PSYCHOMETRIKA—VOL. 89, NO. 1, 241–266
MARCH 2024
https://doi.org/10.1007/s11336-023-09944-3

GENERALIZED STRUCTURED COMPONENT ANALYSIS ACCOMMODATING CONVEX COMPONENTS: A KNOWLEDGE-BASED MULTIVARIATE METHOD WITH INTERPRETABLE COMPOSITE INDEXES

GYEONGCHEOL CHO
THE OHIO STATE UNIVERSITY

HEUNGSUN HWANG
MCGILL UNIVERSITY

Generalized structured component analysis (GSCA) is a multivariate method for examining theory-driven relationships between variables including components. GSCA can provide the deterministic component score for each individual once model parameters are estimated. As the traditional GSCA always standardizes all indicators and components, however, it could not utilize information on the indicators' scale in parameter estimation. Consequently, its component scores could just show the relative standing of each individual for a component, rather than the individual's absolute standing in terms of the original indicators' measurement scales. In the paper, we propose a new version of GSCA, named *convex GSCA*, which can produce a new type of unstandardized components, termed convex components, which can be intuitively interpreted in terms of the original indicators' scales. We investigate the empirical performance of the proposed method through the analyses of simulated and real data.

Key words: generalized structured component analysis, convex component, multivariate analysis, composite index, interpretability.

Generalized structured component analysis (GSCA; Hwang & Takane, 2004; 2014) is a multivariate method that allows for specifying and testing path-analytic relationships between observed variables and components (i.e., weighted sums of observed variables). Observed variables forming components are called composite indicators (Bollen & Bauldry, 2011). Given a theory-driven model, GSCA constructs components from composite indicators such that the components can explain the total variances of all dependent variables in the model as much as possible.

As in many component analysis techniques, GSCA has typically assumed that all components and indicators were standardized to have zero means and unit variances. This traditional, standardized version of GSCA shall be called GSCA_{std} hereafter. GSCA_{std} begins by standardizing indicators prior to estimating parameters and updates component weights in such a way that they produce standardized components during the estimation process. Such standardization can be useful for the interpretation and comparison of GSCA_{std}'s estimates because the GSCA_{std} model is equivalent to a system of multiple regression equations for standardized components and indicators, indicating that its loadings and path coefficients can be interpreted as standardized regression coefficients.

Nonetheless, the conventional standardization of components makes it difficult to interpret *component scores* in terms of the original indicators' measurement scales. The standardized component score for an individual merely shows the individual's relative location to the other individuals in the sample and the absolute score itself is not interpretable. This is less attractive to researchers who are interested in the absolute level of a component for each individual. For

Correspondence should be made to Gyeongcheol Cho, Department of Psychology, The Ohio State University, 1827 Neil Avenue, Columbus, OH 43210, USA. Email: cho.1240@osu.edu

example, if a standardized component is used to measure the level of life satisfaction, an individual's component score can inform whether s/he has a relatively lower or higher level of life satisfaction than the others. However, it cannot tell exactly what the level of her/his life satisfaction is, reflecting whether s/he is satisfied or dissatisfied with her/his life.

Moreover, if indicators for each component are measured on the same scale, which is often observed in practice, standardizing the indicators may not be recommended because it can eliminate "the natural and relevant variability present" (Naik & Khattree, 1996) in each of the indicators, forcing them to have the same variance, although their variances may not be the same in reality. For illustration, suppose that we made two versions of test batteries to assess children's intelligence, both of which were measured on a 0 to 100 scale. Three children took these tests and obtained {49, 50, 51} for Test 1 and {0, 50, 100} for Test 2. The results show that Test 1 almost fails to differentiate the children's intelligence levels, whereas Test 2 differentiates their intelligence level very well, indicating that the difference in score variability between the two tests is interpretable and contains meaningful information. However, when we standardize these scores, such information disappears since both score sets become identical (i.e., {–1, 0, 1}). If GSCA_{std} is applied to the tests, the same standardized weight values (i.e., .5) will be assigned to the two tests.

To obtain unstandardized component scores from original indicators, GSCA_{std} applies an additional rescaling of weight estimates after convergence (Hwang & Takane, 2014, Chapter 2). As will be discussed in more detail in Sect. 1, each indicator's weight estimate is rescaled by dividing it by the indicator's standard deviation. Subsequently, unstandardized component scores are obtained by pre-multiplying the rescaled weights by their indicators' original scores.

However, this rescaling procedure has two issues. Firstly, the procedure is carried out while keeping the variances of components fixed to one. Thus, the variances of the resultant unstandardized components are likely to be different from those of the original indicators, so that it is not guaranteed that the unstandardized component scores would vary within the same range of the original indicators. Secondly, the rescaling procedure tends to have indicators with relatively small variances influence the construction of their unstandardized component more heavily. In the above example, as the sample standard deviations of the two test batteries were 1 and 50, the unstandardized weights obtained from this ad-hoc rescaling procedure would be .5 and .01 for Tests 1 and 2, respectively. This indicates that Test 1 is 50 times more influential for forming children's unstandardized component scores than Test 2, even though Test 2 differentiates children's intelligence levels much better than Test 1. In Sect. 1, we will explain why this issue occurs in the rescaling procedure.

To address these issues, we propose a different version of GSCA, named *convex GSCA* or GSCA_{cvx} for short, which can estimate unstandardized components of original indicators. Specifically, GSCA_{cvx} obtains an unstandardized component as a convex combination of original indicators, termed a *convex component*, if the indicators for the component have the same measurement scale. A convex combination of a set of vectors refers to a special linear combination whose weights are non-negative and summed up to one (Lay et al., 2015, Chapter 8). As will be shown in Sect. 1, a convex component's scores are within the same range of its indicators' scores. This property of the convex component facilitates the interpretation of its component scores with reference to the indicators' scales. Moreover, GSCA_{cvx} avoids the unnecessary standardization of indicators when they are on the same measurement scales, allowing for utilizing information on their variances in parameter estimation.

The remaining sections of the paper are organized as follows. In Sect. 1, we briefly describe GSCA_{std} and explain its ad-hoc procedure of computing unstandardized components and the procedure's limitation. In Sect. 2, we introduce a convex component and explain its six properties. In Sect. 3, we present the GSCA_{cvx} model that accommodates convex components and propose an iterative algorithm for estimating model parameters. We also provide a set of overall goodness-

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

of-fit and cross-validation indexes for model evaluation and comparison. In Sect. 4, we conduct a Monte-Carlo simulation study to examine GSCA$_{cvx}$'s parameter recovery. In Sect. 5, we apply GSCA$_{cvx}$ to real data to demonstrate its practical usefulness. In the final section, we summarize the previous sections and discuss the method's implications and prospective extensions.

1. Traditional GSCA with Standardized Variables

1.1. Model and Parameter Estimation

GSCA$_{std}$ involves three sub-models—weighted relation, component measurement, and structural models (Hwang & Takane, 2004, 2014) . Let $\mathbf{z}_{std} = [z_{std,1}, z_{std,2}, \cdots, z_{std,J}]'$ denote a J by 1 random vector of standardized indicators, where $z_{std,j}$ is the jth standardized indicator, i.e., $E(z_{std,j}) = 0$ and $var(z_{std,j}) = 1$ ($j = 1, 2, \cdots, J$). The mean of $\mathbf{z}_{std}$ is a zero vector and the correlation matrix of $\mathbf{z}_{std}$ is denoted by $\mathbf{\Sigma}_{std}$. Let $\boldsymbol{\gamma}_{std} = [\gamma_{std,1}, \gamma_{std,2}, \cdots, \gamma_{std,P}]'$ denote a P by 1 random vector of standardized components, where $\gamma_{std,p}$ is the pth standardized component, i.e., $E(\gamma_{std,p}) = 0$, $var(\gamma_{std,p}) = 1$ ($p = 1, 2, \cdots, P$). Let $\mathbf{W}_{std}$ denote a J by P matrix consisting of component weights assigned to indicators. Let $\mathbf{C}_{std}$ denote a P by J matrix of loadings relating components to indicators. Let $\mathbf{B}_{std}$ denote a P by P matrix of path coefficients relating components to each other. Let $\boldsymbol{\xi} = [\xi_1, \xi_2, \cdots, \xi_J]'$ denote a J by 1 random vector of errors in the component measurement model, where ξ_j is an error for the jth indicator. Let $\boldsymbol{\zeta} = [\zeta_1, \zeta_2, \cdots, \zeta_P]'$ denote a P by 1 random vector of errors in the structural model, where ζ_p is an error for the pth component. The three sub-models of GSCA$_{std}$ are expressed as follows.

$$\boldsymbol{\gamma}_{std} \equiv \mathbf{W}'_{std}\mathbf{z}_{std} \text{ (weighted relation model)} \tag{1}$$

$$\mathbf{z}_{std} = \mathbf{C}'_{std}\boldsymbol{\gamma}_{std} + \boldsymbol{\xi} \text{ (component measurement model)} \tag{2}$$

$$\boldsymbol{\gamma}_{std} = \mathbf{B}'_{std}\boldsymbol{\gamma}_{std} + \boldsymbol{\zeta} \text{ (structural model).} \tag{3}$$

The weighted relation model (1) shows that (standardized) components are defined as a linear combination of standardized indicators. The component measurement and structural models (2) and (3) express the directional relationships between the indicators and components and those among the components, respectively. As (2) and (3) can be seen as systems of linear regression equations, their model parameters, including loadings and path coefficients, can be interpreted in the same manner as standardized regression coefficients. The three sub-models are combined into the following equation,

$$\begin{aligned}
&[\mathbf{z}_{std}; \boldsymbol{\gamma}_{std}] = [\mathbf{C}_{std}, \mathbf{B}_{std}]'\boldsymbol{\gamma}_{std} + [\boldsymbol{\xi}; \boldsymbol{\zeta}] \\
&\quad \leftrightarrow [\mathbf{I}_J, \mathbf{W}_{std}]'\mathbf{z}_{std} = [\mathbf{C}_{std}, \mathbf{B}_{std}]'\mathbf{W}'_{std}\mathbf{z}_{std} + [\boldsymbol{\xi}; \boldsymbol{\zeta}] \\
&\quad \leftrightarrow \mathbf{V}'_{std}\mathbf{z}_{std} = \mathbf{A}'_{std}\mathbf{W}'_{std}\mathbf{z}_{std} + \mathbf{e},
\end{aligned} \tag{4}$$

where $\mathbf{I}_J$ is the identity matrix of order J, $\mathbf{V}_{std} \equiv [\mathbf{I}_J, \mathbf{W}_{std}]$, $\mathbf{A}_{std} \equiv [\mathbf{C}_{std}, \mathbf{B}_{std}]$, $\mathbf{e} \equiv [\boldsymbol{\xi}; \boldsymbol{\zeta}]$ and a semicolon within brackets is an operator to vertically concatenate two vectors in the array. The equation (4) is called the GSCA$_{std}$ model.

Let $\mathbf{1}_Q$ denote a column vector of Q ones. Let $SS(\mathbf{X}) \equiv tr(\mathbf{X}'\mathbf{X})$ for any matrix $\mathbf{X}$. Let *vecdiag*() denote an operator that returns a column vector stacking the diagonal elements of a square matrix one below another. GSCA$_{std}$ estimates model parameters ($\mathbf{W}_{std}$ and $\mathbf{A}_{std}$) by

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

minimizing the following objective function

$$
\begin{aligned}
& f_{std}(\mathbf{W}_{std}, \mathbf{A}_{std}) \\
& = tr(E(\mathbf{e}_{std}\mathbf{e}'_{std})) \\
& = E(SS([\mathbf{z}_{std}; \boldsymbol{\gamma}_{std}]' - \mathbf{z}'_{std}\mathbf{W}_{std}\mathbf{A}_{std}), \\
& = E(SS(\mathbf{z}'_{std}([\mathbf{I}_J, \mathbf{W}_{std}] - \mathbf{W}_{std}\mathbf{A}_{std})) \\
& = tr((\mathbf{V}_{std} - \mathbf{W}_{std}\mathbf{A}_{std})'\boldsymbol{\Sigma}_{std}(\mathbf{V}_{std} - \mathbf{W}_{std}\mathbf{A}_{std}))
\end{aligned} \tag{5}
$$

subject to $vecdiag(\mathbf{W}'_{std}\boldsymbol{\Sigma}_{std}\mathbf{W}_{std}) = \mathbf{1}_p$. Thus, GSCA$_{std}$ estimates the model parameters by minimizing the sum of error variances for all variables in the model given $\boldsymbol{\Sigma}_{std}$. In general, $\boldsymbol{\Sigma}_{std}$ is replaced with the sample correlation matrix of indicators, denoted by $\mathbf{S}_{std}$. The objective function (5) also shows that GSCA$_{std}$ aims to create components that explain the total variances of variables in the model rather than their covariances, as with PCA or other component-based methods. The error terms in the GSCA$_{std}$ model are not considered independent entities that cause the variation of indicators but simply treated as residuals that are unexplained by independent components. Thus, GSCA$_{std}$ typically makes no assumptions about the correlation structure of the error terms of indicators, leaving them freely correlated. This is distinct from the common factor model, where the error terms are typically assumed to be uncorrelated. Nonetheless, no error covariances between different blocks of indicators may be assumed in some special cases of GSCA (Cho et al., 2020, 2022).

Note that (1) defines a component as a weighted sum of indicators, which is also the case in PCA. However, this equation itself is not identified because there would exist infinitely different ways of deciding the component weights. Thus, we need a certain rule or criterion to determine the component weights. PCA's criterion is one of the most widely used ones in statistics that the weights are to be determined in such a way that their corresponding components explain the maximum total variance of the indicators. The regression coefficients of indicators on their component are (component) loadings. These relationships between components and their indicators are expressed in the component measurement model (2). Thus, GSCA can have confirmatory PCA (Takane, Kiers, & de Leeuw, 1995) as a special case when it considers (1) and (2) only.

As the minimization problem (5) cannot be solved in closed form, an alternating least squares (ALS) algorithm was developed for iteratively finding the minimum point of (5). In the ALS algorithm, $\mathbf{W}_{std}$ and $\mathbf{A}_{std}$ are updated alternately with the other fixed until the difference in (5) between consecutive iterations decreases beyond a pre-specified tolerance level (e.g., 10^{-5}) (see Hwang & Takane, 2014, Chapter 2, for a full description of the ALS algorithm). Let $\widehat{\boldsymbol{\Gamma}}_{std}$ denote an N by P matrix of the standardized score estimates of components, $\mathbf{D}_{std}$ denote an N by J matrix of the standardized scores of indicators, and N is the number of cases in the sample. Let us suppose that we obtain the estimates of $\mathbf{W}_{std}$ and $\mathbf{A}_{std}$ that minimize (5), denoted by $\widehat{\mathbf{W}}_{std}$ and $\widehat{\mathbf{A}}_{std}$. Then, a matrix of standardized component scores is obtained by

$$
\widehat{\boldsymbol{\Gamma}}_{std} \equiv \mathbf{D}_{std}\widehat{\mathbf{W}}_{std}. \tag{6}
$$

1.2. Unstandardized Weight Estimates in GSCA $_{std}$

Let $\mathbf{D} = \mathbf{1}_N\hat{\boldsymbol{\mu}}' + \mathbf{D}_{std}\widehat{\boldsymbol{\Delta}}_z$ denote an N by J matrix of the unstandardized scores of indicators, where $\hat{\boldsymbol{\mu}}$ is a J by 1 sample mean vector and $\widehat{\boldsymbol{\Delta}}_z$ is a diagonal matrix whose entries are sample standard deviations of unstandardized indicators. Conventionally, unstandardized component

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

means are subsequently computed by transforming $\widehat{\mathbf{W}}_{std}$ as follows. As it follows from (6) that

$$\begin{aligned}
&\widehat{\mathbf{\Gamma}}_{std} = (\mathbf{D} - \mathbf{1}_N\hat{\boldsymbol{\mu}}')\widehat{\boldsymbol{\Delta}}_{\mathrm{z}}^{-1}\widehat{\mathbf{W}}_{std} \\
&\leftrightarrow 1_N\hat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Delta}}_{\mathrm{z}}^{-1}\widehat{\mathbf{W}}_{std} + \widehat{\mathbf{\Gamma}}_{std} = \mathbf{D}\widehat{\boldsymbol{\Delta}}_{\mathrm{z}}^{-1}\widehat{\mathbf{W}}_{std} \\
&\leftrightarrow 1_N\hat{\boldsymbol{\mu}}'\widehat{\mathbf{W}}_{uni} + \widehat{\mathbf{\Gamma}}_{std} = \mathbf{D}\widehat{\mathbf{W}}_{uni},
\end{aligned} \tag{7}$$

where $\widehat{\mathbf{W}}_{uni} \equiv \widehat{\boldsymbol{\Delta}}_{\mathrm{z}}^{-1}\widehat{\mathbf{W}}_{std}$, GSCA$_{std}$ computes unstandardized component scores, denoted here by $\widehat{\mathbf{\Gamma}}_{uni}$, as $\widehat{\mathbf{\Gamma}}_{uni} \equiv \mathbf{D}\widehat{\mathbf{W}}_{uni}$ (Hwang & Takane, 2014, p. 26).

As shown in the last line of (7), however, $\widehat{\mathbf{\Gamma}}_{uni}$ can be simply seen as a variant of standardized component scores whose means are only relocated *a posteriori* by $\mathbf{1}_N\hat{\boldsymbol{\mu}}'\widehat{\mathbf{W}}_{uni}$ in that $\widehat{\mathbf{\Gamma}}_{std}$ remains standardized irrespective of the sample variances of the original indicators. Consequently, it is not guaranteed that the scores of $\widehat{\mathbf{\Gamma}}_{uni}$ are within the same range of the unstandardized scores of their indicators, which will be empirically shown in Sect. 4. Also, as illustrated in Sect. 1, GSCA$_{std}$ tends to assign smaller unstandardized weights to original indicators with relatively large variances in forming $\widehat{\mathbf{\Gamma}}_{uni}$. That is because minimizing (5) involves imposing a relatively large penalty on an original indicator with a relatively large variance, which is shown in Appendix 1. This disproportionate penalization for original indicators can inadvertently amplify the influence of an original indicator with a small variance on GSCA$_{std}$'s parameter estimation. Such an approach could be deemed unsuitable when one aims to obtain an unstandardized component of original indicators on a single scale.

1.3. Convex Component and Its Six Properties

Let γ_p denote the pth component ($p = 1, 2, \cdots, P$) that is assumed to have the mean τ_p and variance ϕ_p. Let $\mathbf{z}_p$ denote a J_p by 1 vector of indicators for γ_p, where J_p is the number of indicators for γ_p. We call the vector $\mathbf{z}_p$ *a block of indicators* for γ_p, which is assumed to have the mean vector $\boldsymbol{\mu}_p$ and covariance matrix $\boldsymbol{\Sigma}_p$. Let $\mathbf{w}_p$ denote a J_p by 1 vector of weights for $\mathbf{z}_p$. Let $\mathbf{0}_{k\times l}$ denote a k by l matrix of zeros, where k and l are any scalars. If γ_p is defined as a convex component, it can be expressed as

$$\gamma_p \equiv \mathbf{w}_p'\mathbf{z}_p \text{ subject to } \mathbf{w}_p'\mathbf{1}_{Jp} = 1 \text{ and } \mathbf{w}_p \geq \mathbf{0}_{J\times 1}. \tag{8}$$

A convex component has six useful properties as follows.

Proposition 1. *A convex component has scores within the range of its indicators' scores.*

Proposition 2. *Each score of a convex component corresponds to a component score of an individual whose scores for indicators are all the same as the component score.*

Proposition 3. *The mean of a convex component is not fixed to zero but is determined by weights within the range of its indicators' means.*

Proposition 4. *The standard deviation of a convex component is not fixed to one but is determined by weights within the range from 0 to the maximum standard deviation of its indicators.*

Proposition 5. *Given a linearly independent set of indicators' scores, a set of convex component scores has a unique set of weights that are nonnegative and summed up to one.*

Proposition 6. *The path coefficient of a convex component on an outcome variable indicates the expected amount of change in the outcome variable for a unit change in each indicator of the convex component while holding other variables fixed.*

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

We provide proofs for the six propositions in Appendix 2. The first four properties make a convex component' scores, mean, and standard deviation interpretable with reference to its indicators' scale when its indicators are on the same scale. The fifth property allows interpreting weight parameters as the contribution rates of indicators to forming their component. The last property allows for interpreting the path coefficient of a convex component with respect to its indicators' scale. We here illustrate these properties with an example of (major) depression.

Let us assume that depression can be represented by a convex component (γ) with three symptom-related indicators (z_1 = depressed affect, z_2 = somatic discomfort, and z_3 = interpersonal problem), which are commonly rated on a seven-point Likert scale (0 = "none", 1 = "minimal", 2 = "mild", 3 = "moderate", 4 = "moderately severe", 5 = "severe", and 6 = "extremely severe"). It is generally considered safe to treat ordinal variables with five or more categories as continuous (Johnson & Creech, 1983; Norman, 2010; Sullivan & Artino, 2013; Zumbo & Zimmerman, 1993). Then, this depression component serves as a summary index whose score indicates the overall severity level of the three depressive symptoms for each individual. Specifically, once weight parameters are estimated, a score set of depression component is obtained given a dataset of its indicators. Proposition 1 indicates that all individuals' scores of depression component will be within the range of the measurement scale of its indicators (e.g., [0, 6]). Proposition 2 implies that each individual's score of depression component within the range can be interpreted as the depression level of an individual whose indicators' scores are all the same as the depression component score. For example, if a patient's depression component score is 3, it implies that their depression level can be considered equivalent to that of depression of a patient whose symptom levels are all moderate (i.e., 3), suggesting that their depression is generally moderate. By Propositions 3 and 4, the means and the standard deviations of depression component are determined by weight parameter estimates within the range of its indicators' original scales (e.g., [0, 6]) as well, which can also be interpreted in relation to those scales. For instance, if the mean of depression component scores turns out to be 5, it means that the average depression level of patients in the sample can be considered equivalent to the depression level of a patient whose symptom levels are all severe, or that the patients' depression is severe on average. Also, if the standard deviation of depression component scores turns out to be 1, it implies that the depression severity levels of patients in the sample were one-unit lower or higher than the moderate level on average.

By Proposition 5, it is guaranteed that once a set of depression component scores is obtained with a set of weight estimates, any other set of weight estimates does not exist that makes the same score set of depression component while satisfying the constraint in (8). As these weight estimates are always non-negative and summed up to one, they can be interpreted as the indicators' contribution 'rates' of forming the convex component. For example, suppose that the weight estimates for z_1, z_2 and z_3 are .41, .24, and .35, respectively. It indicates that when the severity level of depression component increases by one unit due to a one-unit increase in all the three symptom-related indicators, the contribution rates of z_1, z_2 and z_3 to the one-unit increase of depression severity are 41%, 24%, and 35%, respectively. Such interpretation was not applicable to weight of standardized components, as their values can be negative and not necessarily summed up to one. Note that this proposition is satisfied only if a linearly independent set of indicators' scores is given as a dataset. A set of indicators' scores being linearly independent means that a score vector of an indicator cannot be expressed as a linear combination of score vectors of the other indicators, which further implies that sample covariance matrix of the indicators is positive definite.

By Proposition 6, the path coefficient of a convex component on an outcome variable can be interpreted as an aggregate effect of the indicators of the convex component on the outcome variable, given that the structural model holds. For example, let's consider a situation where a path coefficient of a depression component on employment earnings for the year of depression

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

reported is identified -$5000 (e.g., Dobson et al., 2021). This would suggest that a one unit increase across all depression symptoms, such as a shift in all depression symptom levels from mild to moderate, would be associated with a $5000 loss for the individual experiencing depression. Such an interpretation was not feasible for path coefficients of standardized components.

2. Convex GSCA

2.1. Model Specification

Convex GSCA ($GSCA_{cvx}$) introduces a convex component with original indicators into the GSCA model. The $GSCA_{cvx}$ model also consists of three sub-models: weighted relation, component measurement, and structural models (Hwang & Takane, 2004, 2014) . Let $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \cdots, \gamma_P]'$ denote a P by 1 random vector of components. Each component (γ_p) can be either a convex or standardized component. If a block of indicators ($\mathbf{z}_p$) has the same measurement unit within the block, γ_p is defined as a convex component as expressed in (8). Otherwise, $\boldsymbol{\gamma}_p$ is defined as a standardized component, whose indicators ($\mathbf{z}_p$) are also assumed to be standardized such that $\tau_p = 0$, $\phi_p = 1$, $\boldsymbol{\mu}_p = \mathbf{0}_{Jp\times 1}$, and $vecdiag\ (\boldsymbol{\Sigma}_p) = \mathbf{1}_{Jp}$. Let $\mathbf{W}$ denote a J by P matrix consisting of component weights assigned to $\mathbf{z}$. Let $\mathbf{C}$ denote a P by J matrix of loadings relating $\boldsymbol{\gamma}$ to $\mathbf{z}$. Let $\mathbf{B}$ denote a P by P matrix of path coefficients relating γ to each other. Let $\mathbf{c}_0$ and $\mathbf{b}_0$ denote the column vectors of intercepts for the component measurement and structural models, respectively. The three sub-models of $GSCA_{cvx}$ are expressed as follows.

$$\boldsymbol{\gamma} \equiv \mathbf{W}'\mathbf{z}\ \text{(weighted relation model)} \tag{9}$$

$$\mathbf{z} = \mathbf{c}_0 + \mathbf{C}'\boldsymbol{\gamma} + \boldsymbol{\xi}\ \text{(component measurement model)} \tag{10}$$

$$\boldsymbol{\gamma} = \mathbf{b}_0 + \mathbf{B}'\boldsymbol{\gamma} + \boldsymbol{\zeta}\ \text{(structural model)}. \tag{11}$$

In $GSCA_{cvx}$, the weighted relation model (9) shows that each component is defined as a weighted sum of standardized or unstandardized indicators. As $GSCA_{cvx}$ may involve unstandardized variables, intercept terms ($\mathbf{c}_0$ and $\mathbf{b}_0$) are newly included into the component measurement and structural model (10) and (11). Each model parameter in (10) and (11)—intercepts, loadings, and path coefficients—can be interpreted in the same manner as the intercepts and regression coefficients in linear regression model with unstandardized variables. The three sub-models are combined into the following equation,

$$\begin{aligned}
&[\mathbf{z};\boldsymbol{\gamma}] = [\mathbf{c}_0;\mathbf{b}_0] + [\mathbf{C},\mathbf{B}]'\boldsymbol{\gamma} + \left[\boldsymbol{\xi};\boldsymbol{\zeta}\right] \\
&\leftrightarrow [\mathbf{I}_J,\mathbf{W}]'\mathbf{z} = [\mathbf{c}_0;\mathbf{b}_0] + [\mathbf{C},\mathbf{B}]'\mathbf{W}'\mathbf{z} + \left[\boldsymbol{\xi};\boldsymbol{\zeta}\right] \\
&\leftrightarrow \mathbf{V}'\mathbf{z} = \mathbf{a}_0 + \mathbf{A}'\mathbf{W}'\mathbf{z} + \mathbf{e},
\end{aligned} \tag{12}$$

where $\mathbf{a}_0 \equiv [\mathbf{c}_0;\mathbf{b}_0]$, $\mathbf{V} \equiv [\mathbf{I}_J, \mathbf{W}]$, $\mathbf{A} \equiv [\mathbf{C}, \mathbf{B}]$, and $\mathbf{e} \equiv [\boldsymbol{\xi};\boldsymbol{\zeta}]$. The equation (12) is called the $GSCA_{cvx}$ model. If every indicator and component is standardized, the $GSCA_{cvx}$ model (12) becomes identical to the $GSCA_{std}$ model (4)

2.2. Parameter Estimation

Let $\boldsymbol{\sigma}_p$ denote a J_p by 1 vector of standard deviations (SD) of $\mathbf{z}_p$. If the pth component is defined as standardized ones, $\boldsymbol{\sigma}_p$ is equivalent to $\mathbf{1}_{Jp}$. Let $\mathbf{O}_z$ denote a J by J diagonal matrix whose jth element is $J_p^{-1}\mathbf{1}_{Jp}/\boldsymbol{\sigma}_p$ if the jth indicator in the pth block is a dependent variable and zero otherwise. Let $\mathbf{O}_\gamma$ denote a P by P diagonal matrix whose pth element is $J_p^{-1}\mathbf{1}_{Jp}/\boldsymbol{\sigma}_p$ if the

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

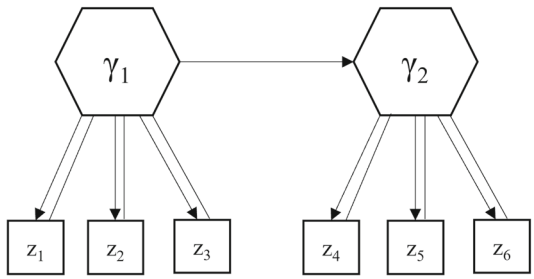

FIGURE 1.
An illustrative GSCA$_{cvx}$ model. Hexagons represent components, squares denote indicators, straight lines indicate weights, single-headed arrows denote loadings and path coefficients. All intercepts and error terms are omitted to make the figure concise.

pth component is a dependent variable and zero otherwise. Let $\mathbf{O} \equiv blkdiag(\mathbf{O}_z, \mathbf{O}_\gamma)$. GSCA$_{cvx}$ estimates parameters by minimizing the following objective function

$$\begin{aligned} & f_{cvx}(\mathbf{W}, \mathbf{A}, \mathbf{a}_0) \\ & = tr(\mathbf{O}E(\mathbf{e}\mathbf{e}')\mathbf{O}) \\ & = E(SS(([\mathbf{z}; \boldsymbol{\gamma}]' - (\mathbf{a}_0' + \mathbf{z}'\mathbf{W}\mathbf{A}))\mathbf{O})). \end{aligned} \tag{13}$$

subject to $\mathbf{w}_p' \boldsymbol{\Sigma}_p' \mathbf{w}_p = 1$ or $\mathbf{1}_{J_p}{}' \mathbf{w}_p = 1$ ($p = 1, 2, \cdots, P$). The objective function (13) shows that components in GSCA$_{cvx}$ are constructed such that they can minimize the "weighted" sum of error variances for all dependent variables under the constraints. Specifically, the objective function (13) penalizes each prediction error for dependent variables differentially by dividing it by the average SD of the corresponding block of indicators. This prevents prediction errors for a block of indicators with large variances from dominating the estimation of parameters.

To help understand the role of $\mathbf{O}$ in (13), we illustrate how $\mathbf{O}$ is determined based on the standard deviations of indicators. This will also explain the characteristic of the objective function described above. Figure 1 presents an illustrative GSCA$_{cvx}$ model involving two convex components (γ_1 and γ_2), each measured by three indicators that share the same scale, while the scales of two indicator blocks differ. Let us assume that $\boldsymbol{\sigma}_1 = [1; 2; 3]$ and $\boldsymbol{\sigma}_2 = [100; 200; 300]$, indicating that the differences in the overall magnitude of indicators' variances between the two blocks arises from the difference in scale. In this case, without $\mathbf{O}$ in (13) (i.e., $\mathbf{O} = \mathbf{I}$), the value of (13) would predominantly rely on the error variances for $\mathbf{z}_2$ and γ_2, implying that the error variances for $\mathbf{z}_1$ would be rarely considered in parameter estimation due to their scale. However, GSCA$_{cvx}$ determines $\mathbf{O} = blkdiag(\mathbf{O}_z, \mathbf{O}_\gamma)$, where $\mathbf{O}_z = blkdiag(2, 2, 2, 200, 200, 200)^{-1}$ and $\mathbf{O}_\gamma = blkdiag(0, 200^{-1})$, and then uses it to penalize the error variances for $\mathbf{z}_2$ and γ_2 to adjust their effects on (13). For instance, given $\mathbf{A} = \mathbf{0}$ and $\mathbf{a}_0 = E([\mathbf{z}; \boldsymbol{\gamma}])$, there are substantial differences in error variances between $\mathbf{z}_1$ and $\mathbf{z}_2$ (i.e., $[1^2; 2^2; 3^2]$ for $\mathbf{z}_1$ and $[100^2; 200^2; 300^2]$ for $\mathbf{z}_2$), but their error variances contribute equally to the value of (13) (i.e., $(1^2 + 2^2 + 3^2)/2^2 = [100^2; 200^2; 300^2]/200^2$). This suggests that introducing $\mathbf{O}$ into (13) enables GSCA$_{cvx}$ to consider prediction errors for both $\mathbf{z}_1$ and $\mathbf{z}_2$ during the parameter estimation process.

Conversely, as illustrated above, the objective function (13) does not impose different penalties on indicators within the same block to take into account potential differences in their variances. Furthermore, the objective function (13) is *partially scale-invariant*, which means that the minimum value of (13) does not vary with a linear change of measurement scales of each block of indicators that share the same scale (e.g., a scale range from 1–10 to 0–100), leading to the same weight estimates. This property is distinct from a property of *(full) scale invariant* (Swaminathan

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

& Algina, 1978) in that changing the measurement scales of each indicator differentially (e.g., standardization) is not considered. The proof for the property is provided in Appendix 3.

As the minimum point of (13) cannot be found in closed form, we developed an ALS algorithm for iteratively finding its minimum point. A detailed description of the ALS algorithm is provided in Appendix 4. Note that we do not constrain the weights for convex components to be non-negative in (13) to make the method more flexible. In some cases, researchers may wish to examine which indicators contribute to forming a component in the opposite direction to the other indicators and may be excluded during model re-specification. The negative weight of an indicator for a convex component may signify that the indicator is not suitable to form the component along with other indicators. As discussed in Appendix 4, the ALS algorithm allows for the imposition of the additional non-negativity constraints on weights, forcing the weights to be always positive.

2.3. Model Evaluation Indexes

GSCA$_{std}$ provides four overall goodness-of-fit measures, including FIT, AFIT, GFI, and SRMR, and one overall cross-validation index, out-of-bag prediction error (OPE). The FIT indicates the average explained variance of all variables in the model, whereas the AFIT is an adjusted version of FIT that takes into account the number of model parameters and sample size (Hwang & Takane, 2014, pp. 26–29). The GFI and SRMR evaluate the discrepancy between the sample and implied covariance matrices(Cho et al., 2020). The OPE aims to measure the average out-of-sample prediction error of the model for all variables via a bootstrapping-based cross validation and can be used for comparing models in terms of predictive generalizability (Cho et al., 2019). Whereas the GFI and SRMR can be used for GSCA$_{cvx}$ without modification, the FIT, AFIT, and OPE need to be modified for GSCA$_{cvx}$ because these measures were developed only for the condition where all variables are standardized. We revised FIT and OPE such that they can be applied for the GSCA$_{cvx}$ model with both standardized and unstandardized variables, taking into account the variances of dependent variables only.

We propose a modified version of FIT, termed FIT for unstandardized dependent variables (FIT$^{\mathrm{UD}}$), as follows.

$$\mathrm{FIT}^{\mathrm{UD}} = 1 - \frac{SS(([\mathbf{D}, \mathbf{D}\hat{\mathbf{W}}] - (\mathbf{1}_N\hat{\mathbf{a}}_0' + \mathbf{D}\hat{\mathbf{W}}\hat{\mathbf{A}}))\hat{\mathbf{O}})}{SS(([\mathbf{D}, \mathbf{D}\hat{\mathbf{W}}] - \mathbf{1}_N\hat{\boldsymbol{\mu}}'[\mathbf{I}_J, \hat{\mathbf{W}}])\hat{\mathbf{O}})}. \tag{14}$$

The FIT$^{\mathrm{UD}}$ indicates the proportion of the explained variance of all dependent variables (including dependent convex components) to their weighted total variance. If every component and indicator is standardized, FIT$^{\mathrm{UD}} = \frac{T}{T_Y}$FIT, where $T \equiv P + J$ and T_{Y} is the total number of dependent variables in the model. Also, we provide the following two local fit measures of FIT$^{\mathrm{UD}}$

$$\mathrm{FIT}_{\mathrm{M}}^{\mathrm{UD}} = 1 - \frac{SS((\mathbf{D} - (\mathbf{1}_N\hat{\mathbf{c}}_0' + \mathbf{D}\hat{\mathbf{W}}\hat{\mathbf{C}}))\hat{\mathbf{O}}_{\mathrm{z}})}{SS((\mathbf{D} - \mathbf{1}_N\hat{\boldsymbol{\mu}}')\hat{\mathbf{O}}_{\mathrm{z}})}, \tag{15}$$

$$\mathrm{FIT}_{\mathrm{S}}^{\mathrm{UD}} = 1 - \frac{SS((\mathbf{D}\hat{\mathbf{W}} - (\mathbf{1}_N\hat{\mathbf{b}}_0' + \mathbf{D}\hat{\mathbf{W}}\hat{\mathbf{B}}))\hat{\mathbf{O}}_{\gamma})}{SS((\mathbf{D}\hat{\mathbf{W}} - \mathbf{1}_N\hat{\boldsymbol{\mu}}'\hat{\mathbf{W}})\hat{\mathbf{O}}_{\gamma})}, \tag{16}$$

where $\hat{\mathbf{O}}_{\mathrm{z}}$ and $\hat{\mathbf{O}}_{\gamma}$ are sample analogies of $\mathbf{O}_{\mathrm{z}}$ and $\mathbf{O}_{\gamma}$. We refer to "local fit" as the goodness-of-fit of GSCA's sub-models. The FIT$_{\mathrm{M}}^{\mathrm{UD}}$ and FIT$_{\mathrm{S}}^{\mathrm{UD}}$ can be used for evaluating the component measurement and structural models, respectively. The FIT$_{\mathrm{M}}^{\mathrm{UD}}$ indicates the proportion of the explained variance of all dependent indicators to their weighted total variance, whereas the FIT$_{\mathrm{S}}^{\mathrm{UD}}$

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

indicates the proportion of the explained variance of all dependent (convex) components to their weighted total variance.

Moreover, we propose a revised version of OPE, termed OPE for dependent variables (OPE^{UD}), to evaluate the predictive generalizability of models involving convex components, as follows.

$$\text{OPE}^{\text{UD}} = \frac{1}{K}\sum_{k=1}^{K}\frac{SS(([\mathbf{D}_k^*, \mathbf{D}_k^*\hat{\mathbf{W}}_k] - (\mathbf{1}_{N_k}\hat{\mathbf{a}}_0' + \mathbf{D}_k^*\hat{\mathbf{W}}_k\hat{\mathbf{A}}_k))\hat{\mathbf{O}}_k)}{SS(([\mathbf{D}_k^*, \mathbf{D}_k^*\hat{\mathbf{W}}_k] - \mathbf{1}_{N_k}\hat{\boldsymbol{\mu}}_k'[\mathbf{I}_J, \hat{\mathbf{W}}_k])\hat{\mathbf{O}}_k)}, \tag{17}$$

where $\widehat{\mathbf{W}}_k$, $\widehat{\mathbf{A}}_k$, $\hat{\mathbf{a}}_k$, and $\hat{\boldsymbol{\mu}}_k$ are the parameter estimates obtained from the kth bootstrap sample ($k = 1,2,\cdots,K$), $\hat{\mathbf{O}}_k$ is the penalty term that rescales prediction errors for all dependent variables in the kth bootstrap sample, $\mathbf{D}_k^*$ is the kth test sample consisting of observations that are not included in the kth bootstrap sample, and N_k is the number of observations in the kth test sample. As shown in (17), the bootstrap sampling procedure generates pairs of mutually exclusive samples (bootstrap and test samples), over which a specified GSCA model is cross-validated (for a detailed description of OPE's computation, refer to Cho et al., 2019). The OPE^{UD} represents the weighted average out-of-sample prediction error of the model for dependent variables. The value of the OPE^{UD} ranges from 0 to infinity, where 0 means that a specified model perfectly predicts every dependent variable, and a value over 1 indicates that the prediction accuracy of a specified model is worse than that of the null model, where all dependent variables are predicted by their sample means. Again, when every variable is standardized, $\text{OPE}^{\text{UD}} = \frac{T}{T_Y}\text{OPE} - \frac{(T-T_Y)}{T_Y}$. In addition, we provide the following two local cross-validation indexes of OPE^{UD}

$$\text{OPE}_{\text{M}}^{\text{UD}} = \frac{1}{K}\sum_{k=1}^{K}\frac{SS((\mathbf{D}_k^* - (\mathbf{1}_{N_k}\hat{\mathbf{c}}_{0,k}' + \mathbf{D}_k^*\hat{\mathbf{W}}_k\hat{\mathbf{C}}_k))\hat{\mathbf{O}}_{z,k})}{SS((\mathbf{D}_k^* - \mathbf{1}_{N_k}\hat{\boldsymbol{\mu}}_k')\hat{\mathbf{O}}_{z,k})}, \tag{18}$$

$$\text{OPE}_{\text{S}}^{\text{UD}} = \frac{1}{K}\sum_{k=1}^{K}\frac{SS((\mathbf{D}_k^*\hat{\mathbf{W}}_k - (\mathbf{1}_{N_k}\hat{\mathbf{b}}_{0,k}' + \mathbf{D}_k^*\hat{\mathbf{W}}_k\hat{\mathbf{B}}_k))\hat{\mathbf{O}}_{\gamma,k})}{SS((\mathbf{D}_k^*\hat{\mathbf{W}}_k - \mathbf{1}_{N_k}\hat{\boldsymbol{\mu}}_k'\hat{\mathbf{W}}_k)\hat{\mathbf{O}}_{\gamma,k})}, \tag{19}$$

where $\hat{\mathbf{O}}_{z,k}$ and $\hat{\mathbf{O}}_{\gamma,k}$ are the penalty terms that rescale prediction errors for dependent indicators and components, respectively, in the kth bootstrap sample. The $\text{OPE}_{\text{M}}^{\text{UD}}$ and $\text{OPE}_{\text{S}}^{\text{UD}}$ can be used for evaluating the predictive generalizability of the component measurement and structural models, respectively.

3. Simulated Data Analysis

We conduct a simulation study to examine the parameter recovery of the proposed method. Figure 2 depicts the population GSCA_{cvx} model used in our simulation study. The population model involves four convex components, each of which is measured by four composite indicators. Indicators per block had different mean vectors: the mean vectors of indicators are [6, 5, 4, 3] for γ_1, [5.5, 4.5, 3.5, 2.5] for γ_2, [5, 4, 3, 2] for γ_3, and [4.5, 3.5, 2.5, 1.5] for γ_4, respectively.

We manipulate four experimental factors: the variances of indicators, correlations between indicators per component, distribution of indicators, and correlations among components. We consider the variances of indicators because this is a unique piece of information for the proposed method to use for creating components as compared to GSCA_{std}. The other three factors have been frequently considered in testing the performance of GSCA (e.g., Cho et al., 2022; Cho & Choi,

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

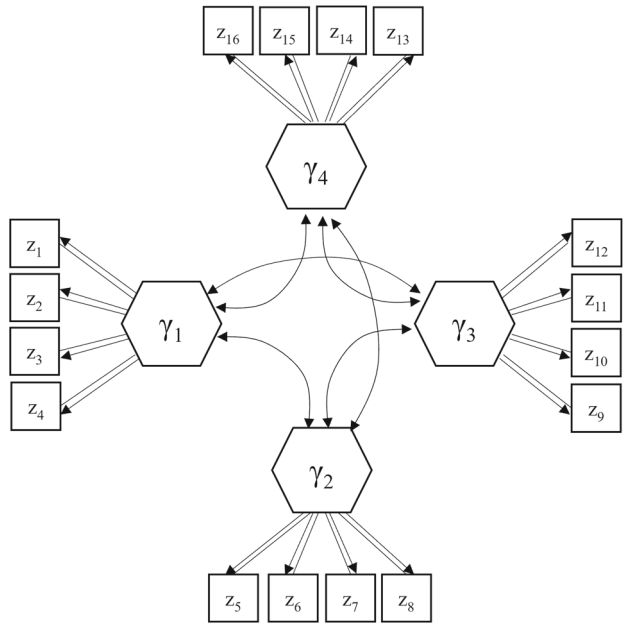

FIGURE 2.
The population GSCA$_{cvx}$ model used in the simulation study. Double-headed arrows represent correlations. All intercepts and error terms are omitted to make the figure concise.

TABLE 1.
Three conditions of the correlation patterns of four indicators per component in the simulation study.

	Condition 1				Condition 2				Condition 3			
	z_1	z_2	z_3	z_4	z_1	z_2	z_3	z_4	z_1	z_2	z_3	z_4
z_1	1				1				1			
z_2	.24	1			.50	1			.49	1		
z_3	.24	.20	1		.43	.47	1		.56	.74	1	
z_4	.17	.21	.13	1	.30	.23	.45	1	.66	.48	.69	1

2020; Hwang, Malhotra, et al., 2010b). Specifically, we consider three levels of the variances of indicators per component: [1, 1, 1, 1], [1 2, 3, 4], and [1, 4, 9, 16]. We take into account three correlation matrices of indicators per component, which are provided in Table 1 (Cho & Choi, 2020). We consider two distributions of indicators: normal and non-normal. The normal distribution has a skewness of 0 and a kurtosis of 3, whereas the non-normal distribution has a skewness of 1.25 and a kurtosis of 3.75 as in Hwang et al. (2010a). Lastly, we consider three levels of correlations among components (0, .2, and .4) as in Cho et al. (2022). In total, we consider 54 population GSCA models with convex components (3 levels of indicators' variances × 2 types of indicators' distribution × 3 levels of indicators' correlations × 3 levels of components' correlations).

Per population model, we consider five sample sizes (N = 100, 200, 400, 800, and 1500), for each of which 1000 samples are randomly generated from the multivariate distribution with the population mean vector and covariance matrix of indicators. The procedure of deriving the population covariance matrix of indicators from the prescribed parameter values of a population

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

GSCA$_{cvx}$ model is explained in Appendix 5. We apply GSCA$_{cvx}$[1] to each sample and obtain parameter estimates.

As parameter recovery measures, we empirically compute the absolute bias and root mean squared error (RMSE) of each parameter estimator. These measures are defined as

$$\text{Absolute bias} = \left|\mathrm{E}(\hat{\theta}) - \theta\right| \approx \left|\frac{1}{1000}(\sum_{i=1}^{1000}\hat{\theta}_i) - \theta\right|, \tag{20}$$

$$\text{RMSE} = \sqrt{\mathrm{E}(\hat{\theta} - \theta)^2} \approx \sqrt{\frac{1}{1000}\sum_{i=1}^{1000}(\hat{\theta}_i - \theta)^2}, \tag{21}$$

where θ is the value of each parameter, $\hat{\theta}$ is the estimator of θ, and $\hat{\theta}_i$ is the estimate of θ obtained from the ith sample. We focus here on reporting the average absolute bias and RMSE values of the estimators of weights, loadings, intercepts, component means, and component variances over the population models per sample size, as the sample size is the only factor that substantially influences the absolute bias and RMSE values of the estimators. The results for each population model are provided in Supplementary Material.

Table 2 shows the average absolute bias and RMSE values of the estimators per sample size. In all sample sizes, the absolute biases of the weight, loading, and component mean estimators are small and close to zero on average. For example, when $N = 100$, the average absolute biases of the weight, loading, and component mean estimators are .002, .022, and .008, respectively. They continue to decrease and approach zero when the sample size increases. The average RMSE values of the same estimators show a similar pattern. When $N = 100$, the average RMSE values are around .047, .134, and .216, respectively, and becomes close to zero as the sample size increases. The average absolute bias and RMSE values of the intercept and component variance estimators are relatively large, compared to those of the other parameter estimators in the same condition. For instance, when $N = 100$, the average absolute biases of the intercept and component variance estimators are .107 and .178, respectively, and their average RMSE values are .668 and .859, respectively. However, both of them also decrease with the sample size and become close to zero. Taken together, GSCA$_{cvx}$ estimators are empirically unbiased on average, improving their parameter recovery as the sample size increases.

4. Illustration with Empirical Data

To illustrate its empirical utility, we apply GSCA$_{cvx}$ to American customer satisfaction index (ACSI) data. The ACSI model (Fornell et al., 1996) is built on the established theories and has been used to produce index scores for customer satisfaction in the United States since 1994. The present ACSI data are comprised of 774 customers' responses for fourteen items: z_1 = expectation for overall quality, z_2 =expectation for reliability, z_3 = expectation for customization, z_4 = overall quality, z_5 = reliability, z_6 = customization, z_7 = price given quality, z_8 = quality given price, z_9 = perceived overall satisfaction, z_{10} = fulfilment of expectations, z_{11} = distance to the ideal, z_{12} = complaint behavior, z_{13} = repurchase intention, z_{14} = price tolerance. Twelve of the items (z_1,z_2, z_3, z_4, z_5, z_6, z_7, z_8, z_9, z_{10}, z_{11}, and z_{13}) are measured on a 10-point Likert scale (e.g., 1 = "very negative" and 10 = "very positive"). Within the interval [1, 5], a smaller point reflects a stronger negative response, whereas within the interval [6, 10], a larger point indicates a stronger positive response. On the other hand, z_{12} is a binary variable (1 = formally complained and 0

[1] The MATLAB code is available at https://osf.io/y75kg/?view_only=0d02aea6aaaa4aa29d405172544aae7d.

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

TABLE 2.
The average absolute bias and RMSE values of the estimators of weights, loadings, intercepts, component means, and component variances per sample size.

N	Absolute Bias					RMSE				
	Weights	Loadings	Intercepts	Component Means	Component Variances	Weights	Loadings	Intercepts	Component Means	Component Variances
100	0.002	0.022	0.107	0.008	0.178	0.047	0.134	0.668	0.216	0.859
200	0.001	0.011	0.052	0.005	0.078	0.030	0.092	0.460	0.146	0.515
400	0.001	0.005	0.026	0.003	0.037	0.021	0.064	0.319	0.102	0.346
800	0.000	0.003	0.013	0.002	0.018	0.014	0.045	0.223	0.072	0.240
1500	0.000	0.002	0.008	0.002	0.010	0.010	0.033	0.162	0.052	0.174

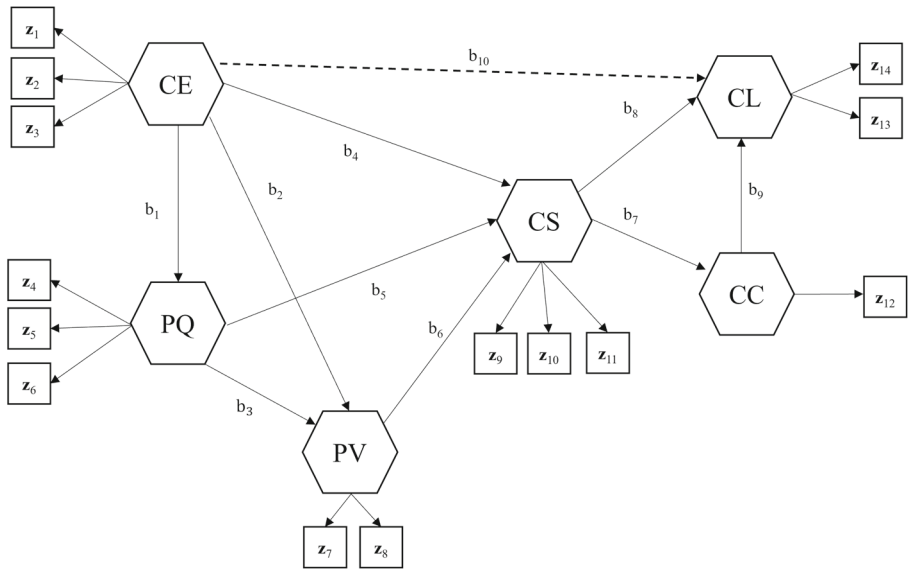

FIGURE 3.
The ACSI model. The dashed line labeled b_{10} signifies an incorrectly specified path coefficient. All weights and error terms are omitted to make the figure concise. CE = customer expectations, PQ = perceived quality, PV = perceived value, CS = customer satisfaction, CC = customer complaints, CL = customer loyalty.

= otherwise) and z_{14} is a composite of two price tolerance measures in different metrics, which is expressed as a percentage ranging from 0 to 50 (the higher, the more tolerant). The means, covariances, minimums, and maximums of the items are provided in Table 3. Refer to Fornell et al. (1996) for more detailed information on the items.

Figure 3 depicts the relationships among the six components and their indicators. The 14 items are used as composite indicators of the following six components: γ_1 = customer expectations (CE), γ_2 = perceived quality (PQ), γ_3 = perceived value (PV), γ_4 = customer satisfaction (CS), γ_5 = customer complaints (CC), and γ_6 = customer loyalty (CL). We represent all the constructs by convex components with unstandardized indicators except for the customer loyalty. As two indicators (z_{13} and z_{14}) for customer loyalty are not measured on the same scale, we set this component as a standardized one with the indicators standardized.

We use 4000 bootstrap samples for computing the standard error and 95% confidence interval of each parameter estimate. For comparison, we also apply GSCA_{std} to the same data and compute unstandardized weight estimates and unstandardized component scores based on the procedure discussed in Sect. 1. As customer satisfaction is the focal component in the ACSI model, we concentrate on interpreting the scores of customer satisfaction, its statistics, and the relevant model parameters.

The model fitted by GSCA_{cvx} shows $\text{FIT}^{\text{UD}} = .714$, indicating that the ACSI model accounts for 71.4% of the weighted total variance of all dependent variables in the model. It also provides GFI = .987 and SRMR = .022, pointing to an acceptable level of model fit (Cho et al., 2020). In addition, it provides that $\text{FIT}^{\text{UD}}_{\text{M}} = .802$ and $\text{FIT}^{\text{UD}}_{\text{S}} = .438$. This indicates that the component measurement model explains 80.2% of the weighted total variance of all dependent indicators, whereas the structural model explains 43.8% of the weighted total variance of all dependent components.

Table 4 provides the weight and loading estimates, and their standard errors and 95% confidence intervals obtained from GSCA_{cvx}, along with the intercept estimates in the measurement

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

TABLE 3.
Sample covariances (in upper triangular), correlations (in lower triangular), variances (in diagonal), means, minimums, and maximums of the fourteen indicators in the ACSI example.

	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}
z_1	5.81	3.61	2.71	2.92	2.91	2.10	1.96	2.79	3.01	2.73	3.23	–.14	2.62	11.78
z_2	.65	5.38	2.88	2.69	2.92	2.36	1.48	2.63	2.79	2.30	2.64	–.13	2.05	7.65
z_3	.43	.47	6.90	2.03	2.15	2.56	1.49	2.18	2.16	1.96	2.41	–.12	1.74	7.77
z_4	.53	.50	.33	5.31	4.77	3.76	2.52	4.15	4.87	4.35	4.42	–.32	3.94	16.86
z_5	.49	.51	.33	.84	6.11	3.98	2.48	4.37	5.23	4.80	5.01	–.36	4.29	17.20
z_6	.33	.39	.37	.62	.61	6.93	2.46	3.65	3.95	3.67	3.75	–.23	3.18	13.58
z_7	.31	.25	.22	.42	.39	.36	6.67	3.43	3.05	2.88	2.78	–.14	2.42	11.31
z_8	.47	.46	.33	.72	.71	.56	.53	6.18	4.74	4.42	4.70	–.29	3.60	16.60
z_9	.50	.48	.33	.85	.85	.60	.47	.77	6.19	5.06	5.09	–.35	4.55	19.55
z_{10}	.45	.40	.30	.75	.78	.56	.44	.71	.81	6.27	4.96	–.29	3.98	16.71
z_{11}	.51	.43	.35	.73	.77	.54	.41	.72	.78	.75	6.93	–.32	4.74	20.79
z_{12}	–.17	–.17	–.13	–.40	–.42	–.25	–.16	–.34	–.41	–.34	–.36	.12	–.33	-1.55
z_{13}	.37	.30	.22	.58	.58	.41	.32	.49	.62	.54	.61	–.33	8.79	35.21
z_{14}	.31	.21	.19	.47	.45	.33	.28	.43	.51	.43	.51	–.29	.76	241.11
Mean	7.34	7.75	6.67	7.66	7.59	7.39	5.96	7.12	7.59	6.82	6.76	.14	7.73	31.82
Min	1	1	1	1	1	1	1	1	1	1	1	0	1	0
Max	10	10	10	10	10	10	10	10	10	10	10	1	10	50

TABLE 4.
The weights, loading, and intercept estimates of the fourteen indicators in the ACSI model and their standard errors (SE) and 95% confidence intervals (CI) obtained from $GSCA_{cvx}$, along with the unstandardized weight estimates obtained from $GSCA_{std}$ ($\hat{\mathbf{W}}_{uni}$).

Indicator	Component	Weights			$\hat{\mathbf{W}}_{uni}$	Loadings			Intercepts($\hat{\mathbf{c}}_0$)
		Estimate	SE	95% CI		Estimate	SE	95% CI	
z_1	CE	.345	.013	[.320, .372]	.180	1.008	.025	[.957, 1.054]	.018
z_2		.337	.014	[.309, .366]	.188	.982	.027	[.925, 1.033]	.616
z_3		.317	.013	[.292, .343]	.128	1.011	.036	[.937, 1.077]	–.674
z_4	PQ	.387	.018	[.353, .425]	.184	.979	.013	[.953, 1.004]	.260
z_5		.342	.017	[.307, .374]	.170	1.043	.013	[1.018, 1.071]	–.303
z_6		.271	.007	[.257, .285]	.101	.976	.024	[.927, 1.021]	.012
z_7	PV	.404	.010	[.384, .423]	.154	.960	.020	[.921, .997]	–.427
z_8		.596	.010	[.577, .616]	.293	1.027	.013	[1.002, 1.051]	.289
z_9	CS	.422	.015	[.393, .454]	.188	1.004	.011	[.982, 1.025]	.433
z_{10}		.254	.013	[.229, .279]	.107	.965	.016	[.933, .996]	–.052
z_{11}		.324	.012	[.300, .348]	.131	1.022	.016	[.990, 1.053]	–.524
z_{12}	CL	1.000	.000	[1.000, 1.000]	2.909	1.000	.000	[1.000, 1.000]	.000
z_{13}	CC	.610	.015	[.579, .639]	.206	.956	.004	[.949, .963]	.000
z_{14}		.453	.016	[.424, .484]	.029	.920	.007	[.906, .932]	.000

model. The unstandardized weight estimates obtained from $GSCA_{std}$ are also provided for comparison. Overall, all the weight and loading estimates obtained from $GSCA_{cvx}$ are large and statistically significant, indicating that all the indicators contribute to forming their components, which in turn, explain the variances of their indicators well. Among the three indicators (z_9, z_{10}, and z_{11}) for customer satisfaction, z_9 (perceived overall satisfaction) are the largest contributor ($w_9 = .422$, SE = .015, 95% CI = [.393, .454]). This indicates that when each of the three indicators equally increases, leading to an increase in customer satisfaction, the contribution rate of $\mathbf{z}_9$ for the increase in customer satisfaction was 42.2%, which is greater than those of the two others ($z_{10} = 25.4\%$ and $z_{11} = 32.4\%$). Similarly, the unstandardized weight estimate of $\mathbf{z}_9$ obtained from $GSCA_{std}$ is the largest among the three ($w_9 = .188$, $w_{10} = .107$, and $w_{11} = .131$). In contrast, it is uncertain how to interpret the unstandardized weight estimates obtained from $GSCA_{std}$.

Table 5 presents the path coefficient estimates and their standard errors and 95% confidence intervals obtained from $GSCA_{cvx}$. Overall, the patterns of all the path coefficient estimates are consistent with those from previous studies (e.g., Hwang & Takane, 2014, Chapter 2). For instance, perceived quality and perceived value have statistically significant influences on customer satisfaction ($b_5 = .723$, SE = .033, 95% CI = [.659, .786]; $b_6 = .275$, SE = .035, 95% CI = [.204, .344]). Customer satisfaction have statistically significant effects on customer complaints ($b_7 = -.059$, SE = .006, 95% CI = [–.072, –.047]) and customer loyalty ($b_8 = .252$, SE = .015, 95% CI = [.222, .279]). Each individual path coefficient estimate is indicative of the expected change of the dependent component for a one-unit change in indicators of a predictor component. For instance, the estimate of the path coefficient, $b_8 = .252$, implies that a one-unit increase in z_9 (perceived overall satisfaction), z_{10} (expectation fulfillment), and z_{11} (distance to the ideal) would be associated with an increase of .252 unit in customer loyalty. The R^2 value is .331 for perceived quality, .511 for perceived value, .812 for customer satisfaction, .164 for customer complaints, and .404 for customer loyalty. Also, the intercept estimates for the dependent components in the same order as above are 3.014, .793, –.501, .558, and −1.756.

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

TABLE 5.
The path efficient estimates and their standard errors (SE) and 95% confidence intervals (CI) obtained from $GSCA_{cvx}$.

		Estimate	SE	95% CI
b_1	CE → PQ	.626	.037	[.551, .696]
b_2	CE → PV	.134	.039	[.058, .209]
b_3	PQ → PV	.646	.038	[.573, .721]
b_4	CE → CS	.045	.026	[–.005, .095]
b_5	PQ → CS	.723	.033	[.659, .786]
b_6	PV → CS	.275	.035	[.204, .344]
b_7	CS → CC	–.059	.006	[–.072, –.047]
b_8	CS → CL	.252	.015	[.222, .279]
b_9	CC→ CL	–.267	.104	[–.471, –.064]

TABLE 6.
The means, standard deviations (SD), and ranges of the unstandardized component scores estimated from $GSCA_{cvx}$ and $GSCA_{std}$. The last component (CL) is defined as a standardized component in $GSCA_{cvx}$.

	$GSCA_{cvx}$			$GSCA_{std}$		
	Mean	SD	Range	Mean	SD	Range
CE	7.265	2.014	[1.000, 10.000]	3.633	1.000	[.496, 4.961]
PQ	7.564	2.194	[1.000, 10.000]	3.444	1.000	[.455, 4.546]
PV	6.652	2.223	[1.000, 10.000]	3.008	1.000	[.448, 4.475]
CS	7.125	2.353	[1.000, 10.000]	3.037	1.000	[.425, 4.253]
CC	0.137	0.344	[0.000, 1.000]	.398	1.000	[.000, 2.909]
CL	0.000	1.000	[–2.311, .998]	2.518	1.000	[.206, 3.516]

Table 6 presents the estimated means, standard deviations, and ranges of unstandardized component scores obtained from $GSCA_{cvx}$ and $GSCA_{std}$. As expected, the individual scores of each convex component obtained from $GSCA_{cvx}$ are within the range of their indicators' scores. The individual scores of customer expectation, perceived quality, perceived value, and customer satisfaction all range from 1 to 10 and those of customer complaint were between 0 and 1, which are equivalent to the ranges of their indicators' measurement scales. The mean of customer satisfaction from $GSCA_{cvx}$ is 7.125, indicating that the average satisfaction level in the sample is moderately positive or equivalent to the satisfaction level of a customer whose indicator scores are all 7.125. This mean of customer satisfaction appears to be congruent with the means of its original indicators (7.585, 6.824, and 6.760). The standard deviation of customer satisfaction is 2.353, suggesting that the scores of customer satisfaction are somewhat widely spread out from the mean. This standard deviation value also seems to conform to those of its original indicators (2.489, 2.504, and 2.632).

On the contrary, unstandardized components' scores obtained from $GSCA_{std}$ are not always within the range of their indicators' scores. Some scores of customer expectation, perceived quality, perceived value, and customer satisfaction are smaller than 1, which is the minimum value of their indicators on the scale. Moreover, the means of unstandardized components are also far from those of their original indicators. For instance, the mean of customer satisfaction obtained from $GSCA_{std}$ is just 3.037, even though its indicators' means are around 7 as stated above. Thus, it is questionable whether the mean of customer satisfaction obtained from $GSCA_{std}$ can be a

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

good representation of the average level of customer satisfaction in the sample. Furthermore, all the standard deviations of unstandardized components are fixed to one, even though none of their indicators have standard deviations being around 1.

To illustrate the usage of OPE^{UD} as a model comparison criterion, we additionally contemplate two misspecified models of the ACSI model, while assuming the original ACSI model as the true model (denoted by Model 1). One misspecified model (Model 2) is an under-specified one, where a path coefficient (b_6) is omitted from Model 1. The other misspecified model (Model 3) is an over-specified one that includes an additional path coefficient from customer expectation to customer loyalty in Model 1, as displayed in Fig. 3. We apply GSCA_{cvx} to fit the three models to the data and compute their OPE^{UD} values based on 4000 bootstrap samples. Model 1 provides the smallest OPE^{UD} value (Model 1 = .2883, Model 2 = .2901, and Model 3 = .2887), indicating that the original ACSI model has the highest predictive generalizability among the three models. The OPE^{UD} value of Model 2 is larger than that of Model 1 (.2901 >.2883), suggesting that excluding a path coefficient (b_6) from Model 1 rather decreases the prediction accuracy of the model. On the other hand, the OPE^{UD} value of Model 3 is larger than that of Model 1 (.2887 >.2883), indicating that specifying an additional path coefficient (b_{10}) to Model 1 is not helpful to improve the predictive generalizability of the model.

5. Concluding Remarks

We proposed convex GSCA that can accommodate a new type of unstandardized components, named convex components. A convex component is defined as a convex combination of original indicators whose weights are all non-negative and summed up to one. Every individual score of a convex component is always within the range of its indicators' scores and can be interpreted as a construct's specific level of a person who has the same score for all its indicators as his/her component score. Moreover, the means and standard deviations of convex components are estimated along with other parameters through a single optimization procedure, which can also be interpreted in terms of indicators' scales. Thus, introducing convex components to the GSCA model will enhance the practical utility of component scores and their summary statistics, for instance, in investigating individuals' levels of a construct or comparing the average levels of a construct between groups.

We developed an alternating least squares (ALS) algorithm for estimating parameters of the convex GSCA model, which does not require standardizing blocks of indicators that have the same measurement scales within the blocks. The algorithm not only enables information on the variances of each block of indicators to be additionally utilized in parameter estimation, but also prevents indicators with small variances from influencing more heavily the construction of an unstandardized component than those with large variances. Furthermore, its objective function is partially scale-invariant, indicating that the minimum value of the objective function remains unchanged with a linear change in the measurement scale of each block of indicators, giving rise to the same weight estimates.

We evaluated the parameter recovery of the proposed method in a simulation study and further illustrated the merits of the proposed method via a real data analysis. In the simulation study, the proposed method empirically produced unbiased parameter estimates on average under nine GSCA models with convex components and its accuracy was further improved with large sample size. In the real data analysis, the patterns of the parameter estimates were consistent with those from previous studies, and the benefits of convex components were pronounced, compared to the unstandardized components obtained from the conventional ad-hoc procedure of rescaling weight estimates. Unlike these unstandardized components, convex components' weight estimates were interpretable, all their individual scores fell within the range of indicators' measurement scales or

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

their scores, and their estimated means and standard deviations were congruous with those of their indicators. Therefore, we are confident to recommend that researchers employ the method when they are interested in the GSCA model with unstandardized components of original indicators.

Note that as an anonymous reviewer pointed out, researchers may still want to consider standardizing observed variables that are measured on the same scale. We recommend considering this option only if researchers are not interested in unstandardized component scores. If researchers apply GSCA to estimate the scores of unstandardized components after standardizing indicators of the same scale, an indicator with a small variance can be assigned a relatively large unstandardized weight, leading to a potentially inflated influence of the indicator on the estimation of the component scores, as shown in Sect. 1. This issue does not occur when researchers keep the original scales of indicators and apply convex GSCA with convex components.

In future research, we may consider incorporating convex components into various extensions of GSCA, which deal with more complex analyses, for instance, those of involving higher-order components (Hwang & Takane, 2014, Chapter 3), missing observations (Hwang & Takane, 2014, Chapter 3), multilevel components (Hwang et al., 2007), components with categorical indicators (Hwang & Takane, 2010), component interaction terms (Hwang et al., 2021a, 2010a), or factors (Hwang et al., 2021b). Such additional extensions will improve the usefulness of GSCA, placing components on their indicators' scales while having their means and variances free parameters to be estimated along with others.

Acknowledgments

The authors are grateful to Claes Fornell for providing the ACSI data.

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

Appendix 1. A Proof of Disproportional Penalty Imposition on Indicators During the Minimization of the Objective Function (5)

Let $\mathbf{z} = \boldsymbol{\mu} + \boldsymbol{\Delta}_z \mathbf{z}_{std}$ is a random vector of original indicators, where $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, and $\boldsymbol{\Delta}_z$ denote a column mean vector of $\mathbf{z}$, the covariance matrix of $\mathbf{z}$, and a diagonal matrix consisting of each indicator's standard deviation, respectively. Let $\boldsymbol{\gamma}_{uni} \equiv \mathbf{W}'_{uni}\mathbf{z}$ denote a random vector of unstandardized components, where $\mathbf{W}_{uni}$ is a matrix of unstandardized weight parameters and $vecdiag(\mathbf{W}'_{uni}\boldsymbol{\Sigma}\mathbf{W}_{uni}) = \mathbf{1}_P$. Let $\mathbf{e}_{uni}$ is a random vector of prediction errors for $[\mathbf{z}; \boldsymbol{\gamma}_{uni}]$. Let $\boldsymbol{\Delta} \equiv$

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

$blkdiag(\boldsymbol{\Delta}_z, \mathbf{I}_P)$ is a diagonal matrix of penalty parameters for $\mathbf{e}_{uni}$, where $blkdiag()$ is an operator to convert input matrices into a block-diagonal matrix. Here, the penalty parameters refer to the parameters that rescale prediction error for each dependent variable in the model. Let $\mathbf{A}_{uni}$ denote a matrix of unstandardized loading and path coefficients in $GSCA_{std}$. Let $\mathbf{a}_{0,uni}$ denote a column vector of the unstandardized intercepts in $GSCA_{std}$. When $\mathbf{W}_{uni} = \boldsymbol{\Delta}_z^{-1}\mathbf{W}_{std}$, $\mathbf{A}_{uni} = \mathbf{A}_{std}\boldsymbol{\Delta}$, and $\mathbf{a}_{0,uni} = ([\mathbf{I}_J, \mathbf{W}_{uni}] - \mathbf{W}_{uni}\mathbf{A}_{uni})'\,\boldsymbol{\mu}$, (5) is equivalent to the following objective function,

$$\begin{aligned} &f_{uni}(\mathbf{W}_{uni}, \mathbf{A}_{uni}, \mathbf{a}_{0,uni}) \\ &\quad = tr(\boldsymbol{\Delta}^{-1}E(\mathbf{e}_{uni}\mathbf{e}'_{uni})\boldsymbol{\Delta}^{-1}) \end{aligned} \tag{A.1}$$

subject to $vecdiag(\mathbf{W}'_{uni}\boldsymbol{\Sigma}\mathbf{W}_{uni}) = \mathbf{1}_P$, which can be proved as follows.

$$\begin{aligned} &f_{std}(\mathbf{W}_{std}, \mathbf{A}_{std}) \\ &\quad = tr(E(\mathbf{e}_{std}\mathbf{e}_{std}{}')) \\ &\quad = E(SS(\mathbf{z}'_{std}\mathbf{V}_{std} - \mathbf{z}'_{std}\mathbf{W}_{std}\mathbf{A}_{std}))) \\ &\quad = E(SS((\mathbf{z}-\boldsymbol{\mu})'\boldsymbol{\Delta}_z{}^{-1}(\mathbf{V}_{std} - \mathbf{W}_{std}\mathbf{A}_{std}))) \\ &\quad = E(SS((\mathbf{z}-\boldsymbol{\mu})'([\boldsymbol{\Delta}_z^{-1}, \boldsymbol{\Delta}_z^{-1}\mathbf{W}_{std}] - \boldsymbol{\Delta}_z^{-1}\mathbf{W}_{std}\mathbf{A}_{std}\boldsymbol{\Delta}\boldsymbol{\Delta}^{-1}))) \\ &\quad = E(SS((\mathbf{z}-\boldsymbol{\mu})'([\mathbf{I}_J, \mathbf{W}_{uni}]\boldsymbol{\Delta}^{-1} - \mathbf{W}_{uni}\mathbf{A}_{uni}\boldsymbol{\Delta}^{-1}))) \\ &\quad = E(SS((\mathbf{z}-\boldsymbol{\mu})'(\mathbf{V}_{uni} - \mathbf{W}_{uni}\mathbf{A}_{uni})\boldsymbol{\Delta}^{-1})) \\ &\quad = E(SS((\mathbf{z}'\mathbf{V}_{uni} - (\mathbf{z}'\mathbf{W}_{uni}\mathbf{A}_{uni} + \mathbf{a}_{0,uni}{}'))\boldsymbol{\Delta}^{-1})), \\ &\quad = tr(\boldsymbol{\Delta}^{-1}E(\mathbf{e}_{uni}\mathbf{e}_{uni}{}')\boldsymbol{\Delta}^{-1}) \\ &\quad = f_{std}(\mathbf{W}_{uni}, \mathbf{A}_{uni}, \mathbf{a}_{0,uni}), \end{aligned} \tag{A.2}$$

where $\mathbf{V}_{uni} \equiv [\mathbf{I}_J, \mathbf{W}_{uni}]$. The equivalence between (5) and (A.1) indicates that $GSCA_{std}$'s parameters are actually the standardized versions of $\mathbf{W}_{uni}$ and $\mathbf{A}_{uni}$ that are obtained by minimizing the sum of penalized error variances for the original indicators and unstandardized components. While minimizing (A.1), a relatively large penalty will be imposed on an indicator with a relatively large variance, potentially inflating the influence of an indicator with a small variance on $GSCA_{std}$'s parameter estimation.

Appendix 2. Proofs of the Six Propositions that Characterize a Convex Component

Let us suppose that the pth component (γ_p) is a convex component defined with J_p indicators ($\mathbf{z}_p$), indicating that the sum of weights assigned to the indicators is equal to one (i.e., $\mathbf{1}'_{J_p}\mathbf{w}_p = 1$) and all the weights are non-negative (i.e., $\mathbf{w}_p \geq \mathbf{0}_{Jp\times 1}$). Let $z_{i,p}$ denote the ith random variable in $\mathbf{z}_p$ ($i = 1, 2, \cdots, J_p$), which takes a value in $z_{i,p} \subset \mathbb{R}$. Let $w_{i,p}$ denote the ith element of $\mathbf{w}_p$ ($i = 1, 2, \cdots, J_p$).

Proposition 1. *A convex component has scores within the range of its indicators' scores.*

Proof. Let $m_1 \equiv inf\{inf\ z_{1,p}, inf\ z_{2,p}, \cdots, inf\ z_{J_p,p}\}$ and $m_2 \equiv sup\{sup\ z_{1,p}, sup\ z_{2,p}, \cdots, sup\ z_{J_p,p}\}$. Then, $m_1 = m_1\sum_{i=1}^{J_p} w_{i,p} = \sum_{i=1}^{J_p} m_1 w_{i,p} \leq \gamma_p = \sum_{i=1}^{J_p} z_{i,p}w_{i,p} \leq \sum_{i=1}^{J_p} m_2 w_{i,p} = m_2\sum_{i=1}^{J_p} w_{i,p} = m_2$.

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

Proposition 2. *Each score of a convex component corresponds to a component score of an individual whose scores for indicators are all the same as the component score.*

Proof. Let $g \in \mathcal{G}_p$ denote a value of γ_p, where $\mathcal{G}_p \subset \mathbb{R}$ is the set of all possible values γ_p can take in $\mathbb{R}$. If $\mathbf{z}_p = [g, g, \cdots, g]' = g\mathbf{1}_{J_p}$, then $\gamma_p = \mathbf{w}_p{}' g\mathbf{1}_{J_p} = g$.

Proposition 3. *The mean of a convex component is not fixed to zero but is determined by weights within the range of its indicators' means.*

Proof. $E(\gamma_p) = \mathbf{w}_p{}' E(\mathbf{z}_p) = \mathbf{w}_p{}' \boldsymbol{\mu}_p$. Thus, $E(\gamma_p)$ varies depending on $\mathbf{w}_p$ unless $\boldsymbol{\mu}_p = \mathbf{0}$. Let $\mu_{i,p}$ denote the ith element of $\mathbf{u}_p$. Let $m_3 \equiv inf\{\mu_{1,p}, \mu_{2,p}, \cdots, \mu_{J_p,p}\}$ and $m_4 \equiv sup\{\mu_{1,p}, \mu_{2,p}, \cdots, \mu_{J_p,p}\}$. Then, $m_3 = m_3 \sum_{i=1}^{J_p} w_{i,p} = \sum_{i=1}^{J_p} m_3 w_{i,p} \le E(\gamma_p) = \mathbf{w}_p{}' \boldsymbol{\mu}_p = \sum_{i=1}^{J_p} \mu_{i,p} w_{i,p} \le \sum_{i=1}^{J_p} m_4 w_{i,p} = m_4 \sum_{i=1}^{J_p} w_{i,p} = m_4$.

Proposition 4. *The standard deviation of a convex component is not fixed to one but is determined by weights within the range from 0 to the maximum standard deviation of its indicators.*

Proof. $\mathrm{var}(\gamma_p)^{1/2} = (\mathbf{w}_p{}' \mathrm{var}(\mathbf{z}_p)\mathbf{w}_p)^{1/2} = (\mathbf{w}_p{}' \boldsymbol{\Sigma}_p \mathbf{w}_p)^{1/2}$, indicating that the standard deviation of γ_p depends on $\mathbf{w}_p$. Let $\sigma_{k,l,p}$ denote the (k,l)th element of $\boldsymbol{\Sigma}_p$. Let $m_5 \equiv sup\{\sigma_{1,1p}, \sigma_{2,2,p}, \cdots, \sigma_{J_p,J_p,p}\}$. Then, $\mathrm{var}(\gamma_p)^{1/2} = (\mathbf{w}_p{}' \boldsymbol{\Sigma}_p \mathbf{w}_p)^{1/2} = (\sum_{k=1}^{J_p} \sum_{l=1}^{J_p} w_{k,p} w_{l,p} \sigma_{k,l,p})^{1/2} \le (\sum_{k=1}^{J_p} \sum_{l=1}^{J_p} w_{k,p} w_{l,p} m_5)^{1/2} = (m_5 \sum_{k=1}^{J_p} \sum_{l=1}^{J_p} w_{k,p} w_{l,p})^{1/2} = (m_5 \sum_{k=1}^{J_p} w_{k,p} (\sum_{l=1}^{J_p} w_{l,p}))^{1/2} = (m_5 \sum_{k=1}^{J_p} w_{k,p})^{1/2} = m_5^{1/2}$. Therefore, $0 < var(\gamma_p)^{1/2} \le m_5^{1/2}$.

Proposition 5. *Given a linearly independent set of indicators' scores, a set of convex component scores has a unique set of weights that are nonnegative and summed up to one.*

Proof. Let $\mathbf{D}_p = [\mathbf{d}_{\cdot 1,p}, \mathbf{d}_{\cdot 2,p}, \cdots, \mathbf{d}_{\cdot J_p,p}]$ denote a N by J_p data matrix of $\mathbf{z}_p$, where N is the total number of individuals and $\mathbf{d}_{\cdot i,p}$ is the score set of $z_{i,p}$ $(i = 1, 2, \cdots, J_p)$. Then, the score set of the pth convex component for N individuals, denoted by $\mathbf{g}_{\cdot p}$, can be expressed as $\mathbf{g}_{\cdot p} = \mathbf{D}_p \mathbf{w}_p$. Suppose that there exists a different set of weights, $\mathbf{w}_{p+} = [w_{1,p+}, w_{2,p+}, \cdots, w_{J_p,p+}]'$, such that $\mathbf{g}_{\cdot p} = \mathbf{D}_p \mathbf{w}_{p+}{}'$ and $\mathbf{w}_{p+} \ne \mathbf{w}_p$. Then, $0 = \mathbf{g}_{\cdot p-} \mathbf{g}_{\cdot p} = \mathbf{D}_p \mathbf{w}_p - \mathbf{D}_p \mathbf{w}_{p+} = \mathbf{D}_p(\mathbf{w}_p - \mathbf{w}_{p+}) = \mathbf{d}_{\cdot 1,p}(w_{1,p} - w_{1,p+}) + \mathbf{d}_{\cdot 2,p}(w_{2,p} - w_{2,p+}) + \cdots + \mathbf{d}_{\cdot J_p,p}(w_{J_p,p} - w_{J_p,p+})$. By the assumption that $\{\mathbf{d}_{\cdot 1,p}, \mathbf{d}_{\cdot 2,p}, \cdots, \mathbf{d}_{\cdot J_p,p}\}$ is linearly independent, $w_{1,p} = w_{1,p+}$, $w_2, p = w_{2,p+}, \cdots, w_{J_p,p} = w_{J_p,p+}$, which contradicts the assumption. By the definition of a convex component, $\mathbf{w}_p{}' \mathbf{1}_{J_p} = 1$ and $\mathbf{w}_p \ge \mathbf{0}_{J_p \times 1}$.

Proposition 6. *The path coefficient of a convex component on an outcome variable indicates the expected amount of change in the outcome variable for a unit change in each indicator of the convex component while holding other variables fixed.*

Proof. Let $\gamma_q = b_{0,q} + b_{p,q}\gamma_p + \boldsymbol{\alpha}'\mathbf{x} + \zeta_q$ denote a structural model equation of the outcome variable γ_q on γ_p and a vector of covariates $\mathbf{x}$ for γ_q, where $b_{0,q}$ is an intercept for γ_q, $b_{p,q}$ is the path coefficient from γ_p to γ_q, $\boldsymbol{\alpha}$ is a vector of path coefficients of $\mathbf{x}$, and ζ_q is an error term for γ_q. As this model equation can be re-expressed as $\gamma_q = b_{0,q} + b_{p,q}\, \mathbf{w}_p{}' \mathbf{z}_p + \boldsymbol{\alpha}'\mathbf{x} + \zeta_q$, an expected change of γ_q for a one-unit change in every element of $\mathbf{z}_p$ with the values of $\mathbf{x}$ fixed can be expressed as $E((b_{0,q} + b_{p,q}\mathbf{w}_p{}'(\mathbf{z}_p + \mathbf{1}_p) + \boldsymbol{\alpha}'\mathbf{x} + \zeta_q) - (b_{0,q} + b_{p,q}\mathbf{w}_p{}'\mathbf{z}_p + \boldsymbol{\alpha}'\mathbf{x} + \zeta_q))$, which is equivalent to $E(b_{p,q}\mathbf{w}_p{}'(\mathbf{z}_p + \mathbf{1}_p) - b_{p,q}\mathbf{w}_p{}'\mathbf{z}_p) = E(b_{p,q}(\mathbf{w}_p{}'\mathbf{1}_p)) = E(b_{p,q}) = b_{p,q}$.

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

Appendix 3. A Proof that the Optimization Function of Convex GSCA is Partially Scale-Invariant

Suppose that for each block of indicators that are on the same scale, the measurement scales are linearly transformed arbitrarily. Let $\mathbf{z}_{new} = \mathbf{\Omega}_z(\mathbf{z} + \boldsymbol{\lambda})$ denote a vector of rescaled indicators, where $\boldsymbol{\lambda}$ is a J by 1 constant vector for relocation and $\mathbf{\Omega}_z$ is a diagonal matrix for scalar multiplication for each indicator. This linear transformation of the measurement scales of indicators does not change the minimum value of the objective function (13) and the corresponding weight values, which can be proven as follows. Let $\boldsymbol{\gamma}_{new} = \mathbf{W}'\mathbf{z}_{new}$ denote a vector of components defined with rescaled indicators. Let $\mathbf{e}_{new}$ denote a vector of prediction errors for $[\mathbf{z}_{new}; \boldsymbol{\gamma}_{new}]$. Let $\mathbf{A}_{new}$ denote a matrix of unstandardized loading and path coefficients for $\mathbf{z}_{new}$. Let $\mathbf{a}_{0,new}$ denote a vector of unstandardized intercepts for $\mathbf{z}_{new}$. Let $\mathbf{O}_{new}$ denote a diagonal matrix of penalty parameters for prediction errors given $\mathbf{z}_{new}$. Let ω_p is the scalar multiplier that is applied the pth block of indicators. Let $\mathbf{\Omega}_\gamma \equiv blkdiag(\omega_1, \omega_2, \cdots, \omega_P)$ and $\mathbf{\Omega} \equiv blkdiag(\mathbf{\Omega}_z, \mathbf{\Omega}_\gamma)$.

When $\mathbf{A}_{new} = \mathbf{\Omega}_\gamma^{-1}\mathbf{A}\mathbf{\Omega}$, $\mathbf{a}_{0,new} = (\boldsymbol{\lambda}'(\mathbf{V} - \mathbf{WA}) + \mathbf{a}_0)\mathbf{\Omega}$, and $\mathbf{O}_{new} \equiv \mathbf{\Omega}^{-1}\mathbf{O}$, the objective function (13) can be re-written as

$$
\begin{aligned}
&f_{cvx}(\mathbf{W}, \mathbf{A}, \mathbf{a}_0) \\
&= tr(\mathbf{O}E(\mathbf{e}\mathbf{e}')\mathbf{O}) \\
&= E(SS((\mathbf{z}'\mathbf{V} - (\mathbf{z}'\mathbf{WA} + \mathbf{a}_0{}'))\mathbf{O}) \\
&= E(SS((\mathbf{z}'(\mathbf{V} - \mathbf{WA}) - \mathbf{a}_0{}')\mathbf{O}) \\
&= E(SS((\mathbf{z}' + \boldsymbol{\lambda}' - \boldsymbol{\lambda}')\mathbf{\Omega}_z\mathbf{\Omega}_z^{-1}(\mathbf{V} - \mathbf{WA}) - \mathbf{a}_0{}')\mathbf{\Omega}\mathbf{\Omega}^{-1}\mathbf{O}) \\
&= E(SS((\mathbf{z} + \boldsymbol{\lambda})'\mathbf{\Omega}_z(\mathbf{\Omega}_z^{-1}\mathbf{V} - \mathbf{\Omega}_z^{-1}\mathbf{WA})\mathbf{\Omega} - (\boldsymbol{\lambda}'(\mathbf{V} - \mathbf{WA}) + \mathbf{a}_0{}')\mathbf{\Omega})\mathbf{\Omega}^{-1}\mathbf{O}) \\
&= E(SS(\mathbf{z}'_{new}([\mathbf{\Omega}_z^{-1}, \mathbf{\Omega}_z^{-1}\mathbf{W}]\mathbf{\Omega} - \mathbf{\Omega}_z^{-1}\mathbf{WA}\mathbf{\Omega}) - \mathbf{a}_{0,new}{}'))\mathbf{O}_{new}) \\
&= E(SS(\mathbf{z}_{new}{}'([\mathbf{I}_J, \mathbf{W}]\mathbf{\Omega}^{-1}\mathbf{\Omega} - \mathbf{W}\mathbf{\Omega}_\gamma^{-1}\mathbf{A}\mathbf{\Omega}) - \mathbf{a}_{0,new}{}'))\mathbf{O}_{new}) \\
&= E(SS(\mathbf{z}_{new}{}'(\mathbf{V} - \mathbf{WA}_{new}) - \mathbf{a}_{0,new}{}'))\mathbf{O}_{new}) \\
&= E(SS(\mathbf{z}_{new}{}'\mathbf{V} - (\mathbf{z}_{new}{}'\mathbf{WA}_{new} + \mathbf{a}_{0,new}{}'))\mathbf{O}_{new}) \\
&= tr(\mathbf{O}_{new}E(\mathbf{e}_{new}\mathbf{e}_{new}{}')\mathbf{O}_{new}) \\
&= f_{cvx}(\mathbf{W}, \mathbf{A}_{new}, \mathbf{a}_{0,new}).
\end{aligned}
\tag{A.3}
$$

The seventh equality in (A.3) holds because $\mathbf{\Omega}_z^{-1}\mathbf{W} = \mathbf{W}\mathbf{\Omega}_\gamma^{-1}$.

Appendix 4. A Description of GSCA $_{cvx}$'s ALS Algorithm

The objective function (13) can be re-written as

$$
\begin{aligned}
&f_{cvx}(\mathbf{W}, \mathbf{A}, \mathbf{a}_0) \\
&= E(SS((\mathbf{z}'(\mathbf{V} - \mathbf{WA}) - \mathbf{a}_0{}')\mathbf{O})) \\
&= E(SS((\mathbf{z} - \boldsymbol{\mu})'\mathbf{LO} - (\mathbf{a}_0' - \boldsymbol{\mu}'\mathbf{L})\mathbf{O})) \\
&= E(tr(((\mathbf{z} - \boldsymbol{\mu})'\mathbf{LO} - (\mathbf{a}_0{}' - \boldsymbol{\mu}'\mathbf{L})\mathbf{O})'((\mathbf{z} - \boldsymbol{\mu})'\mathbf{LO} - (\mathbf{a}_0' - \boldsymbol{\mu}'\mathbf{L})\mathbf{O}))) \\
&= tr(\mathbf{OL}'\mathbf{\Sigma}\mathbf{LO}) - 2tr(\mathbf{O}(\mathbf{a}_0' - \boldsymbol{\mu}'\mathbf{L})'E(\mathbf{z} - \boldsymbol{\mu})'\mathbf{LO}) + SS((\mathbf{a}_0' - \boldsymbol{\mu}'\mathbf{L})\mathbf{O})) \\
&= tr(\mathbf{OL}'\mathbf{\Sigma}\mathbf{LO}) + SS((\mathbf{a}_0' - \boldsymbol{\mu}'\mathbf{L})\mathbf{O})),
\end{aligned}
\tag{A.4}
$$

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

where $\mathbf{L} \equiv \mathbf{V} - \mathbf{WA}$. Let $\mathbf{S}$ denote the positive definite sample covariance matrix of indicators and $\widehat{\mathbf{O}}$ denote the sample analogy of $\mathbf{O}$. As $\boldsymbol{\Sigma}$, $\boldsymbol{\mu}$, and $\mathbf{O}$ are typically not available, GSCA_{cvx} replaces $\boldsymbol{\Sigma}$, $\boldsymbol{\mu}$, and $\mathbf{O}$ in (A.4) with $\mathbf{S}$, $\hat{\boldsymbol{\mu}}$, and $\widehat{\mathbf{O}}$, respectively, as follows,

$$\begin{aligned} & f^*_{cvx}(\mathbf{W}, \mathbf{A}, \mathbf{a}_0) \\ & \quad = tr(\widehat{\mathbf{O}}\mathbf{L}'\mathbf{S}\mathbf{L}\widehat{\mathbf{O}}) + SS((\mathbf{a}_0{}' - \hat{\boldsymbol{\mu}}'\mathbf{L})\widehat{\mathbf{O}}), \end{aligned} \tag{A.5}$$

and applies the ALS algorithm to find the minimum point of (A.5) with respect to $\mathbf{W}$, $\mathbf{A}$, and $\mathbf{a_0}$ subject to $\mathbf{w}_p{}'\mathbf{S}_p\mathbf{w}_p = 1$ or $\mathbf{1}_{J_p}{}'\mathbf{w}_p = 1$ ($p = 1, 2, \cdots, P$), where $\mathbf{S}_p$ is an J_p by J_p sample covariance matrix of $\mathbf{z}_p$.

The proposed ALS algorithm begins by assigning random initial values to $\mathbf{A}$ and repeats two steps until convergence. In the first step, $\mathbf{W}$ and $\mathbf{a}_0$ are updated with $\mathbf{A}$ fixed. By solving $\frac{1}{2}\frac{\partial f^*_{cvx}}{\partial \mathbf{a}_0} = 0$, the least square estimates of $\mathbf{a}_0$ given $\mathbf{W}$ and $\mathbf{A}$ can be obtained as

$$\hat{\mathbf{a}}_0 = \hat{\boldsymbol{\mu}}(\mathbf{V} - \mathbf{WA}) \tag{A.6}$$

This implies that the least squares estimate of $\mathbf{a}_0$ can be expressed as a function of $\mathbf{W}$ and $\mathbf{A}$. In other words, if we can find the least square estimate of $\mathbf{W}$ given $\mathbf{A}$ under the constraint (A.6), we can obtain $\hat{\mathbf{a}}_0$ as well by (A.6). Inserting (A.6) into (A.5) makes (A.5) be simplified as

$$\begin{aligned} & f^*_{cvx}(\mathbf{W}, \mathbf{A}) \\ & \quad = tr(\widehat{\mathbf{O}}(\mathbf{V} - \mathbf{WA})'\mathbf{S}(\mathbf{V} - \mathbf{WA})\widehat{\mathbf{O}}) \\ & \quad = N^{-1}SS(\mathbf{D}_{ct}(\mathbf{V} - \mathbf{WA})\widehat{\mathbf{O}}), \end{aligned} \tag{A.7}$$

where $\mathbf{D}_{ct} \equiv \mathbf{D} - \mathbf{1}_N\hat{\boldsymbol{\mu}}'$. Let $\widehat{\mathbf{O}}_\text{Y}$ denote a T by T_Y matrix consisting of all nonzero columns of $\widehat{\mathbf{O}}$, where $T \equiv P + J$ and T_Y is the number of dependent variables in the model. Let $\mathbf{I}_0 \equiv [\mathbf{I}_J, \mathbf{0}_{J\times P}]$ and $\mathbf{A}_\text{I} \equiv \mathbf{A} - [\mathbf{0}_{P\times J}, \mathbf{I}_P]$. Let $\mathbf{W}_{-p}$ denote a J by $(P-1)$ matrix formed by the columns of $\mathbf{W}$ except for its pth column. Let $\mathbf{A}_{\text{I},-p}$ denote a $(P-1)$ by T matrix formed by the rows of $\mathbf{A}_\text{I}$ except for its pth row and $\mathbf{a}_p$ denote a row vector whose entries are the non-zero elements of the pth row of $\mathbf{A}_{\text{I},-p}$ corresponding to $\mathbf{w}_p$. Let $vec()$ denote an operator that returns a column vector obtained by stacking the columns of input matrix vertically. Given $\mathbf{A}$, (A.7) can be re-expressed as

$$\begin{aligned} & f^*_{cvx}(\mathbf{w}_p; \mathbf{A}, \mathbf{W}_{-p}) \\ & \quad = N^{-1}SS(\mathbf{D}_{ct}(\mathbf{V} - \mathbf{WA})\widehat{\mathbf{O}}_\text{Y}) \\ & \quad = N^{-1}SS(\mathbf{D}_{ct}(\left[\mathbf{I}_J, \mathbf{0}_{J\times P}\right] + [\mathbf{0}_{J\times J}, \mathbf{W}] - \mathbf{WA})\widehat{\mathbf{O}}_\text{Y}) \\ & \quad = N^{-1}SS(\mathbf{D}_{ct}(\mathbf{I}_0 - \mathbf{WA}_\text{I})\widehat{\mathbf{O}}_\text{Y}) \\ & \quad = N^{-1}SS(\mathbf{D}_{ct}(\mathbf{I}_0 - \mathbf{W}_{-p}\mathbf{A}_{\text{I},-p} - \mathbf{w}_p\mathbf{a}_p)\widehat{\mathbf{O}}_\text{Y}) \\ & \quad = N^{-1}SS(vec(\mathbf{D}_{ct}(\mathbf{I}_0 - \mathbf{W}_{-p}\mathbf{A}_{\text{I},-p})\widehat{\mathbf{O}}_\text{Y}) - ((\mathbf{a}_p\widehat{\mathbf{O}}_\text{Y})' \otimes \mathbf{D}_{ct})\mathbf{w}_p)) \\ & \quad = N^{-1}SS(\boldsymbol{\psi}_1 - \boldsymbol{\Xi}_1\mathbf{w}_p)), \end{aligned} \tag{A.8}$$

where $\boldsymbol{\psi}_1 \equiv vec(\mathbf{D}_{ct}(\mathbf{I}_0 - \mathbf{W}_{-p}\mathbf{A}_{\text{I},-p})\widehat{\mathbf{O}}_\text{Y})$ and $\boldsymbol{\Xi}_1 \equiv (\mathbf{a}_p\widehat{\mathbf{O}}_\text{Y})' \otimes \mathbf{D}_{ct}$.

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

If γ_p is standardized component, the unstandardized least square estimate of $\mathbf{w}_p$ is obtained by

$$\widehat{\mathbf{w}}_{p*} = (\boldsymbol{\Xi}_1' \boldsymbol{\Xi}_1)^{-1} \boldsymbol{\Xi}_1 \boldsymbol{\psi}_1. \tag{A.9}$$

Then, the standardized least square estimate of $\mathbf{w}_p$ is obtained by $\hat{\mathbf{w}}_p = (\hat{\mathbf{w}}_{p*}{}' \mathbf{S}_p \hat{\mathbf{w}}_{p*})^{1/2}$ such that $\hat{\mathbf{w}}_p$ can satisfy $\hat{\mathbf{w}}_p{}' \mathbf{S}_p \hat{\mathbf{w}}_p = 1$. If every element of $\hat{\mathbf{w}}_p$ is forced to be positive, finding $\hat{\mathbf{w}}_p$ that minimizes (A.8) becomes a well-known *nonnegative least squares problem* (NNLS; Lawson & Hanson, 1974, Chapter 23), which should be solved numerically. For instance, the function *lsqnonneg* in MATLAB or the *nnls* package in R can be utilized under this condition.

If γ_p is a convex component, the ALS algorithm finds the solution for (A.8) subject to $\mathbf{1}_{Jp}{}' \mathbf{w}_p = 1$. This minimization is a linearly constrained least squares problem (Boyd & Vandenberghe, 2018, Chapter 16). As the product of the ranks of two matrices equals to the rank of the Kronecker product of the two matrices and $\mathbf{a}_p \hat{\mathbf{O}}_Y$ has one row, $(\mathbf{a}_p \hat{\mathbf{O}}_Y)' \otimes \mathbf{D}_{ct}$ has linearly independent columns, thereby having the columns of $\begin{bmatrix} \boldsymbol{\Xi}_1 \\ \mathbf{1}_{Jp}' \end{bmatrix}$ are also linearly independent. Thus, there exists $\boldsymbol{\delta}$ satisfying

$$\begin{bmatrix} \boldsymbol{\Xi}_1{}' \boldsymbol{\Xi}_1 & \mathbf{1}_{Jp} \\ \mathbf{1}_{Jp}{}' & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w}_p \\ \boldsymbol{\delta} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Xi}_1{}' \boldsymbol{\Psi}_1 \\ 1 \end{bmatrix}, \tag{A.10}$$

where $\begin{bmatrix} \boldsymbol{\Xi}_1{}' \boldsymbol{\Xi}_1 & \mathbf{1}_{Jp} \\ \mathbf{1}_{Jp}{}' & 0 \end{bmatrix}$ is invertible. Let $\boldsymbol{\Xi}_2 \equiv \begin{bmatrix} \boldsymbol{\Xi}_1{}' \boldsymbol{\Xi}_1 & \mathbf{1}_{Jp} \\ \mathbf{1}_{Jp}{}' & 0 \end{bmatrix}$ and $\boldsymbol{\psi}_2 \equiv \begin{bmatrix} \boldsymbol{\Xi}_1{}' \boldsymbol{\Psi}_1 \\ 1 \end{bmatrix}$. Then, $\hat{\mathbf{w}}_p$ can be obtained by the first J_p entries of $\boldsymbol{\Xi}_2^{-1} \boldsymbol{\psi}_2$, from which $\hat{\mathbf{W}}$ is updated. If $\mathbf{w}_p$ is forced to be positive, minimizing (A.8) becomes a quadratic programming problem with a linear constraint and an inequality constraint (e.g., Floudas & Visweswaran, 1995; Frank & Wolfe, 1956). This problem does not have closed-form solution. Instead, it can be solved numerically via interior point methods (e.g., Altman & Gondzio, 1999; Vanderbei & Carpenter, 1993). For instance, the function *lsqlin* in MATLAB or the *quadprog* package in R can be utilized to minimize (A.8) numerically. This process repeats for every $\mathbf{w}_p$ ($p = 1, 2, \cdots, P$). Then $\hat{\mathbf{a}}_0$ is updated by (A.6).

In the second step, $\mathbf{A}$ and $\mathbf{a_0}$ are updated with $\mathbf{W}$ fixed. Given $\mathbf{W}$, (A.8) can be re-expressed as

$$\begin{aligned} & f_{cvx}^*(\mathbf{A}; \mathbf{W}) \\ & \quad = N^{-1} SS(vec(\mathbf{D}_{ct} \mathbf{V} \hat{\mathbf{O}}_Y) - (\hat{\mathbf{O}}_Y{}' \otimes (\mathbf{D}_{ct} \mathbf{W})) vec(\mathbf{A})) \\ & \quad = N^{-1} SS(vec(\mathbf{D}_{ct} \mathbf{V} \hat{\mathbf{O}}_Y) - \boldsymbol{\Xi}_3 \boldsymbol{\rho}), \end{aligned} \tag{A.11}$$

where $\boldsymbol{\Xi}_3$ is the matrix formed by the columns of$(\hat{\mathbf{O}}_Y{}' \otimes (\mathbf{D}_{ct} \mathbf{W}))$ corresponding to the nonzero elements in $vec(\mathbf{A})$, and $\boldsymbol{\rho}$ is the column vector of the nonzero elements of $vec(\mathbf{A})$. Then, the value of $\boldsymbol{\rho}$ that minimizes (A.11) is obtained by

$$\boldsymbol{\rho} \quad = (\boldsymbol{\Xi}_3' \boldsymbol{\Xi}_3)^{-1} \boldsymbol{\Xi}_3' vec(\mathbf{D}_{ct} \mathbf{V} \hat{\mathbf{O}}_Y) \tag{A.12}$$

from which the non-zero elements of $\hat{\mathbf{A}}$ are updated. Then $\hat{\mathbf{a}}_0$ is updated by (A.6).

https://doi.org/10.1007/s11336-023-09944-3 Published online by Cambridge University Press

Appendix 5. A Procedure for Deriving the Population Covariance Matrix of Indicators from the Prescribed Parameter Values of the GSCA Model with Convex Components

The proposed procedure imitates the one suggested by Cho and Choi's (2020) one, while simply replacing standardized components and correlation matrix of indicators with convex components and covariance matrix of indicators, respectively. Let $\mathbf{c}_p$ is a J_p by 1 vector of loadings for $\mathbf{z}_p$ Let $\boldsymbol{\xi}_p$ is a J_p by 1 vector of error terms for $\mathbf{z}_p$. Let $\boldsymbol{\Phi}_{std}$ denote a P by P correlation matrix of components. Let $\boldsymbol{\Delta}$ denote a P by P diagonal matrix whose pth entry is the standard deviation of γ_p, denoted by ϕ_p. Let $\boldsymbol{\Theta} = blkdiag(\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2, \cdots, \boldsymbol{\Theta}_P)$ denote a J by J covariance matrix of errors in the measurement model, where $\boldsymbol{\Theta}_p$ is a J_p by J_p covariance matrix of errors for the pth block of indicators. Let $\boldsymbol{\tau} = [\tau_1, \tau_2, \cdots, \tau_P{}']$ denote a P by 1 vector of component means.

Given the prescribed values of $\boldsymbol{\Sigma}_p$, $\boldsymbol{\mu}_p$, and $\boldsymbol{\Phi}_{std}(p = 1, 2, \cdots, P)$, $\mathbf{w}_p$ is obtained by $\mathbf{w}_p = (\boldsymbol{\Sigma}_p^{-1/2}\mathbf{u}_{1p}) / (\mathbf{1}_p{}' \boldsymbol{\Sigma}_p^{-1/2}\mathbf{u}_{1p})$, where $\mathbf{u}_{1p}$ is the eigenvector corresponding to the largest eigenvalue of $\boldsymbol{\Sigma}_p$, indicating that $\mathbf{w}_p$ maximizes the sum of explained variances of $\mathbf{z}_p$ given $\boldsymbol{\Sigma}_p$ subject to $\mathbf{1}_p{}'\mathbf{w}_p = 1$. Then, τ_p and ϕ_p are calculated as $\tau_p = \mathbf{w}_p{}' \boldsymbol{\mu}_p$ and $\phi_p = \mathbf{w}_p{}' \boldsymbol{\Sigma}_p\mathbf{w}_p$, respectively, based on which $\mathbf{c}_p$ are obtained by $\mathbf{c}_p = \phi_p^{-2}\boldsymbol{\Sigma}_p\mathbf{w}_p$, implying that $\mathbf{c}_p$ is a vector of least-square loading values of $\mathbf{z}_p$ on γ_p. In turn, $\boldsymbol{\Theta}_p$ is obtained by $\boldsymbol{\Theta}_p = (\mathbf{I}_{Jp} - \mathbf{c}_p\mathbf{w}_p{}')\boldsymbol{\Sigma}_p(\mathbf{I}_{Jp} - \mathbf{w}_p\mathbf{c}_p{}')$. Then, all the block-diagonal elements of $\mathbf{W}$, $\mathbf{C}$, and $\boldsymbol{\Theta}$ can be filled in with $\mathbf{w}_p$, $\mathbf{c}_p$, and $\boldsymbol{\Theta}_p(p = 1, 2, \cdots, P)$, respectively. Also, $\boldsymbol{\tau}$ is computed by $\boldsymbol{\tau} = \mathbf{W}'\boldsymbol{\mu}$ and then, $\mathbf{c}_0$ is calculated as $\mathbf{c}_0 = \boldsymbol{\mu} - \mathbf{C}'\boldsymbol{\tau}$. Next, $\boldsymbol{\Phi}$ is derived by $\boldsymbol{\Phi} = \boldsymbol{\Delta}\boldsymbol{\Phi}_{std}\boldsymbol{\Delta}$. Finally, we obtain $\boldsymbol{\Sigma}$ by $\boldsymbol{\Sigma} = \mathbf{C}'\boldsymbol{\Phi}\mathbf{C} + \boldsymbol{\Theta}$. A more detailed explanation on each step of this procedure can be found in Cho and Choi (2020).

References

Altman, A., & Gondzio, J. (1999). Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization. *Optimization Methods and Software, 11*(1–4), 275–302. https://doi.org/10.1080/10556789908805754

Bollen, K. A., & Bauldry, S. (2011). Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods, 16*(3), 265–284. https://doi.org/10.1037/a0024448

Boyd, S. P., & Vandenberghe, L. (2018). Introduction to applied linear algebra: Vectors, matrices, and least squares. *Cambridge University Press*. https://doi.org/10.1017/9781108583664

Cho, G., & Choi, J. Y. (2020). An empirical comparison of generalized structured component analysis and partial least squares path modeling under variance-based structural equation models. *Behaviormetrika, 47*(1), 243–272. https://doi.org/10.1007/s41237-019-00098-0

Cho, G., Hwang, H., Sarstedt, M., & Ringle, C. M. (2020). Cutoff criteria for overall model fit indexes in generalized structured component analysis. *Journal of Marketing Analytics, 8*, 189–202. https://doi.org/10.1057/s41270-020-00089-1

Cho, G., Jung, K., & Hwang, H. (2019). Out-of-bag prediction error: A cross validation index for generalized structured component analysis. *Multivariate Behavioral Research, 54*(4), 505–513. https://doi.org/10.1080/00273171.2018.1540340

Cho, G., Sarstedt, M., & Hwang, H. (2022). A comparative evaluation of factor- and component-based structural equation modeling methods under (in)consistent model specifications. *British Journal of Mathematical and Statistical Psychology, 75*(2), 220–251. https://doi.org/10.1111/bmsp.12255

Dobson, K. G., Vigod, S. N., Mustard, C., & Smith, P. M. (2021). Major depressive episodes and employment earnings trajectories over the following decade among working-aged Canadian men and women. *Journal of Affective Disorders, 285*, 37–46. https://doi.org/10.1016/j.jad.2021.02.019

Floudas, C. A., & Visweswaran, V. (1995). Quadratic optimization. In R. Horst & P. M. Pardalos (Eds.), *Handbook of global optimization* (pp. 217–269). Springer US. https://doi.org/10.1007/978-1-4615-2025-2_5

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American customer satisfaction index: Nature, purpose, and findings. *Journal of Marketing, 60*(4), 7–18. https://doi.org/10.2307/1251898

Frank, M., & Wolfe, P. (1956). An algorithm for quadratic programming. *Naval Research Logistics Quarterly, 3*(1–2), 95–110. https://doi.org/10.1002/nav.3800030109

Hwang, H., Cho, G., Jin, M. J., Ryoo, J. H., Choi, Y., & Lee, S.-H. (2021). A knowledge-based multivariate statistical method for examining gene-brain-behavioral/cognitive relationships: Imaging genetics generalized structured component analysis. *PLoS ONE, 16*(3), e0247592. https://doi.org/10.1371/journal.pone.0247592

Hwang, H., Cho, G., Jung, K., Falk, C. F., Flake, J., & Jin, M. J. (2021). An approach to structural equation modeling with both factors and components: Integrated generalized structured component analysis. *Psychological Methods, 26*(3), 273–294. https://doi.org/10.1037/met0000336

Hwang, H., Ho, M.-H.R., & Lee, J. (2010). Generalized structured component analysis with latent interactions. *Psychometrika, 75*(2), 228–242. https://doi.org/10.1007/s11336-010-9157-5
Hwang, H., Malhotra, N. K., Kim, Y., Tomiuk, M. A., & Hong, S. (2010). A comparative study on parameter recovery of three approaches to structural equation modeling. *Journal of Marketing Research, 47*(4), 699–712. https://doi.org/10.2139/ssrn.1585305
Hwang, H., & Takane, Y. (2004). Generalized structured component analysis. *Psychometrika, 69*(1), 81–99. https://doi.org/10.1007/BF02295841
Hwang, H., & Takane, Y. (2010). Nonlinear generalized structured component analysis. *Behaviormetrika, 37*(1), 1–14. https://doi.org/10.2333/bhmk.37.1
Hwang, H., & Takane, Y. (2014). *Generalized structured component analysis: A component-based approach to structural equation modeling*. Chapman and Hall/CRC Press.
Hwang, H., Takane, Y., & Malhotra, N. (2007). Multilevel generalized structural component analysis. *Behaviormetrika, 34*(2), 95–109. https://doi.org/10.2333/bhmk.34.95
Johnson, D. R., & Creech, J. C. (1983). Ordinal measures in multiple indicator models: A simulation study of categorization error. *American Sociological Review, 48*, 398–407. https://doi.org/10.2307/2095231
Lawson, C. L., & Hanson, R. J. (1974). *Solving least squares problems*. Abingdon: Prentice Hall.
Lay, D. C., Lay, S. R., & McDonald, J. J. (2015). *Linear algebra and its applications* (5th ed.). Pearson Education.
Naik, D. N., & Khattree, R. (1996). Revisiting olympic track records: Some practical considerations in the principal component analysis. *The American Statistician, 50*(2), 140–144. https://doi.org/10.1080/00031305.1996.10474361
Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education: Theory and Practice,15*(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y
Sullivan, G. M., & Artino, A. R. J. (2013). Analyzing and interpreting data from likert-type scales. *Journal of Graduate Medical Education, 5*(4), 541–542. https://doi.org/10.4300/JGME-5-4-18
Swaminathan, H., & Algina, J. (1978). Scale freeness in factor analysis. *Psychometrika, 43*(4), 581–583. https://doi.org/10.1007/BF02293816
Vanderbei, R. J., & Carpenter, T. J. (1993). Symmetric indefinite systems for interior point methods. *Mathematical Programming, 58*(1), 1–32. https://doi.org/10.1007/BF01581257
Zumbo, B. D., & Zimmerman, D. W. (1993). Is the selection of statistical methods governed by level of measurement? *Canadian Psychology / Psychologie Canadienne, 34*, 390–400. https://doi.org/10.1037/h0078865

Manuscript Received: 19 DEC 2021
Accepted: 25 NOV 2023
Published Online Date: 16 FEB 2024