(A^{'})$, it must hold that $(A, \mathbf{d}, \Sigma_L, \boldsymbol{\lambda}) = (A^{'}, \mathbf{d}^{'}, \Sigma_L^{'}, \boldsymbol{\lambda}^{'})$.

Here, note that a factor model is only identifiable up to some rotation. In the definition, the requirement for equal factor loading configurations, $\mathbf{I}(A^{'}) = \mathbf{I}(A)$, is included to resolve the rotational indeterminacy.Note too that the identifiability of the two-tier model is nontrivial, in the sense that the model could fail to be identifiable, even if (1) the loading matrix of main factors, $A_{:,1:L}$, satisfies the usual identifiability conditions for multivariate factor models, and (2) the testlets satisfy the identifiability conditions for the bifactor model. See the examples provided in the Supplementary Material.

The proposed sufficient conditions for two-tier model identifiability build upon the sufficient conditions for unique variance identifiability under factor models with uncorrelated errors, which can be found in Theorem 5.1 of Anderson and Rubin (1956), and is rephrased below.

Theorem 4. *(Anderson and Rubin (1956)). Consider a general factor model with implied covariance matrix $\Psi = A\Sigma A^T + \Lambda$, where the item error covariance matrix, Λ, is diagonal with $\mathrm{diag}(\Lambda) = \boldsymbol{\lambda}$. Then, $A\Sigma A^T$ and Λ are identifiable if the following holds:*

$C0$ If any row of A is deleted, there remain two disjoint submatrices of A with full column rank.

Under the two-tier model, let $\bar{A}_g$ be the submatrix of A corresponding to items in testlet g; that is, $\bar{A}_g = (A[\mathcal{B}_g, 1 : L], A[\mathcal{B}_g, L + g])$. Similarly, for a subset of testlets $\mathcal{G}_1 \subseteq \{1, \ldots, G\}$, denote the submatrix of A corresponding to testlets in $\mathcal{G}_1$ by $\bar{A}_{\mathcal{G}_1} = (A[\mathcal{B}_{\mathcal{G}_1}, 1 : L], A[\mathcal{B}_{\mathcal{G}_1}, L + \mathcal{G}_1])$, where $\mathcal{B}_{\mathcal{G}_1} = \bigcup_{g \in G_1} \mathcal{B}_g$ and $L + \mathcal{G}_1 = \{L + g^* \mid g^* \in \mathcal{G}_1\}$. We introduce two sets, $\mathcal{H}_4$ and $\mathcal{H}_5$, that are essential to the identifiability results for the two-tier model:

- $\mathcal{H}_4 = \{g \mid A[\mathcal{B}_g, 1 : L]$ is of full column rank$\}$: the set of testlets with non-degenerate main factor information; that is, the L columns corresponding to the main factor loadings are linearly independent;

- $\mathcal{H}_5 = \{g \mid \bar{A}_g$ is of full column rank$\}$: the set of testlets with nondegenerate all-factor information, that is, with linearly independent main factor and testlet factor loadings.

The following theorem provides the sufficient conditions for the identifiability of the linear two-tier model.

Theorem 5. *Under the linear two-tier model, if the true loading matrix, A, satisfies Condition $C0$ and one of Conditions $T1S$-$T3S$, then the parameters are identifiable.*

$T1S$ $|\mathcal{H}_4| \geq 3$, *$A[\mathcal{B}_{\mathcal{H}_4}, 1:L]$ contains an identity, where $\mathcal{B}_{\mathcal{H}_4}$ is the set of items that makes up the testlets in $\mathcal{H}_4$.*

$T2S$ $|\mathcal{H}_4| \geq 2$, $|\mathcal{H}_5| \geq 1$, *$A[\mathcal{B}_{\mathcal{H}_4}, 1:L]$ contains an identity.*

$T3S$ *A[:,1:L] contains an identity, and there exists a partition of testlets $\{1, \ldots, G\} = \mathcal{G}_1 \dot{\cup} \mathcal{G}_2$, such that (1) $\bar{A}_{\mathcal{G}_1}$ has full column rank, and (2) $A[\mathcal{B}_{\mathcal{G}_2}, 1:L]$ has full column rank.*

In other words, the linear two-tier model is identifiable if the following two conditions are simultaneously met:

- After removing any row of the loading matrix A, the remaining rows of A can be partitioned into two disjoint submatrices, both of which contain $L + G$ linearly independent columns.
- One of the following is satisfied:

$T1S$: (1) The test contains at least three testlets with linearly independent main factor loadings, and (2) within these testlets satisfying (1), for each main factor, there exists at least one item that exclusively measures this main factor (i.e., having nonzero loadings only on this main factor and possibly the testlet factor);

$T2S$: (1) The test contains two testlets with linearly independent main factor loadings, (2) within these testlets satisfying (1), for each main factor, there exists at least one item that exclusively measures this main factor, and (3) at least one of the testlets satisfying (1) has linearly independent main *and* testlet factor loadings;

$T3S$: (1) For each main factor, at least one item in the test exclusively measures that main factor (aside from the testlet factor), and (2) the set of all testlets can be partitioned to two disjoint subsets, $\mathcal{G}_1$ and $\mathcal{G}_2$, such

that (a) the loading matrix corresponding to the first subset of testlets, $\bar{A}_{\mathcal{G}_1}$, has $L + |G_1|$ linearly independent columns, and (b) for the second subset, the columns of the main factor loadings, $A[\mathcal{B}_{\mathcal{G}_2}, 1 : L]$, are linearly independent.

3. Extensions to Dichotomous Responses

Notwithstanding the wide application of linear factor models in the social science literature, a large proportion of educational and psychological assessments consist of items with dichotomous responses. For instance, these may include cognitive questions where an examinee responds either correctly (1) or incorrectly (0), or clinical screening questions where a participant either exhibits certain behavior (1) or not (0). Probit item response models, also known as normal ogive models (e.g., Thurstone (1927); Lawley (1943); Lord (1952); Christoffersson (1975)), have been widely adopted for dichotomous responses. In general, consider a test of J items. The responses to the J items are assumed to be locally independent, given the respondents' latent traits, $\boldsymbol{\eta}$, and the probability of responding "1" to the jth item is given by

$$P(Y_j = 1|\boldsymbol{\eta}) = P_\epsilon(\epsilon_j \leq d_j + \mathbf{a}_j^T\boldsymbol{\eta}) = \Phi(d_j + \mathbf{a}_j^T\boldsymbol{\eta}), \tag{3.1}$$

where $\epsilon_j \sim N(0, 1)$, $\Phi(\cdot)$ is the standard normal cumulative distribution function (i.e., the probit link), and $\boldsymbol{\eta} \sim N(0, \Sigma), d_j$, and $\mathbf{a}_j$ are the person's latent traits, item intercept, and item slopes/loadings, respectively. As before, let A denote the matrix of factor loadings.

Item bifactor models and extensions have also been proposed to accommodate dichotomous response tests with underlying bifactor-like latent structures. Examples of such models include the item bifactor model (Gibbons and Hedeker (1992)), extended item bifactor model (Jeon, Rijmen and Rabe-Hesketh (2013)), and two-tier item factor model (Cai (2010)). The identifiability results for the linear bifactor model and its extensions do not directly apply to dichotomous item bifactor-type models, owing to different parameterizations and forms of observed data. This section presents the results on the identifiability of the dichotomous bifactor model, extended bifactor model, and two-tier model with probit links.

Before introducing the identifiability conditions for each of the specific models, it is worth mentioning a few identities on the items' first, second, and kth moments implied by the general probit item factor model in (3.1), as well as their relationships with the thresholds and tetrachoric correlations (Pearson (1900)) under the probit model. Unlike linear models, we do not directly observe the mean and covariance matrix implied by the linear component (i.e., $d_j + \mathbf{a}_j^T \boldsymbol{\eta}_j$) under the probit model. However, the threshold and tetrachoric correlations can be identified (Kendall and Stuart (1958)) and estimated using various approximation methods (e.g., Castellan (1966); Olsson (1979)). In the following, we explain how the tetrachoric correlations relate to the identifiability problem.

Let ξ denote a standard normal random variable. Note that, at the population level (i.e., for a randomly chosen $\boldsymbol{\eta} \sim MVN(0, \Sigma)$), the marginal probability of observing a response of "1" on item j is given by

$$\begin{aligned} P(Y_j = 1) &= \mathbb{E}_{\boldsymbol{\eta}} P(Y_j = 1|\boldsymbol{\eta}) = \mathbb{E}_{\boldsymbol{\eta}} \mathbb{E}_{\epsilon_j} [\mathbf{1}\{\epsilon_j \leq d_j + \mathbf{a}_j^T \boldsymbol{\eta}\}|\boldsymbol{\eta}] = P(\epsilon_j \leq d_j + \mathbf{a}_j^T \boldsymbol{\eta}) \\ &= P\Big(d_j + \sqrt{\mathbf{a}_j^T \Sigma \mathbf{a}_j + 1}\xi \geq 0\Big) = 1 - \Phi\left(- \frac{d_j}{\sqrt{\mathbf{a}_j^T \Sigma \mathbf{a}_j + 1}}\right). \end{aligned} \tag{3.2}$$

In addition, the probability that the responses to items j_1 and j_2 are both "1" is

$$\begin{aligned} &P(Y_{j_1} = 1, Y_{j_2} = 1) = \mathbb{E}_{\boldsymbol{\eta}} P(Y_{j_1} = 1, Y_{j_2} = 1|\boldsymbol{\eta}) \\ &= P(\epsilon_{j_1} \leq d_{j_1} + \mathbf{a}_{j_1}^T \boldsymbol{\eta}, \epsilon_{j_2} \leq d_{j_2} + \mathbf{a}_{j_2}^T \boldsymbol{\eta}) \\ &= P\Big(d_{j_1} + \sqrt{\mathbf{a}_{j_1}^T \Sigma \mathbf{a}_{j_1} + 1}\xi_{j_1} \geq 0, d_{j_2} + \sqrt{\mathbf{a}_{j_2}^T \Sigma \mathbf{a}_{j_2} + 1}\xi_{j_2} \geq 0\Big) \\ &= \Phi_2\left(- \frac{d_{j_1}}{\sqrt{\mathbf{a}_{j_1}^T \Sigma \mathbf{a}_{j_1} + 1}}, - \frac{d_{j_2}}{\sqrt{\mathbf{a}_{j_2}^T \Sigma \mathbf{a}_{j_2} + 1}}, \frac{\mathbf{a}_{j_1}^T \Sigma \mathbf{a}_{j_2}}{\sqrt{\mathbf{a}_{j_1}^T \Sigma \mathbf{a}_{j_1} + 1}\sqrt{\mathbf{a}_{j_2}^T \Sigma \mathbf{a}_{j_2} + 1}}\right), \end{aligned} \tag{3.3}$$

where $\Phi_2(a, b, \rho) = \mathbb{E}(X_1 \geq a, X_2 \geq b)$, $X_1, X_2 \sim N(0, 1)$ and $\mathrm{corr}(X_1, X_2) = \rho$. Here, the a, b, and ρ are commonly referred to as the thresholds and the tetrachoric correlation under the probit framework. The probability of responding "1" simultaneously on k items $(j_1, \ldots, j_k)$ is given by

$$\begin{aligned} &P(Y_{j_1} = 1, \ldots, Y_{j_k} = 1) = \mathbb{E}_{\boldsymbol{\eta}} P(Y_{j_1} = 1, \ldots Y_{j_k} = 1|\boldsymbol{\eta}) \\ &= P(\epsilon_{j_1} \leq d_{j_1} + \mathbf{a}_{j_1}^T \boldsymbol{\eta}, \ldots, \epsilon_{j_k} \leq d_{j_k} + \mathbf{a}_{j_k}^T \boldsymbol{\eta}) \\ &= P\Big(d_{j_1} + \sqrt{\mathbf{a}_{j_1}^T \Sigma \mathbf{a}_{j_1} + 1}\xi_{j_1} \geq 0, \ldots, d_{j_k} + \sqrt{\mathbf{a}_{j_k}^T \Sigma \mathbf{a}_{j_k} + 1}\xi_{j_k} \geq 0\Big) \\ &= \Phi_k\left(- \frac{d_{j_1}}{\sqrt{\mathbf{a}_{j_1}^T \Sigma \mathbf{a}_{j_1} + 1}}, \ldots, - \frac{d_{j_2}}{\sqrt{\mathbf{a}_{j_2}^T \Sigma \mathbf{a}_{j_2} + 1}}, C_\rho\right), \end{aligned} \tag{3.4}$$

with tetrachoric correlation matrix $C_\rho[j_1, j_2] = \mathbf{a}_{j_1}^T \Sigma \mathbf{a}_{j_2} / \sqrt{\mathbf{a}_{j_1}^T \Sigma \mathbf{a}_{j_1} + 1}\sqrt{\mathbf{a}_{j_2}^T \Sigma \mathbf{a}_{j_2} + 1}$ and $C_\rho[j, j] = 1$. Here, $\Phi_k(a_1, \ldots, a_k, C_\rho) = \mathbb{E}(X_1 \geq a_1, X_k \geq a_k)$, $X_1, \ldots, X_k \sim N(0, 1)$ and $\mathrm{corr}(X_{k_1}, X_{k_2}) = C_\rho[k_1, k_2]$.

In the following, we show that threshold and tetrachoric correlations provide full information on probit binary item responses.

Proposition 2. *Two sets of parameters define the same model if and only if their thresholds and tetrachoric correlations are equal; that is,*

$$\frac{d_j}{\sqrt{\mathbf{a}_j^T\Sigma\mathbf{a}_j+1}}=\frac{d_j^{'}}{\sqrt{(\mathbf{a}_j^{'})^T\Sigma^{'}\mathbf{a}_j^{'}+1}}\quad \forall j,$$

and

$$\frac{\mathbf{a}_{j_1}^T\Sigma\mathbf{a}_{j_2}}{\sqrt{\mathbf{a}_{j_1}^T\Sigma\mathbf{a}_{j_1}+1}\sqrt{\mathbf{a}_{j_2}^T\Sigma\mathbf{a}_{j_2}+1}}=\frac{(\mathbf{a}_{j_1}^{'})^T\Sigma^{'}\mathbf{a}_{j_2}^{'}}{\sqrt{(\mathbf{a}_{j_1}^{'})^T\Sigma^{'}\mathbf{a}_{j_1}^{'}+1}\sqrt{(\mathbf{a}_{j_2}^{'})^T\Sigma^{'}\mathbf{a}_{j_2}^{'}+1}}\quad \forall j_1\neq j_2.$$

It follows from the above proposition that checking the identifiability of probit bifactor models reduces to checking whether the probit threshold and tetrachoric correlations admit only one set of parameters. In other words, the probit bifactor models can be identified if $(d_j,\mathbf{a}_j)$ can be identified based on the thresholds (i.e., $d_j/(\mathbf{a}_j^T\Sigma\mathbf{a}_j+1)^{1/2},\forall j$) and the pairwise tetrachoric correlations (i.e., $(\mathbf{a}_{j_1}^T\Sigma\mathbf{a}_{j_2})/\big((\mathbf{a}_{j_1}^T\Sigma\mathbf{a}_{j_1}+1)(\mathbf{a}_{j_2}^T\Sigma\mathbf{a}_{j_2}+1)\big)^{1/2},\forall j_1\neq j_2$). The theoretical results on the sufficient conditions turned out to be very similar to those under the linear bifactor model and its extensions.

3.1. Standard bifactor model

Adopting the same notation as for the linear bifactor model, under the probit bifactor model, the probability of a response of "1" on item $j\in\{1,\ldots,J\}$ in testlet g_j is given by

$$P(Y_j=1|\eta_0,\eta_1,\ldots,\eta_G)=\Phi\left(d_j+a_0\eta_0+\sum_{g=1}^{G}a_{jg}\eta_g\right)=\Phi\left(d_j+a_0\eta_0+a_{jg_j}\eta_{g_j}\right),\tag{3.5}$$

where, similarly to the linear case, $a_{jg}=0$ for all $g\neq g_j$, and $\boldsymbol{\eta}=(\eta_0,\eta_1,\ldots,\eta_G)^T \sim MVN(\mathbf{0},\Sigma)$, with $\Sigma=\boldsymbol{I}_{(1+G)\times(1+G)}$. With A and $\mathbf{d}$ denoting the loading matrix and the vector of intercepts, respectively, the parameter space of the probit standard bifactor model is given by $\mathfrak{P}=\{(A,\mathbf{d})\mid \text{first nonzero element in every column of } A \text{ is positive}\}$. Based on this, the probit bifactor model identifiability is defined as follows.

Definition 4. We say a probit bifactor model is identifiable at $(A,\mathbf{d})$ if for any other set of parameters $(A^{'},\mathbf{d}^{'})$ that defines the same probability distribution, it must hold that

$$(A, \mathbf{d}) = (A^{'}, \mathbf{d}^{'}). \tag{3.6}$$

Adopting the same definitions as those of sets $\mathcal{H}_1, \mathcal{H}_2$, and $\mathcal{Q}_g$ in section 2.1. Theorem 6 provides the sufficient and necessary conditions for the identifiability of dichotomous bifactor models with a probit link.

Theorem 6. *Under a standard bifactor model with a probit link, the model parameter is identifiable if and only if it satisfies one of the follow conditions:*

$P1$ $|\mathcal{H}_1| \geq 3$; $|\mathcal{Q}_g| \geq 3$, *for all* $g = 1, \ldots, G$.

$P2$ $|\mathcal{H}_1| = 2$; $\mathcal{H}_2$ *is nonempty;* $|\mathcal{Q}_g| \geq 3$ *for* $g = 1, \ldots, G$.

The interpretations of $P1$ and $P2$ remain the same as for the linear bifactor model in Section 2.1. In the Supplementary Material, a few examples are provided to illustrate how the identifiability of the probit bifactor model can be checked.

3.2. Extended bifactor model

With the same item response function as the standard bifactor model, the probit extended bifactor model relaxes the assumption of $\Sigma = \boldsymbol{I}_{(1+G)\times(1+G)}$ by allowing correlations among $\eta_1, \ldots, \eta_G$. Hence, the covariance matrix for $\boldsymbol{\eta}$, Σ, takes the form of $\Sigma = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{\Sigma}_G \end{pmatrix}$, with Σ_G positive definite, with diagonal entries of one and no further restriction on the off-diagonal entries. Under the probit extended bifactor model, the parameter space is given by $\mathfrak{P} = \{(A, \mathbf{d}, \Sigma_G) \mid$ first nonzero element in every column of A is positive, $\operatorname{diag}(\Sigma_G) = 1, \Sigma_G$ positive definite.$\}$. The probit extended bifactor model identifiability is defined as follows.

Definition 5. We say a probit extended bifactor model is identifiable at $(A, \mathbf{d}, \Sigma_G)$ if, for any other set of parameters $(A^{'}, \mathbf{d}^{'}, \Sigma_G^{'})$ that define the same probability distribution,

$$(A, \mathbf{d}, \Sigma_G) = (A^{'}, \mathbf{d}^{'}, \Sigma_G^{'}) \tag{3.7}$$

must hold.

Again, it can be shown that the sufficient conditions and the necessary conditions for the linear extended bifactor model still hold when the responses become binary.

Theorem 7. *Under the probit extended bifactor model, the model parameters are identifiable if one of the following requirements are satisfied:*

E1S $|\mathcal{Q}_g| \geq 3$, *for all* $g = 1, \ldots, G$; $|\mathcal{H}_3| \geq 3$.

E2S $|\mathcal{Q}_g| \geq 3$, *for all* $g = 1, \ldots, G$; $|\mathcal{H}_3| = 2$; $|\mathcal{H}_2| \geq 1$.

Theorem 8. *Under the extended bifactor model with a probit link, the model parameters are identifiable only if both conditions below are satisfied:*

E1N $|\mathcal{Q}_g| \geq 2$, *for all* $g = 1, \ldots, G$; *and* $|\mathcal{Q}_g| \geq 3$ *for* $g : \Sigma_G[g, -g] = \mathbf{0}$.

E2N $|\mathcal{H}_3| \geq 2$.

Here, the definitions of the sets $\mathcal{Q}_g, \mathcal{H}_3$, and $\mathcal{H}_2$ and the interpretations of the conditions remain the same as those for the linear extended bifactor model in section 2.2.

3.3. Two-tier model

A two-tier probit model with J items, L main factors, and G testlets has the following item response function for a particular item j in testlet g_j:

$$P(Y_j = 1|\boldsymbol{\eta}) = \Phi\left(d_j + \sum_{l=1}^{L} a_{jl}\eta_l + \sum_{g=L+1}^{L+G} a_{jg}\eta_g\right) = \Phi\left(d_j + \sum_{l=1}^{L} a_{jl}\eta_l + a_{jg_j}\eta_{g_j}\right), \tag{3.8}$$

where $a_{jg} = 0, \forall g \neq g_j$. As in the linear two-tier model, the latent traits $\boldsymbol{\eta} = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2)$, where $\boldsymbol{\eta}_1 = (\eta_1, \ldots, \eta_L)^T$ and $\boldsymbol{\eta}_2 = (\eta_{L+1}, \ldots, \eta_{L+G})^T$, are assumed to follow a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix Σ, which takes the form of $\begin{pmatrix} \Sigma_L & \mathbf{0}^T \\ \mathbf{0}^T & I_{G\times G} \end{pmatrix}$, with Σ_L positive definite, with diagonal elements of one and off-diagonal elements between -1 and 1. Hence, the parameter space for the probit two-tier model is given by $\mathfrak{P} = \{(A, \mathbf{d}, \Sigma_L)$ first nonzero element in every column of A is positive, $\mathrm{diag}(\Sigma_L) = 1, \Sigma_L$ positive definite$\}$, and the probit two-tier model identifiability is defined as follows.

Definition 6. A probit two-tier model is identifiable at $(A, \mathbf{d}, \Sigma_L)$ if there is another set of parameters $(A', \mathbf{d}', \Sigma_L')$ such that $\mathbf{I}(A) = \mathbf{I}(A')$ and they define the same distribution, then it must hold that $(A, \mathbf{d}, \Sigma_L) = (A', \mathbf{d}', \Sigma_L')$.

Below, we provide a set of sufficient conditions for the identifiability of the probit two-tier model.

Theorem 9. *Under the probit two-tier model, suppose the true parameter satisfies Condition $C1$ and one of Conditions $T1$-$T3$. Then, the parameter is identifiable:*

T1S $|\mathcal{H}_4| \geq 3$, $A[\mathcal{B}_{\mathcal{H}_4}, 1:L]$ *contains an identity, where* $\mathcal{B}_{\mathcal{H}_4}$ *is the set of items that make up the testlets in* $\mathcal{H}_4$.

T2S $|\mathcal{H}_4| \geq 2$, $|\mathcal{H}_5| \geq 1$, $A[\mathcal{B}_{\mathcal{H}_4}, 1:L]$ *contains an identity.*

T3S *A[:,1:L] contains an identity, and there exists a partition of testlets* $\{1, \ldots, G\} = \mathcal{G}_1 \cup \mathcal{G}_2$, *such that (a)* $\bar{A}_{\mathcal{G}_1}$ *has full column rank, and (b)* $A[\mathcal{B}_{\mathcal{G}_2}, 1:L]$ *has full column rank.*

Here, the definitions of the sets ($\mathcal{H}$s) and the interpretations of the conditions remain as they were for the linear two-tier model in section 2.3.

4. Remarks

4.1. Orthogonality between primary and testlet dimensions

Discussions on the identification restriction for bifactor models can be found in Rijmen (2009), where it is pointed out that three types of identification restrictions are required:

- $G+1$ restrictions for fixing the origins of general and testlet effects.
- $G+1$ restrictions for fixing the scales of general and testlet effects.
- G restrictions for dealing with the rotation issue.

By translating the restrictions into mathematical expressions, the above three conditions are equivalent to

$$\boldsymbol{\eta} \sim N(\mathbf{0}, \Sigma), \quad \Sigma = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \Sigma_G \end{pmatrix}, \tag{4.1}$$

with $\Sigma_G[g,g] = 1$, for all $g \in \{1, \ldots, G\}$. However, they do not provide a rigorous proof of why we need the third type of restriction. Below, we provide Theorem 10 to answer this question. Consider the parameter space

$$\mathfrak{P} = \{(A, \mathbf{d}, \Sigma_G, \boldsymbol{\rho}, \boldsymbol{\lambda}) \mid \operatorname{diag}(\Sigma) = 1, \Sigma \text{ is positive definite}\}, \tag{4.2}$$

where $\Sigma = \begin{pmatrix} 1 & \boldsymbol{\rho}^T \\ \boldsymbol{\rho} & \Sigma_G \end{pmatrix}$.

Theorem 10. *The bifactor model is not identifiable at any* $(A, d, \Sigma_G, \boldsymbol{\rho}, \boldsymbol{\lambda})) \in \mathfrak{P}$, *as defined in* (4.2).

The implication of Theorem 10 is that there is no identifiable model in $\mathfrak{P}$, as defined in (4.2), that is, when the orthogonality restriction between the primary and the testlet dimensions is dropped. This explains why the identification results can only be extended to correlated testlet dimensions.

4.2. Extensions

The results here can be extended to more general settings:

- The normality assumptions in the linear bifactor model can be removed. That is, we do not require $\eta_g \sim N(0,1)$ and $\epsilon_j \sim N(0,\lambda_j)$, but instead assume $\mathrm{Var}(\eta_g) = 1$ and $\mathrm{Var}(\epsilon_j) = \lambda_j$. By checking the first and second moments, it is not hard to see that the sufficient conditions in the previous theorems still guarantee identification.

- For the ordinal probit model, each Y_j takes values in $\{1,\ldots,K_j\}$ $(K_j \geq 2)$ and follows the probability distribution

$$P(Y_j > k \mid \boldsymbol{\eta}) = \Phi(d_j^{(k)} + \mathbf{a}_j^T \boldsymbol{\eta}), \tag{4.3}$$

  for $k = 1,\ldots,K_j - 1$, with $d_j^{(1)} \geq d_j^{(2)} \geq \cdots \geq d_j^{(K_j-1)}$. Under the same set of sufficient conditions, we can easily obtain the identifiability results.

4.3. Connections

Under the linear bifactor model setting, the sufficient condition in Theorem 4 given by Anderson and Rubin (1956) can be simplified, in the sense that *there are at least three items in each testlet, i.e.* $|Q_g| \geq 3$, *for all* g. (Suppose there exists a testlet with at most two items. Then, it is impossible to find two disjoint submatrices of A with full column rank after deleting an item within that testlet.) It can also be checked that this sufficient condition is satisfied by E1S and E2S in our Theorem 2.

For general linear factor models, two- and three-indicator rules are two sets of simple sufficient identifiability conditions; see Bollen (1989):

- *Two-indicator rules:* (1) each latent factor is related to three items; (2) each row of A has one and only one nonzero element; (3) the latent factors are uncorrelated; and (4) ϵs are uncorrelated.

- *Three-indicator rules:* (1) each latent factor is related to two items; (2) each row of A has one and only one nonzero element; (3) there are no elements in Σ; and (4) ϵs are uncorrelated.

Although two- and three-indicator rules seem similar to the conditions in Theorems 1 and 2, they cannot be applied to bifactor/two-tier models. By nature, it is impossible to assume each row of A has one and only one nonzero element, because each item has at least two latent dimensions (general factors and testlet-specific factor). Fortunately, three items are enough to identify the testlet effects. Owing to the model structure, the general factor can also be identified when there is a sufficient number of testlets.

5. Discussion

This study addresses the fundamental issue of the dentifiability of a bifactor model and its extensions, under both a linear model with continuous indicators and a probit model with dichotomous responses. The identifiability (or nonidentifiability) of a model can be determined using easily checkable conditions. In particular, conditions $P1$ and $P2$ establish the minimum requirements that can ensure the identifiability of the standard bifactor model. For the extended bifactor model with correlated subdimensions, a set of necessary conditions $(E1N, E2N)$ and a set of sufficent conditions $(E1S - E3S)$ for parameter identifiability are proposed. Sufficient conditions for two-tier model identifiability are presented in $C0$ (or $C1$ for the probit model) and $T1S - T3S$. Theoretical results explain the under-identification phenomena observed in existing literature. Simulation studies demonstrated the effects on parameter estimation when the identifiability conditions were or were not met. From a practical viewpoint, these checkable identifiability conditions can guide test developers through the design and evaluation of bifactor-type assessments.

Note that although both probit and logistic models can be applied for binary outcomes, the current identifiability results for probit models do not directly apply to item bifactor analysis with logistic parameterization, as seen in DeMars (2006), Cai (2010), and Jeon, Rijmen and Rabe-Hesketh (2013). When a normal distribution is assumed for the latent traits, random-effect logistic item factor models involve a convolution of Gaussian and logistic random variables. Hence, this class of models does not imply the same first and second moments for the item responses as those in the probit case. Future research may look into the identifiability conditions for bifactor-type models with a logit link, perhaps adopting similar approaches to those in San Martín, Rolin and Castro (2013) for two-parameter logistic item response models. The current bifactor model identification findings may also be extended to higher-order factor models Yung, Thissen and McLeod (1999), under which latent factors are assumed to exhibit a hierarchical structure, with higher-order latent factors governing secondary, specific factors.

Supplementary Materials

The online Supplementary Material contains the simulation studies, illustrative examples, and technical proofs of the main theoretical results.

Acknowledgments

This research was supported in part by NSF Grants DMS-2015417, SES-1826540, and IIS-1633360.

References

Allman, E. S., Matias, C and Rhodes, J. A. (2009). Identifiability of parameters in latent structure models with many observed variables. *The Annals of Statistics* **37**, 3099–3132.

Anderson, T. W. and Rubin, H. (1956). Statistical inference in factor analysis. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* **5**, 111–150.

Bollen, K. A. (1989). *Structural Equations with Latent Variables.* Wiley, New York.

Bollen, K. A. and Bauldry, S. (2010). Model identification and computer algebra. *Sociological Methods & Research* **39**, 127–156.

Bradlow, E. T., Wainer, H. and Wang, X. (1999). A Bayesian random effects model for testlets. *Psychometrika* **64**, 153–168.

Cai, L. (2010). A two-tier full-information item factor analysis model with applications. *Psychometrika* **75**, 581–612.

Cai, L., Choi, K., Hansen, M. and Harrell, L. (2016). Item response theory. *Annual Review of Statistics and Its Application* **3**, 297–321.

Cai, L., Yang, J. S. and Hansen, M. (2011). Generalized full-information item bifactor analysis. *Psychological Methods* **16**, 221.

Caspi, A., R., Houts, M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S. et al. (2014). The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science* **2**, 119–137.

Castellan, N. J. (1966). On the estimation of the tetrachoric correlation coefficient. *Psychometrika* **31**, 67–73.

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J.-P. and Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality* **80**, 219–251.

Chen, Y., Li, X. and Zhang, S. (2019). Structured latent factor analysis for large-scale data: Identifiability, estimability, and their implications. *Journal of the American Statistical Association* **115**, 1756–1770.

Chen, Y., Liu, J., Xu, G. and Ying, Z. (2015). Statistical analysis of q-matrix based diagnostic classification models. *Journal of the American Statistical Association* **110**, 850–866.

Christoffersson, A. (1975). Factor analysis of dichotomized variables. *Psychometrika* **40**, 5–32.

DeMars, C. E. (2006). Application of the bi-factor multidimensional item response theory model to testlet-based tests. *Journal of Educational Measurement* **43**, 145–168.

DeMars, C. E. (2012). Confirming testlet effects. *Applied Psychological Measurement* **36**, 104–121.

Eid, M., Krumm, S., Koch, T. and Schulze, J. (2018). Bifactor models for predicting criteria by general and specific factors: Problems of nonidentifiability and alternative solutions. *Journal of Intelligence* **6**, 42.

Gibbons, R. D., Bock, R. D., Hedeker, D., Weiss, D. J., Segawa, E., Bhaumik, D. K. et al. (2007). Full-information item bifactor analysis of graded response data. *Applied Psychological Measurement* **31**, 4–19.

Gibbons, R. D. and Hedeker, D. R. (1992). Full-information item bi-factor analysis. *Psychometrika* **57**, 423–436.

Gibbons, R. D., Rush, A. J. and Immekus, J. C. (2009). On the psychometric validity of the domains of the PDSQ: An illustration of the bi-factor item response theory model. *Journal of Psychiatric Research* **43**, 401–410.

Gignac, G. E. and Watkins, M. W. (2013). Bifactor modeling and the estimation of model-based reliability in the WAIS-IV. *Multivariate Behavioral Research* **48**, 639–662.

Green, S. and Yang, Y. (2018). Empirical underidentification with the bifactor model: A case study. *Educational and Psychological Measurement* **78**, 717–736.

Gu, Y. and Xu, G. (2019). The sufficient and necessary condition for the identifiability and estimability of the DINA model. *Psychometrika* **84**, 468–483.

Holzinger, K. J. and Swineford, F. (1937). The bi-factor method. *Psychometrika* **2**, 41–54.

Jennrich, R. I. and Bentler, P. M. (2012). Exploratory bi-factor analysis: The oblique case. *Psychometrika* **77**, 442–454.

Jeon, M., Rijmen, F. and Rabe-Hesketh, S. (2013). Modeling differential item functioning using a generalization of the multiple-group bifactor model. *Journal of Educational and Behavioral Statistics* **38**, 32–60.

Jöreskog, K. G. and Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Scientific Software International, Durham.

Kendall, M. G. and Stuart, A. (1958). *The Advanced Theory of Statistics* **1**. (Chaps. 1-6).

Kenny, D. A. and Milan, S. (2012). Identification: A non-technical discussion of a technical issue. In *Handbook of Structural Equation Modeling* (Edited by R. H. Hoyle), 145–163. Guilford Press, New York.

Kruskal, J. B. (1977). Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra and Its Applications* **18**, 95–138.

Lawley, D. N. (1943). Xxiii.— On problems connected with item selection and test construction. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics* **61**, 273–287.

Lord, F. (1952). *A Theory of Test Scores* (Psychometric Monograph No. 7). Psychometric Corporation, Richmond.

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika* **44**, 443–460.

Pearson, K. (1900). I. mathematical contributions to the theory of evolution.—VII. on the correlation of characters not quantitatively measurable. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **195**, 1–47.

Reilly, T. (1995). A necessary and sufficient condition for identification of confirmatory factor analysis models of factor complexity one. *Sociological Methods & Research* **23**, 421–441.

Reilly, T. and O'Brien, R. M. (1996). Identification of confirmatory factor analysis models of arbitrary complexity: The side-by-side rule. *Sociological Methods & Research* **24**, 473–491.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research* **47**, 667–696.

Reise, S. P., Morizot, J. and Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research* **16**, 19–31.

Rijmen, F. (2009). Efficient full information maximum likelihood estimation for multidimensional IRT models. *ETS Research Report Series*, i–31.

Rijmen, F. (2010). Formal relations and an empirical comparison among the bi-factor, the testlet and a second-order multidimensional IRT model. *Journal of Educational Measurement* **47**, 361–372.

San Martín, E., Rolin, J.-M. and Castro, L. M. (2013). Identification of the 1PL model with guessing parameter: Parametric and semi-parametric results. *Psychometrika* **78**, 341–379.

Thurstone, L. L. (1927). The unit of measurement in educational scales. *Journal of Educational Psychology* **18**, 505.

Xu, G. (2017). Identifiability of restricted latent class models with binary responses. *The Annals of Statistics* **45**, 675–707.

Xu, G. and Zhang, S. (2016). Identifiability of diagnostic classification models. *Psychometrika* **81**, 625–649.

Yung, Y.-F., Thissen, D. and McLeod, L. D. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika* **64**, 113–128.

Guanhua Fang

Cognitive Computing Lab, Baidu Research, Bellevue, CA 98004, USA.

E-mail: gf2340@columbia.edu

Jinxin Guo

School of Mathematics and Statistics, Northeast Normal University, Changchun, Jilin, China 130024.

E-mail: guojx235@nenu.edu.cn

Xin Xu

School of Mathematics and Statistics, Northeast Normal University, Changchun, Jilin, China 130024.

E-mail: xux636@nenu.edu.cn

Zhiliang Ying

Department of Statistics, Columbia University, New York, NY 10027, USA.

E-mail: zy9@columbia.edu

Susu Zhang

Departments of Psychology and Statistics, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA.

E-mail: szhan105@illinois.edu

(Received September 2020; accepted December 2020)