



Modeling indices using partial least squares: How to determine the optimum weights?

Taşkın DİRSEHAN^{1,2} · Jörg HENSELER^{3,4}

Accepted: 9 August 2022
© The Author(s) 2022

Abstract

Indices are often used to model theoretical concepts in economics and finance. Beyond the econometric models used to test the relationships between these variables, partial least squares path modeling (PLS-PM) allows the study of complex models, but it is an estimator that is still in its infancy in economics and finance research. Thus, the use of PLS-PM for composite analysis needs to be explored further. As one such attempt, this paper is focused on the determination of the indices' optimum weights. For this purpose, the effects of the market potential index (MPI) on foreign direct investment (FDI) and gross domestic product (GDP) were analysed by implementing different weighting schemes. The assessment of the model shows that PLS Mode B leads to better model fit.

Keywords PLS Mode A · PLS Mode B · Partial least squares path modeling · Composite measurement · Indices · Weighting schemes

1 Introduction

Socio-economic position (SEP) is a widely used concept in epidemiological research that is used to determine the well-being of societies (Howe et al. 2008). It refers to physical resources, social resources, and status within a social hierarchy (Krieger 2001). Therefore, scholars often use indices to assess relative SEPs within a population. Moreover, national

✉ Taşkın DİRSEHAN
taskin.dirsehan@marmara.edu.tr; t.dirsehan@utwente.nl

Jörg HENSELER
j.henseler@utwente.nl

¹ Department of Business Administration, Faculty of Business Administration, Marmara University, Istanbul, Turkey

² Visiting Researcher at Department of Design, Production and Management, Faculty of Engineering Technology, University of Twente, Enschede, The Netherlands

³ Department of Design, Production and Management, Faculty of Engineering Technology, University of Twente, Enschede, The Netherlands

⁴ Nova Information Management School, Universidade Nova de Lisboa, Lisbon, Portugal

governments are interested both in the positions and potentials of their economies in the world (Bobek and Vide 2005). Thus, they also benefit from the use of indices. The creation of an index requires the use of formative indicators (Diamantopoulos and Winklhofer 2001). Such composite indicators allow the public to easily comprehend the information provided (Freudenberg 2003). They are widely used in the field of economic and business statistics to clarify their political importance and operational adequacy in decision making (Munda and Nardo 2003).

Although the indicators of SEP tend to be monetary measures, such as income or consumption expenditure, more complex conceptualizations of SEP are needed to cover other aspects of well-being (Howe et al. 2008). Therefore, scholars from economics and finance attempt to develop increasingly complex indices and build more complex models to develop socioeconomic theories.

Moreover, when constructing the indices from a set of variables, scholars need to decide about the weights to assign to each indicator. Principal Components Analysis (PCA) has been one of the methods recommended for determining weights for components (Filmer and Pritchett 2001). However, beyond its complexity, PCA is considered as problematic when used with the discrete data included in the indices, and the first principal component often explains only a low proportion of the total variation in asset data. In recent years, structural equation modeling (SEM) has become an important statistical tool used especially in the social and behavioral sciences (Benitez et al. 2020). Estimators for SEM can be classified as either covariance- or variance-based estimators. While the first estimates model parameters using the empirical variance-covariance matrix, the second estimates the model parameters using proxies based on linear combinations of observed variables (Henseler et al. 2016). On the other hand, SEM through PLS is an approach based on regression, which minimizes the residual variances of variables (Hair et al. 2011; Merli et al. 2019).

Even though some other methods have been suggested, none have proved more simple nor more suitable for discrete data (Howe et al. 2008). Therefore, latent variable approaches have been proposed (Ferguson et al. 2003; Montgomery and Hewett 2005). One recent and useful method for this purpose is partial least squares (PLS), which will be discussed in detail in this paper.

As a widely used variance-based SEM technique, PLS enables the researcher to analyze complex relationships between variables while also examining their direct, indirect and moderating relationships (Nitzl et al. 2016). PLS appears to be an excellent method of studying emergent variables (as opposed to latent variables). Emergent variables are defined by their indicators; in other words, they are composites of variables (Benitez et al. 2020; Henseler and Schubert 2020; Reise 1999). They are also termed ‘composite constructs’ (Benitez et al. 2018), ‘aggregate constructs’ (Edwards 2001), and ‘formative constructs’ (Petter et al. 2007). However in this paper, we follow the suggestion of Benitez et al. (2020) and we emergent variables to emphasize that the construct emerges from the indicators. An emergent variable is a composite with the additional property that it follows the axiom of unity (see Henseler and Schubert 2021). If an emergent variable has effects in a larger model, the emergent variable mediates all the effects of its elements (whereas not all composites do that). One may think of the analogy of common factors and latent variables and ask the question “Why do we need the term *latent variable*?” The answer is that all latent variables are common factors, but not all common factors are latent variables. Latent vari-

ables are those common factors that obey the axiom of local independence (not all common factors do that).

The PLS literature before around 2015 mostly uses the term “PLS-SEM”, which reflects an outdated understanding of what formative and reflective mean. They associate reflective with PLS Mode A and formative with PLS Mode B. Consequently, the model to be analyzed does not exactly match the analyzed model (Henseler 2021). This is probably why proponents of “PLS-SEM” discourage the assessment of model fit. As a result, the term PLS path modeling (PLS-PM) was preferred in this study rather than PLS-SEM.

Owing to its alleged ability to model both factors and composites, PLS-PM has been termed a “silver bullet” (Hair et al. 2011) and it is widely used in various fields of business administration research including information systems (Marcoulides & Saunders, 2006), strategic management (Hair et al. 2012), marketing (Hair et al. 2012), operations management (Peng and Lai 2012), human resource management (Ringle et al. 2020), and finance (Avkiran and Ringle 2018).

Whereas factors can be used to model latent variables in behavioral research, including into attitudes and personal traits, composites can be applied to model strong concepts (Höök and Löwgren 2012) such as social and economic indices. Unlike the measurement models, in the composite model the indicators do not cause the construct, but combine to compose it (Benitez et al. 2020). Composite models do not suffer from factor indeterminacy, and the indicator weights determine – for each observation – how to determine the location parameter (Henseler and Schubert 2021). To specify and assess composite models, Jörg Henseler and Theo K. Dijkstra developed confirmatory composite analysis (CCA) as an innovative set of procedures (Henseler and Schubert 2020).

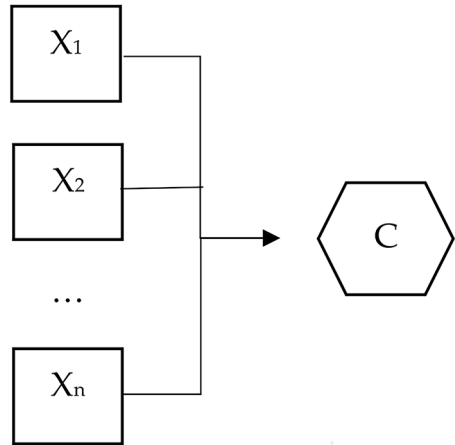
Any choice of weights can be questioned and debated in public discussions (Anand and Sen 1997). By focusing on the composites, this paper illustrates the use of PLS-PM as an analytical tool for determining the weights of indices by comparing alternative weighting schemes that are available in PLS.

2 Foundations of PLS-PM and composite analysis

In recent years, structural equation modeling (SEM) has become an important statistical tool used especially in the social and behavioral sciences (Benitez et al. 2020). Estimators for SEM can be classified as either covariance- or variance-based. While the first estimates model parameters using the empirical variance-covariance matrix, the second estimates the model parameters using the proxies based on linear combinations of observed variables (Henseler et al. 2016). On the other hand, SEM via PLS is an approach based on regression, which minimizes the residual variances of variables (Hair et al. 2011; Merli et al. 2019).

In scientific disciplines, theoretical constructs are the building blocks of theories (Henseler 2017). In general, most theoretical constructs can only indirectly be measured through observable indicators, and no single indicator can capture the full theoretical meaning of an underlying construct (Steenkamp and Baumgartner 2000). Thus, multiple indicators are necessary for measuring a theoretical construct. Therefore, researchers may benefit from SEM to examine the relationships between constructs. SEM is widely preferred by researchers since it permits them to graphically model and estimate parameters for relationships between theoretical constructs and to test behavioral science theories (Bollen 1989).

Fig. 1 Composite model



Based on their estimation objectives, SEM can be categorized as covariance-based and variance-based (Henseler et al. 2009). As a variance-based estimator for SEM, PLS-PM can estimate linear, non-linear, recursive, and nonrecursive structural models (Dijkstra and Henseler 2015; Dijkstra and Schermelleh-Engel 2014).

PLS-PM can be employed for modeling two types of constructs: emergent variables and latent variables (Benitez et al. 2020). Figure 1 shows a composite model in which the hexagon represents an emergent variable, and the rectangles stand for the observable variables. Latent variables are those that are not directly observable but are inferred through a measurement model from directly measured observed variables (Hair et al. 2016). On the other hand, emergent variables are constructs emerging from the indicators; in other words, they are defined by their indicators (Reise 1999). The composite model can be employed to operationalize these concepts (Henseler 2017). The relationship between an emergent variable and its indicators is definitional rather than causal (Henseler 2015). In the literature, the use of indices has gained momentum as an alternative to traditional psychometric measurement models (Diamantopoulos and Winklhofer 2001). Since the indices are composites, PLS is expected to be a useful tool for construct analysis.

3 Application areas of PLS vs. econometric models

In the field of economics and finance, econometric analysis is widely used to analyze the relationships between constructs such as Gross Domestic Product (GDP) (Bajrami et al. 2022), Foreign Direct Investment (FDI) (Song et al. 2021), agricultural employment (Jiang et al. 2022), the unemployment rate (Drachal 2020), total entrepreneurship activity (Gautam and Lal 2021) and those from indices such as the economic inequality index (Brida, Rizzo, Sanchez Carrera, & Segarra, 2021), the global competitiveness index (Sergi et al. 2021), the green finance index (Ye et al. 2022), and the digital index (Litvintseva and Karelin 2020).

Since the performance of econometric models is sensitive to which analytical method is used (Clements and Hendry 1998), it might pay off in some situations to consider the use of PLS-PM. Besides estimating complex models that have many latent and/or emergent variables and handling both reflective (if consistent PLS is employed, see Dijkstra

and Henseler 2015a,b) and composite models, PLS-PM avoids some small sample size problems (Henseler et al. 2009). On the other hand, the visual interface of PLS makes it easier for many users to understand even very complex models and to interpret their results. Moreover, since different econometric models may provide different forecasting accuracies (Song et al. 2003), PLS-PM can also be used as a triangulation for model testing. In addition to cross-sectional studies, PLS-PM has also been proposed to be used for longitudinal studies to analyze unobservable and complex variables over time (Roemer 2016).

4 Various ways of forming composites in PLS: Guidelines for determining optimum weights for indices

PLS-PM includes three approaches to measurement models: PLS Mode A, PLS Mode B, and Sum Scores (Yuan et al. 2020). In PLS Mode A, the weights are calculated as correlations between the observed variables and the corresponding construct scores. In contrast, in PLS Mode B, the weights are calculated as the estimated coefficients of an ordinary least squares regression from the construct scores on the corresponding observed variables (Schuberth et al. 2021). If the constructs are conceptualized as composites, then regression weights (PLS-PM Mode B) should be preferred. On the other hand, if a model includes common factors, the use of correlation weights (PLS Mode A) is recommended in combination with a correction for attenuation as provided by PLSc (Dijkstra and Henseler 2015a,b).

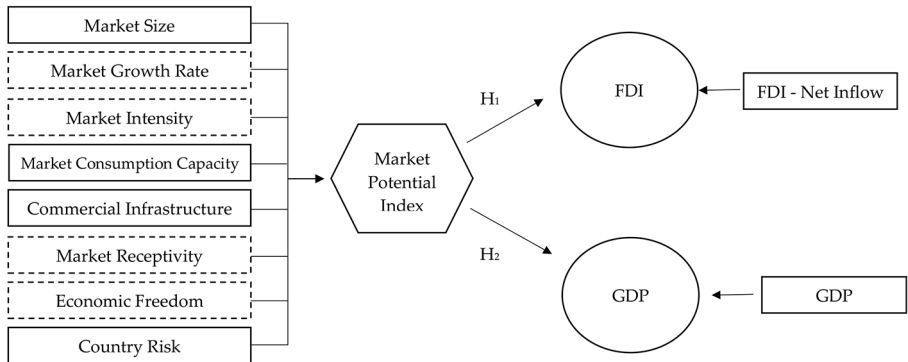
Consequently, scholars tend to choose weighting schemes mostly based on the epistemic relationship between the construct and its indicators, i.e., the choice is mostly based on whether the measurement model is formative or reflective (Diamantopoulos and Winklhofer 2001; Henseler 2010). Often, scholars simply use PLS Mode A for reflective measurements and PLS Mode B for formative measurements of latent variables. However, this oversimplification has casted doubts about PLS-PM's suitability (Aguirre-Urreta and Marakas 2014a, b; Rigdon et al. 2014). In line with the critics of Rönkkö et al. (2016), Schuberth et al. (2021, p. 108) conclude that '*both Mode A and Mode B produce inconsistent parameter estimates for latent variable models such as the reflective measurement and causal formative measurement models.*' Instead, PLS Mode B and PLS Mode A should be understood as different options for determining composite weights. Compared to PLS Mode B, PLS Mode A has the additional constraint that the weight estimates are proportional to the loadings of a common factor model (Henseler and Schuberth 2022).

This study aims to develop guidelines for determining weighting schemes in line with the principles of Occam's razor which are "as parsimonious as possible, as flexible as practical" (Sharma et al. 2019). Many social science scholars give importance to parsimony in theory development since it 'explains much by little' (Benitez et al. 2020; Sharma et al. 2019).

Once a model is estimated, its fit should be assessed to examine whether the model is consistent with the collected data. For this purpose, the bootstrap-based test based on a discrepancy measure – such as the SRMR – can be used (Schuberth et al. 2020). To test the hypothesis of exact model fit, the discrepancy measure should be compared with the 95% or 99% quantiles of its corresponding distribution as obtained through bootstrapping (Henseler et al. 2016). The null hypothesis stating that the indicator population correlation matrix

Table 1 Guidelines to determine weights for indices

Step 1.	Estimate the model based on the extant literature
Step 2.	Assess the overall model fit for different weighting schemes
Step 3.	Update the model with elimination or combination of variables
Step 4.	Keep the weighting scheme that yields insignificant model misfit
Step 5.	Extract construct scores

**Fig. 2** Modeling the Market Potential Index and its impact on foreign direct investment (FDI) and gross domestic product (GDP)

equals the model-implied counterpart is rejected if the discrepancy measure exceeds the 95% or 99% quantiles of its reference distribution (Schuberth et al. 2020).

Based on the explanations above, guidelines are developed to determine the weighting scheme for composites that exhibit an optimum alignment with the data. Those are shown in Table 1. We follow each step of our proposed guidelines to determine the weights of indices as composites.

5 An illustrative example: indicators' weights of the market potential index (MPI)

5.1 Description of the example

We provide an illustrative example to test various weighting schemes in a macroeconomic construct using PLS, where Market Potential Index was assumed to be a predictor of foreign direct investment (FDI) and gross domestic product (GDP). Figure 2 shows the proposed model to be estimated and tested.

As the global economy becomes increasingly interconnected and markets become increasingly accessible, international expansion has become increasingly imperative for companies. Thus, managers need objective criteria to identify, evaluate and select suitable foreign markets (Ozturk et al. 2015). To satisfy this need, the Market Potential Index (MPI) was proposed by Cavusgil (1997) initially for Western firms to assess and compare the

attractiveness of emerging markets. It measures and ranks countries based on eight dimensions (Abbasi and Ansari 2014):

- market size (*the country's population, especially those living in urban areas*)
- market growth rate (*the country's real GDP growth rate*)
- market intensity (*private consumption and gross national income per capita representing the discretionary expenditures of citizens*)
- market consumption capacity (*the percentage share of income earned by the country's middle class*)
- commercial infrastructure (*characteristics such as number of mobile phone subscribers, density of telephone lines, number of PCs, density of paved roads and population per retail outlet*)
- economic freedom (*the degree to which government intervenes in business activities*)
- market receptivity (*the particular emerging market's inclination to trade with the exporter's country as estimated by the volume of imports*)
- country risk (*the degree of political risk*)

The raw data of these variables were rescaled into a scale of 1–100 (Sakarya et al. 2007). Next, the weights of these indicators were determined based on a Delphi process among international business scholars and professionals (Sheng and Mullen 2011). With the aim of guiding U.S. companies' for their expansion in international markets, the index was created for 26 countries classified as "emerging markets" before 2014. By that year, it had been calculated for 87 countries that were in the top 100 performers in terms of total GDP, that had a population of at least one million, and for which reliable data was available for the majority of the indicators used (GlobalEdge, 2021). The index measures have been published every year on the GlobalEdge web site.

Based on a dynamic panel data method, Carstensen and Toubal (2004) showed that market potential has a substantial positive effect on FDI. FDI is a critical component of countries' economies. It has been assumed that FDI leads to increasing income, production, prices, employment, economic growth, development and general welfare of the recipient country (Kok and Ersoy 2009). Moreover, beyond just resources, such investment also brings technology, access to markets, and improvements in human capital (Stiglitz 2000). Since FDI is widely regarded as an aggregation of capital, technology, marketing, and management, many countries consider it as an important part of their economic development strategies (Cheng and Kwan 2000).

Another core indicator in judging the position of a country's economy over time or relative to that of other countries is GDP (Van den Bergh 2009). Since GDP was first suggested an indicator of progress seven decades ago, it has become to be seen as an old-fashion metric for the world today and it has been criticized since it ignores social costs, environmental impacts, and income inequality (Costanza et al. 2014).

Acknowledging the above shortcomings, we consider two basic indicators of economic progress through proposing two hypotheses:

H₁ MPI has a positive effect on FDI.

H₂ FDI has a positive effect on GDP.

Although previous studies that adopted PLS-PM have considered more complex models that included more constructs, second-order constructs and/or moderation effects (Benitez et al. 2020), the hypothesized model seems suitable for our purposes since (1) we aim to provide a guideline for using PLS-PM in confirmatory and explanatory economic research using indices; (2) the considered model with both its latent and emergent variables is a good example of how PLS-PM can leverage its full capacities; and (3) we use endogenous variables with single-indicator constructs.

The composite model is employed to operationalize the indices (Market Potential Index, in the present example). In doing so, Market Potential is assumed to be composed of market size, market growth rate, market intensity, market consumption capacity, commercial infrastructure, economic freedom, market receptivity, and country risk. Scholars and analysts of both economics and finance might also consider other approaches, such as export market potential (Sheng and Mullen 2011), the Bitcoin market potential index (Hileman 2015) or other industry-specific indexes. The hexagon shown in Fig. 2 represents the exogenous construct (here: a composite) and rectangles represents the elements forming the construct (while the dashed rectangles represent the elements eliminated after the first composite analysis).

On the other hand, the endogenous variables in the model, namely FDI and GDP, were modeled as single-indicator constructs.

5.2 Data collection and preparation

To test the hypotheses, we collected secondary data from the following online databases: (1) eight indicators' scores of Market Potential Index were downloaded from the GlobalEdge web site (<https://globaledge.msu.edu/mpi>); (2) FDI data – net inflows from the database of the World Bank (<https://data.worldbank.org/indicator/BX.KLT.DINV.CD.WD?display=default>); and (3) GDP data from the database of the World Bank (<https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?display=default>).

Moreover, we excluded data from 2020 and 2021 to eliminate the lockdown effect of COVID pandemic on economies. Since the MPI calculation method was changed after 2015, we considered data for the years 2016, 2017, 2018 and 2019. We used the averages of the 4-year data for each indicator to smooth any potential bias. Next, we aggregated all the data and we ended up with a list of 82 countries and complete data for all the indicators.

5.3 Estimation of the weights in the composite model

Various software packages – including PLS-Graph, SmartPLS, WarpPLS, XLSTAT-PLS and ADANCO – can be used to estimate the parameters of a PLS path model (Benitez et al. 2020). However, most of them do not allow to test a model's goodness of fit. In this study, we used ADANCO 2.0.1 Professional for Windows (<http://www.composite-modeling.com/>) to estimate the composite model.

Firstly, we needed to set a dominant indicator in the composite model. The dominant indicator is used to dictate the orientation of the construct and it is expected to positively correlate with the construct. Since FDI and GDP were specified as single-indicator constructs, we only needed to determine a dominant indicator for MPI. To select that, face validity can be used (Benitez et al. 2020). Since in the original MPI calculation Market

Table 2 Steps to assess common factor and composite models

Steps	Type of construct	Assessment criteria	Decision criteria
Testing the adequacy of the model	Latent and emergent variable	SRMR d_{ULS} d_G	SRMR < 0.080 SRMR < HI_{95} $d_{ULS} < HI_{95}$ or $d_{ULS} < HI_{99}$ $d_G < HI_{95}$ or $d_G < HI_{99}$
Evaluating reliability of construct scores	Latent variable	ρ_A	$\rho_A > 0.707$
Loadings	Emergent variable	Loading significance	Significant at 5% significance level

Source: Prepared based on Benitez et al. (2020)

Table 3 Results of the Composite Analysis

Weighting Scheme	Discrepancy	Value			Result of assessment	Conclusion
		HI_{95}	HI_{99}			
PLS Mode A	SRMR	0.1293	0.0908	0.1016	SRMR > 0.08; SRMR > HI_{99}	Not Supported
	d_{ULS}	0.3510	0.1730	0.2168	$d_{ULS} > HI_{99}$	Not Supported
	d_G	0.5053	0.2264	0.3530	$d_{ULS} > HI_{99}$	Not Supported
PLS Mode B	SRMR	0.0531	0.0649	0.0843	SRMR < 0.08; SRMR < HI_{95}	Supported
	d_{ULS}	0.0592	0.0884	0.1493	$d_{ULS} < HI_{95}$	Supported
	d_G	0.1566	0.1194	0.1854	$d_{ULS} < HI_{99}$	Acceptable
Sum Scores	SRMR	0.1818	0.1652	0.1944	SRMR > 0.08	Not Supported
	d_{ULS}	0.1818	0.1652	0.1944	$d_{ULS} < HI_{99}$	Acceptable
	d_G	0.7922	0.2636	0.3914	$d_{ULS} > HI_{99}$	Not Supported

Size had received the highest weight (25/100), in our example, we chose it as the dominant indicator for MPI. Next, we used the factor weighting scheme for inner weighting, and statistical inferences were based on the bootstrap procedure, relying on 4,999 bootstrap runs as recommended by Henseler et al. (2016). It is important that each bootstrap sample has the same size as the original sample. ADANCO takes this into account automatically.

When we ran the analysis, the model produced insignificant factor loadings for market growth rate, market intensity, market receptivity and economic freedom. After elimination of these elements, we ran it again.

As the next step, we tried PLS Mode A, PLS Mode B and Sum Score as the weighting schemes for the constructs.¹ Since both FDI and GDP are single-indicator constructs, their loadings equal one and remain thus unaffected by a change in weighting scheme. Therefore, we focus on the weighting schemes for MPI.

5.4 Assessment of the composite model

The steps followed to assess the composite model are summarized in Table 2. The assessment started with the evaluation of the overall fit of the model. Table 3 summarizes the values of the discrepancy measures as well as the 95% and 99% quantiles of their corre-

¹ Since we do not aim to classify nor rank the countries based on their overall MPI score, but instead we aim to use the index as a construct with its indicators and since the PLS can handle non-standardized values, we skip the normalization procedure of the regular overall MPI score calculation.

Table 4 Weight and loading estimates of the Market Potential Index

Weighting Scheme	Indicator	Weight	Loading
PLS Mode A	Market Size	0.5314***	0.7245***
	Market Consumption Capacity	0.3217***	0.8047***
	Commercial Infrastructure	0.3282***	0.7676***
	Country Risk	0.1717**	0.6071***
PLS Mode B	Market Size	0.8240***	0.9283***
	Market Consumption Capacity	0.0654	0.5608***
	Commercial Infrastructure	0.3706*	0.5666***
	Country Risk	-0.0389	0.2965*
Sum Scores	Market Size	0.3333***	0.5379***
	Market Consumption Capacity	0.3333***	0.8392***
	Commercial Infrastructure	0.3333***	0.8522***
	Country Risk	0.3333***	0.7711***

Note: * $p < .05$, ** $p < .01$, *** $p < .001$

sponding reference distribution for our example. Only for PLS Mode B was the value of the SRMR below the recommended threshold value of 0.080 (Henseler et al. 2014). Next, when we considered whether the discrepancy measures were below the 95% quantile of their reference distribution (HI_{95}) or at least below the 99% quantile (HI_{99}), we observed that only in PLS Mode B do the assessment criteria perform well.

Even though composite models are typically estimated by Mode B (regression weights) in PLS-PM (Benitez et al. 2020), we evaluated all the weighting schemes on the same basis. Table 4 shows the weights and loadings obtained through PLS Mode A, PLS Mode B and Sum Scores as weighting schemes. It is important to note the difference between weights and loadings. While weights represent the degree of importance of each indicator (ingredient) to the construct, composite loadings represent the correlation between the indicator and the corresponding emergent variable (Cenfetelli and Bassellier 2009). After evaluating the significance of the weight and loading estimates, scholars can decide whether or not to keep non-significant indicators for a construct's content validity (Hair et al. 2016). Based on this explanation, we observe that the remaining four indicators (Market Size, Market Consumption Capacity, Commercial Infrastructure, and Country Risk) loadings are significant in all the weighting schemes.

6 Discussion and further research

Estimating and assessing models for indices as composites in the fields of economics and finance poses a challenge for scholars employing SEM, particularly in determining a suitable weighting scheme. To overcome this difficulty, this study suggests employing PLS-PM and provides guidelines to determine the optimum weighting scheme to be used in the analysis. For this purpose, we propose a five-step approach to choosing between PLS Mode A, PLS Mode B, and Sum Scores. This enables researchers to determine the optimum weights for the indices in their analyses.

We used widely known constructs in economics and applied each step we proposed in the guidelines. We compared the outcomes of three weighting schemes: PLS Mode A, PLS Mode B, and Sum Scores. Even though an isolated composite is not identified, as soon as an

emergent variable is embedded in a wider model (i.e., it is studied in its nomological net), there is only one possible set of weights that fulfill the requirements of an emergent variable.

Based on the results from the illustrative example that included composites, similar to the findings of Schuberth et al. (2020) and (Becker et al. 2012), PLS Mode B produced not only consistent estimates, but also acceptable discrepancy measures. Therefore, it was preferable over the other weighting schemes. On the other hand, in the case of a large degree of multicollinearity among the indicators of a composite, the use of PLS Mode A might be worthwhile (Schuberth et al. 2020).

Since any preliminary study is limited by its design, we recommend future research to conduct more complex index studies. For instance, researchers could attempt to develop higher-order models and/or test the relationships between various indices.

7 Conclusions

Indices are often used in economics and finance research to test various types of theoretical concepts, such as market potential, as well as to estimate and test their relationships. PLS-PM is a useful estimator for this purpose. However, determining optimum weights for indices remains an open issue in PLS-PM analyses. This study provides guidelines on PLS-PM for composite analysis by testing three weighting schemes (PLS Mode A, PLS Mode B, and Sum Score). To achieve this, an example was provided with an estimation of the effect of MPI on FDI and GDP. Two hypotheses were developed based on the extant literature. As a result, the key contribution of this study to the methodological literature on the empirical research in economics and finance was the attempt to determine the optimum weights of indices. Further research is recommended to test how would PLS-PM perform as an alternative to econometric analyses and how to transfer modeling tactics to panel data.

The empirical study revealed that whereas PLS Mode A and sum scores yielded a significant model misfit, the model's goodness of fit was acceptable when PLS Mode B was employed. Therefore, based on the results, we recommend using PLS Mode B to determine weights in PLS-PM studies that involve the use of indices in economics and finance.

Funding Jörg Henseler gratefully acknowledges financial support from FCT Fundação para a Ciência e a Tecnologia (Portugal), national funding through a research grant from the Information Management Research Center – MagIC/NOVA IMS (UIDB/04152/2020).

Declarations

Conflict of interest Jörg Henseler acknowledges a financial interest in ADANCO and its distributor, Composite Modeling.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abbasi, E.H., Ansari, N.: Market potential indicators: A comparative analysis of Brazil and India. *Journal of Economics & Business Research* 20(1), 65–76 (2014).
- Aguirre-Urreta, M.I., Marakas, G.M.: A rejoinder to Rigdon (2014). *Inform. Syst. Res.* 25(4), 785–788 (2014a). <https://doi.org/10.1287/isre.2014.0545>
- Aguirre-Urreta, M.I., Marakas, G.M.: Research note—Partial least squares and models with formatively specified endogenous constructs: A cautionary note. *Inform. Syst. Res.* 25(4), 761–778 (2014b). <https://doi.org/10.1287/isre.2013.0493>
- Anand, S., Sen, A.: Concepts of human development and poverty! A multidimensional perspective. United Nations Development Programme, Poverty and human development: Human development papers, 1–20. (1997)
- Avkiran, N., Ringle, C.M.: Partial least squares structural equation modeling. In *Handbook of Market Research* (Vol. 267, pp. 1–29). (2018)
- Bajrami, R., Gashi, A., Ukshini, K., Rexha, D.: Impact of the government size on economic growth in the western Balkan countries. *Journal of Governance and Regulation* 11(1), 55–63 (2022). <https://doi.org/10.22495/jgrv11i1art6>
- Becker, J.-M., Klein, K., Wetzels, M.: Hierarchical latent variable models in PLS-SEM: guidelines for using reflective-formative type models. *Long Range Plann.* 45(5–6), 359–394 (2012). <https://doi.org/10.1016/j.lrp.2012.10.001>
- Benitez, J., Henseler, J., Castillo, A., Schubert, F.: How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research. *Inf. Manag.* 57(2), 103168 (2020). <https://doi.org/10.1016/j.im.2019.05.003>
- Benitez, J., Llorens, J., Braojos, J.: How information technology influences opportunity exploration and exploitation firm’s capabilities. *Inf. Manag.* 55(4), 508–523 (2018). <https://doi.org/10.1016/j.im.2018.03.001>
- Bobek, V., Vide, R.K.: The signification and the feasibility of measuring globalization of economy. *Industrial Manage. Data Syst.* 105(5), 596–612 (2005). <https://doi.org/10.1108/02635570510599977>
- Bollen, K.A.: *Structural Equations with Latent Variables*. Wiley, New York (1989)
- Brida, J.G., Risso, W.A., Carrera, S., Segarra, V.: Growth and inequality in the Mexican states: Regimes, thresholds, and traps. *Papers in Regional Science.* 100(5), 1295–1322 (2021). <https://doi.org/10.1111/pirs.12616>
- Carstensen, K., Toubal, F.: Foreign direct investment in Central and Eastern European countries: a dynamic panel analysis. *J. Comp. Econ.* 32(1), 3–22 (2004). <https://doi.org/10.1016/j.jce.2003.11.001>
- Cavusgil, S.T.: Measuring the potential of emerging markets: An indexing approach. *Bus. Horiz.* 40(1), 87–92 (1997). <https://link.gale.com/apps/doc/A19369690/AONE?u=anon~5db338e8&sid=googleScholar&xid=11686bb9>
- Cenfetelli, R.T., Bassellier, G.: Interpretation of formative measurement in information systems research. *MIS Q.* 33(4), 689–707 (2009). <https://doi.org/10.2307/20650323>
- Cheng, L.K., Kwan, Y.K.: What are the determinants of the location of foreign direct investment? The Chinese experience. *J. Int. Econ.* 51(2), 379–400 (2000). [https://doi.org/10.1016/S0022-1996\(99\)00032-X](https://doi.org/10.1016/S0022-1996(99)00032-X)
- Clements, M., Hendry, D.: *Forecasting Economic Time Series*. Cambridge University Press (1998)
- Costanza, R., Kubiszewski, I., Giovannini, E., Lovins, H., McGlade, J., Pickett, K.E., Wilkinson, R.: Development: Time to leave GDP behind. *Nat. News.* 505(7483), 283–285 (2014). <https://doi.org/10.1038/505283a>
- Diamantopoulos, A., Winklhofer, H.M.: Index construction with formative indicators: An alternative to scale development. *J. Mark. Res.* 38(2), 269–277 (2001). <https://doi.org/10.1509/jmkr.38.2.269.18845>
- Dijkstra, T.K., Schermelleh-Engel, K.: Consistent partial least squares for nonlinear structural equation models. *Psychometrika* 79(4), 585–604 (2014). <https://doi.org/10.1007/S11336-013-9370-0>
- Dijkstra, T.K., Henseler, J.: Consistent and asymptotically normal PLS estimators for linear structural equations. *Comput Stat Data Anal* 81, 10–23 (2015a). <https://doi.org/10.1016/j.csda.2014.07.008>
- Dijkstra, T.K., Henseler, J.: Consistent partial least squares path modeling. *MIS Q.* 39(2), 297–316 (2015b). <https://doi.org/10.25300/MISQ/2015/39.2.02>
- Drachal, K.: Forecasting unemployment rate in Poland with dynamic model averaging and internet searches. *Glob. Bus. Econ. Rev.* 23(4), 368–389 (2020). <https://doi.org/10.1504/GBER.2020.110684>
- Edwards, J.R.: Multidimensional constructs in organizational behavior research: An integrative analytical framework. *Organ. Res. Methods* 4(2), 144–192 (2001). <https://doi.org/10.1177/109442810142004>
- Ferguson, B.D., Tandon, A., Gakidou, E., Murray, C.J.: Estimating permanent income using indicator variables. Health systems performance assessment: Debates, methods and empiricism. Geneva: World Health Organization, 747–760 (2003)

- Filmer, D., Pritchett, L.H.: Estimating wealth effects without expenditure data—or tears: An application to educational enrollments in states of India. *Demography* 38(1), 115–132 (2001). <https://doi.org/10.1353/dem.2001.0003>
- Freudenberg, M.: Composite indicators of country performance: A critical assessment. OECD Science, Technology and Industry Working Papers (2003)
- Gautam, S., Lal, M.: Entrepreneurship and economic growth: Evidence from G-20 economies. *J. East-West Bus.* 27(2), 140–159 (2021). <https://doi.org/10.1080/10669868.2020.1859427>
- GlobalEdge: Market Potential Index (MPI) – 2020 (2021). <https://globaleedge.msu.edu/mpii>
- Hair, J.F., Ringle, C.M., Sarstedt, M.: PLS-SEM: Indeed a silver bullet. *J. Mark. Theory Pract.* 19(2), 139–152 (2011). <https://doi.org/10.2753/MTP1069-6679190202>
- Hair, J.F., Sarstedt, M., Pieper, T.M., Ringle, C.M.: The use of partial least squares structural equation modeling in strategic management research: A review of past practices and recommendations for future applications. *Long Range Plann.* 45(5–6), 320–340 (2012). <https://doi.org/10.1016/j.lrp.2012.09.008>
- Hair, J.F., Sarstedt, M., Ringle, C.M., Mena, J.A.: An assessment of the use of partial least squares structural equation modeling in marketing research. *J. Acad. Mark. Sci.* 40(3), 414–433 (2012). <https://doi.org/10.1007/s11747-011-0261-6>
- Hair, J.F., Hult, G.T.M., Ringle, C.M., Sarstedt, M.: A primer on partial least squares structural equation modeling (PLS-SEM). Sage (2016)
- Henseler, J.: On the convergence of the partial least squares path modeling algorithm. *Comput. Stat.* 25(1), 107–120 (2010). <https://doi.org/10.1007/s00180-009-0164-x>
- Henseler, J. Is the whole more than the sum of its parts? On the interplay of marketing and design research. Universiteit Twente (2015). https://research.utwente.nl/files/5119155/oratieboekje_Henseler.pdf
- Henseler, J.: Bridging design and behavioral research with variance-based structural equation modeling. *J. Advertising.* 46(1), 178–192 (2017). <https://doi.org/10.1080/00913367.2017.1281780>
- Henseler, J., Schuberth, F.: Auxiliary theories. *Composite-Based Structural Equation Modeling: Analyzing Latent and Emergent Variables*, pp. 25–37. Guilford Press (2021)
- Henseler, J., Schuberth, F.: Partial least squares as a tool for scientific inquiry: Comments on Cadogan and Lee. *European Journal of Marketing*, in print (2022). <https://doi.org/10.1108/EJM-06-2021-0416>
- Henseler, J., Dijkstra, T.K., Sarstedt, M., Ringle, C.M., Diamantopoulos, A., Straub, D.W.,... Calantone, R.J. Common beliefs and reality about PLS: Comments on Rönkkö and Evermann (2013). *Organ Res Methods*, 17(2), 182–209 (2014). <https://doi.org/10.1177/1094428114526928>
- Henseler, J., Hubona, G., Ray, P.A.: Using PLS path modeling in new technology research: Updated guidelines. *Industrial Manage. Data Syst.* 116(1), 2–20 (2016). <https://doi.org/10.1108/IMDS-09-2015-0382>
- Henseler, J., Ringle, C.M., Sinkovics, R.R.: The use of partial least squares path modeling in international marketing. *Advances in International Marketing* 20, 277–319 (2009). [https://doi.org/10.1108/S1474-7979\(2009\)0000020014](https://doi.org/10.1108/S1474-7979(2009)0000020014)
- Henseler, J., Schuberth, F.: Using confirmatory composite analysis to assess emergent variables in business research. *J. Bus. Res.* 120, 147–156 (2020). <https://doi.org/10.1016/j.jbusres.2020.07.026>
- Henseler, J.: Composite-Based Structural Equation Modeling: Analyzing Latent and Emergent Variables. Guilford Press (2021)
- Hileman, G. The bitcoin market potential index. Paper presented at the International Conference on Financial Cryptography and Data Security (2015)
- Howe, L.D., Hargreaves, J.R., Huttly, S.R.: Issues in the construction of wealth indices for the measurement of socio-economic position in low-income countries. *Emerg. Themes Epidemiol.* 5(1), 1–14 (2008). <https://doi.org/10.1186/1742-7622-5-3>
- Höök, K., Löwgren, J.: Strong concepts: Intermediate-level knowledge in interaction design research. *ACM Trans. Computer-Human Interact. (TOCHI)* 19(3), 1–18 (2012). <https://doi.org/10.1145/2362364.2362371>
- Jiang, G., Alvarado, R., Murshed, M., Tillaguango, B., Toledo, E., Méndez, P., Isik, C.: Effect of agricultural employment and export diversification index on environmental pollution: Building the agenda towards sustainability. *Sustainability* 14(2), 677 (2022). <https://doi.org/10.3390/su14020677>
- Kok, R., Ersoy, B.A.: Analyses of FDI determinants in developing countries. *Int. J. Soc. Econ.* 36(1/2) 105–123 (2009). <https://doi.org/10.1108/03068290910921226>
- Krieger, N.: A glossary for social epidemiology. *J. Epidemiol. Community Health.* 55(10), 693–700 (2001). <https://doi.org/10.1136/jech.55.10.693>
- Litvintseva, G., Karelin, I.: Effects of digital transformation of the economy and quality of life in Russia. *Terra Economicus* 18(3), 53–71 (2020). <https://doi.org/10.18522/2073-6606-2020-18-3-53-71>
- Marcoulides, G.A., Saunders, C. (eds.): Editor's comments: PLS: a silver bullet? *MIS Q.* 30(2), iii–ix (2006). <https://doi.org/10.2307/25148727>
- Merli, R., Preziosi, M., Acampora, A., Ali, F.: Why should hotels go green? Insights from guests experience in green hotels. *Int. J. Hospitality Manage.* 81, 169–179 (2019). <https://doi.org/10.1016/j.ijhm.2019.04.022>

- Montgomery, M.R., Hewett, P.C.: Urban poverty and health in developing countries: Household and neighborhood effects. *Demography* 42(3), 397–425 (2005). <https://doi.org/10.1353/dem.2005.0020>
- Munda, G., Nardo, M.: On the methodological foundations of composite indicators used for ranking countries. *Ispra, Italy: Joint Research Centre of the European Communities*, 1–19 (2003)
- Nitzl, C., Roldan, J.L., Cepeda, G.: Mediation analysis in partial least squares path modeling. *Industrial Manage. Data Syst.* 116(9), 1849–1864 (2016). <https://doi.org/10.1108/IMDS-07-2015-0302>
- Ozturk, A., Joiner, E., Cavusgil, S.T.: Delineating foreign market potential: A tool for international market selection. *Thunderbird Int. Bus. Rev.* 57(2), 119–141 (2015). <https://doi.org/10.1002/tie.21686>
- Peng, D.X., Lai, F.: Using partial least squares in operations management research: A practical guideline and summary of past research. *J. Oper. Manag.* 30(6), 467–480 (2012). <https://doi.org/10.1016/j.jom.2012.06.002>
- Petter, S., Straub, D., Rai, A.: Specifying formative constructs in information systems research. *MIS Q.* 31(4), 623–656 (2007). <https://doi.org/10.2307/25148814>
- Reise, S.P.: Personality measurement issues viewed through the eyes of IRT. In: *The New Rules of Measurement: What Every Psychologist and Educator Should Know*, pp. 219–242. Lawrence Erlbaum Associates (1999)
- Rigdon, E.E., Becker, J.-M., Rai, A., Ringle, C.M., Diamantopoulos, A., Karahanna, E., Dijkstra, T.K.: Conflating antecedents and formative indicators: A comment on Aguirre-Urreta and Marakas. *Inform. Syst. Res.* 25(4), 780–784 (2014). <https://doi.org/10.1287/isre.2014.0543>
- Ringle, C.M., Sarstedt, M., Mitchell, R., Gudergan, S.P.: Partial least squares structural equation modeling in HRM research. *Int. J. Hum. Resource Manage.* 31(12), 1617–1643 (2020). <https://doi.org/10.1080/09585192.2017.1416655>
- Roemer, E.: A tutorial on the use of PLS path modeling in longitudinal studies. *Industrial Manage. Data Syst.* 116(9), 1901–1921 (2016). <https://doi.org/10.1108/IMDS-07-2015-0317>
- Rönkkö, M., McIntosh, C.N., Antonakis, J., Edwards, J.R.: Partial least squares path modeling: Time for some serious second thoughts. *J. Oper. Manag.* 47, 9–27 (2016). <https://doi.org/10.1016/j.jom.2016.05.002>
- Sakarya, S., Eckman, M., Hyllegard, K.H.: Market selection for international expansion: Assessing opportunities in emerging markets. *Int. Mark. Rev.* 24(2) 208–238 (2007). <https://doi.org/10.1108/02651330710741820>
- Schuberth, F., Müller, T., Henseler, J.: Which equations? An inquiry into the equations in partial least squares structural equation modeling. In: *New Perspectives in Serving Customers, Patients, and Organizations: A Festschrift for Judit Simon*, pp. 96–115. Corvinus University of Budapest (2021)
- Schuberth, F., Rademaker, M.E., Henseler, J.: Estimating and assessing second-order constructs using PLS-PM: The case of composites of composites. *Industrial Manage. Data Syst.* 120(12), 2211–2241 (2020). <https://doi.org/10.1108/IMDS-12-2019-0642>
- Sergi, B.S., D'Aleo, V., Konecka, S., Szopik-Depczyńska, K., Dembińska, I., Ioppolo, G.: Competitiveness and the Logistics Performance Index: The ANOVA method application for Africa, Asia, and the EU regions. *Sustain. Cities Soc.* 69, 102845 (2021). <https://doi.org/10.1016/j.scs.2021.102845>
- Sharma, P., Sarstedt, M., Shmueli, G., Kim, K.H., Thiele, K.O.: PLS-based model selection: The role of alternative explanations in information systems research. *J. Association Inform. Syst.* 20(4) (2019). <https://doi.org/10.17005/1.jais.00538>
- Sheng, S.Y., Mullen, M.R.: A hybrid model for export market opportunity analysis. *Int. Mark. Rev.* 28(2), 163–182 (2011). <https://doi.org/10.1108/02651331111122650>
- Song, H., Witt, S.F., Jensen, T.C.: Tourism forecasting: Accuracy of alternative econometric models. *Int. J. Forecast.* 19(1), 123–141 (2003). [https://doi.org/10.1016/S0169-2070\(01\)00134-0](https://doi.org/10.1016/S0169-2070(01)00134-0)
- Song, Y., Zhang, Y., Wang, Y., Zhang, B., Su, J.: The influence of foreign direct investment on the urban–rural income gap: Evidence from China. *Kybernetes* 51(1), 466–484 (2021). <https://doi.org/10.1108/K-07-2020-0488>
- Steenkamp, J.-B.E.M., Baumgartner, H.: On the use of structural equation models for marketing modeling. *Int. J. Res. Mark.* 17(2–3), 195–202 (2000). [https://doi.org/10.1016/S0167-8116\(00\)00016-1](https://doi.org/10.1016/S0167-8116(00)00016-1)
- Stiglitz, J.E.: Capital market liberalization, economic growth, and instability. *World Dev.* 28(6), 1075–1086 (2000). [https://doi.org/10.1016/S0305-750X\(00\)00006-1](https://doi.org/10.1016/S0305-750X(00)00006-1)
- Van den Bergh, J.C.: The GDP paradox. *J. Econ. Psychol.* 30(2), 117–135 (2009). <https://doi.org/10.1016/j.joep.2008.12.001>
- Ye, T., Xiang, X., Ge, X., Yang, K.: Research on green finance and green development based eco-efficiency and spatial econometric analysis. *Sustainability* 14(5), 2825 (2022). <https://doi.org/10.3390/su14052825>
- Yuan, K.-H., Wen, Y., Tang, J.: Regression analysis with latent variables by partial least squares and four other composite scores: Consistency, bias and correction. *Struct. Equation Modeling: Multidisciplinary J.* 27(3), 333–350 (2020). <https://doi.org/10.1080/10705511.2019.1647107>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.