
This is the **accepted version** of the journal article:

Lamberti, Giuseppe. «Hybrid multigroup partial least squares structural equation modelling : an application to bank employee satisfaction and loyalty». *Quality and Quantity*, Vol. 57 (2021), p. 683-705. DOI 10.1007/s11135-021-01096-9

This version is available at <https://ddd.uab.cat/record/322113>

under the terms of the  license

Hybrid multigroup partial least squares structural equation modelling. An application to bank employee satisfaction and loyalty

Lamberti Giuseppe

Abstract

We describe a practical approach to tackling observed heterogeneity using partial least squares structural equation modelling (PLS-SEM) when the number of categorical variables is high and the context of the research is exploratory. The approach is based on combining classical multigroup PLS-SEM approach and pathmax analysis. We provide practical guidance on using our hybrid multigroup PLS-SEM and illustrate its application using real data for bank employees. In investigating work climate, specifically the relationship between satisfaction and loyalty considering specific drivers (empowerment, company reputation, leadership, pay, and work conditions) and different sources of heterogeneity (gender, age, marital status, education, job level, and antiquity), the hybrid multigroup PLS-SEM identified three partitions defined by juniors, seniors, and managers, and identified significant differences between those groups, specifically in indicating that leadership and pay were more important for juniors, empowerment for seniors, and company reputation and work conditions for managers.

Keywords: Heterogeneity • Multigroup analysis • PLS-MGA • Pathmax analysis • Job satisfaction • Employee loyalty.

Giuseppe Lamberti Giuseppe.Lamberti@uab.cat

Department of Business, School of Economics and Business, Universitat Autònoma de Barcelona, Avinguda de L'Eix Central, Room B1-1108, Edifici B2,s/n, Avinguda de l'Eix Central, 08193 Bellaterra, Cerdanyola del Vallès, Spain

1. Introduction

In recent decades, partial least squares structural equation modelling (PLS-SEM) has become well established as a multivariate method for analysing complex causal relationships between variables (Akter et al. 2017; Hair et al. 2019; Hair et al. 2014b; Sarstedt et al. 2014).

PLS-SEM allows multiple blocks of observed variables (or indicators) to be analysed, with each block playing the role of a latent variable when it is hypothesized that a system of linear relationships exists between blocks (Esposito Vinzi et al. 2010). Two models are computed: the measurement (or outer) model that relates observed variables to latent variables, and the structural (or inner) model that computes the strength and direction of relationships among the latent variables (Ciavolino et al. 2018). This approach is different from common factor-based (i.e., covariance-based) structural equation modelling (SEM), which, considering latent variables to be common factors that explain co-variation between the associated indicators (Sarstedt et al. 2016), is based on strict assumptions regarding distributions and sample size. PLS-SEM, in contrast, maximizes the variation in the observed variables following a soft approach that requires no strong assumptions regarding distributions, sample size (Hair et al 2018).

PLS-SEM usefulness for prediction and the fact that no assumptions are required (Chin et al. 2020; Hair et al. 2012; Hwang et al. 2020; Reinartz et al. 2009) have contributed to its popularity and to its application in fields as varied as marketing (Cheah et al. 2020; Cheung et al. 2020; Lim et al. 2019; Osakwe et al. 2020), management (Chin et al. 2020; Hair et al. 2014; Menon et al. 2020), sport (Chih-Hung Wang and Tang 2018; Howat and Assaker 2013), technology (Kim et al. 2020, Hensler et al. 2016; Shiau et al. 2019; Lamberti et al. 2020) and, most recently, psychology (Signore et al. 2019; Macchitella et al. 2020).

The problem remains, however, as to how to apply PLS-SEM in a heterogeneity context. While the emphasis may be on collecting data on the variables corresponding directly to the questions to be answered by a study, it is also usual to collect background information, e.g., on sociodemographic variables (gender, age, social status, education, income, etc) that may be useful in segmenting groups and identifying different patterns in models, i.e., heterogeneity. Using a single model when heterogeneity is present can be

misleading as the causal relationships between constructs are likely to differ depending on differing group profiles defined by specific segmentation variables (Becker et al. 2013). Resolving the heterogeneity problem may require different analyses for heterogeneous and homogeneous groups, as defined by segmentation variables.

Researchers have invested considerable efforts in developing methods capable of taking observed and unobserved (or latent) heterogeneity into account in stepped model estimates (Becker et al. 2013; Chin 2000, 2003a; Esposito Vinzi et al. 2008; Fordellone and Vichi 2020; Henseler et al. 2009; Keil et al. 2000; Lamberti et al. 2016; Ringle et al. 2010; Ringle and Schlittgen 2007; Schlittgen et al. 2016; Tenenhaus et al. 2006). Focusing (for our purposes) on observed heterogeneity, a widely used approach is multigroup analysis (MGA), with applications reported for a multitude of contexts, ranging from marketing to the social sciences (Hair 2018; Matthews 2017; Sarstedt et al. 2011). The principal reasons for its popularity are that it is simple to use and easy to interpret, and has suitable statistical modelling software available, including R (R Development Core Team 2012), XLSTAT (Esposito Vinzi et al. 2007) and SmartPLS (Sarstedt and Cheah 2019; Ringle et al. 2005). However, MGA is less than efficient in three scenarios: (1) when there is a large number of segmentation variables, (2) when there is a large number of segmentation levels, and – related to these issues – (3) when the research aim is to consider the mixed effects of segmentation variables. Those three factors are especially relevant when we are analysing survey-based data with numerous sociodemographic details for respondents.

Pathmax analysis (Lamberti et al. 2016, 2017) was developed as a possible solution to this problem. Using an iterative procedure, all possible binary partitions of the segmentation variables are analysed by an algorithm and significant segments are generated as output. However, the method has the drawback that it cannot verify, a posteriori, whether the path coefficients of the defined segments are statistically significant.

As a solution to overcome the limitations of the MGA and pathmax approaches to observed heterogeneity when the number of segmentation variables is high, we propose a method that we call hybrid multigroup PLS-SEM in our article, organized as follows. We first describe the main approaches to dealing with observed heterogeneity and provide practical guidance on applying MGA and pathmax analysis. We then describe our hybrid multigroup approach and illustrate it with an application to satisfaction and loyalty among

the employees of a large Spanish bank. Finally, we conclude with a discussion of our findings and our conclusions.

2. Heterogeneity in PLS-SEM

Classically, depending on a priori knowledge we can distinguish between two kinds of heterogeneity: (1) observed heterogeneity, where the number of variables is known a priori (e.g., if we know that heterogeneity is defined by age and education we can define segments accordingly); and (2) unobserved (or latent) heterogeneity, where the number of segments and their composition is not known a priori but becomes apparent from differences in structural path coefficients. Here we focus on observed heterogeneity for the MGA and pathmox analysis, underlining the corresponding advantages and limitations.

2.1. Multigroup analysis

MGA, which was developed for PLS-SEM, tests statistical differences in path coefficients between segments. Heterogeneity is usually observed, i.e., the number of groups in which observations can be segmented is known a priori; indeed, most of the MGA applications in the literature are based on a limited nominal number of segmentation variables observed as sources of heterogeneity. The data is separated into segments according to segmentation variables and calculated for each segment are path coefficients, which are then compared to identify significant differences. This comparative analysis is based on testing two hypotheses: the null hypothesis, which is that path coefficients between segment are the same, i.e., $\beta_{G1}^{path} = \beta_{G2}^{path}$, and the alternative hypothesis, which is that path coefficients between segments are different, i.e., $\beta_{G1}^{path} \neq \beta_{G2}^{path}$. The statistics used to test the null hypothesis can be calculated using several procedures.

Tests reported in the literature and exhaustively reviewed by Hair, Sarstedt, Ringle, and Gudergan (2018), include the parametric t-test (Keil et al. 2000; Chin 2000) along with its Smith-Satterthwaite adaptation, the PLS-MGA (Henseler et al. 2009), and the permutation test (Chin 2003a). An additional method is the omnibus test proposed by Sarstedt et al. (2011), which allows the comparison to be extended to more than two groups. The parametric t-test and Satterthwaite adaptation both use a bootstrap re-

sampling procedure; the coefficients for each segment are calculated in each re-sampling and the standard error estimates are parametrically tested using t-tests. The PLS-MGA applies a non-parametric approach, and, like the parametric approach, uses a bootstrap procedure. For each parameter, bootstrap estimates for each segment are compared with bootstrap estimates for other segments, and counting the number of times the bootstrap estimate for a first segment is larger than that of a second segment obtains a p-value for a one-tailed test. The permutation test, which applies approximate randomization (i.e., permutation) to evaluate path coefficient differences across segments, operates on the basic premise that all segments are equivalent, and that every segment member is the same as before sampling began. A calculated statistic is determined to be significant on the basis of the likelihood of its significance if assignments to segments have been jumbled/mixed. Finally, the omnibus test combines bootstrapping and permutation testing to mimic an overall F-test.

The main limitation of these approaches is in relation to dimensionality: making comparisons and understanding the reasons for significant differences become increasingly difficult as the number of segments increases. This happens whenever there is a large number of segmentation variables, when variables have a large number of levels and/or when effects are mixed and models are complex.

Regarding software and the tests they implement, SmartPLS implements the parametric t-test (Chin 2000; Keil et al. 2000), the Smith-Satterthwaite t-test adaptation and the PLS-MGA (Henseler et al. 2009), whereas the *plsmp* package in R (Sanchez et al. 2013) and XLSTAT implement the parametric t-test (Chin 2000; Keil et al. 2000) and the permutation test (Chin 2003a).

2.2. Pathmox analysis

Pathmox analysis (Lamberti et al. 2016, 2017) has recently been proposed to analyse heterogeneity in the PLS-SEM environment, as it is capable of detecting heterogeneity when many segmentation variables are present and when it is difficult to determine which variables are responsible for differences. The principles of binary segmentation are applied in pathmox analysis to produce a tree with different models in each of the obtained nodes. Pathmox functions as follows:

Step 1. A global PLS-SEM model is fit to all the data to define the root of the tree.

Step 2. The pathmox algorithm establishes the set of admissible partitions for each node in the tree considering the available segmentation variables.

Step 3. In an iterative procedure, models with the greatest differences in each child node are identified.

Three parameters need to be pre-defined for the algorithm: (1) the significance threshold, which, according to classical statistical testing, can be fixed as $p=0.05$; (2) the minimum number of elements per node, which depends on sample size, which, in turn, should be large enough to guarantee the goodness of the results; and (3) the depth of the tree's growing level, given that the maximum tree level has to be moderate to avoid overfitting, and this also ensures that the algorithm does not have to search for large trees with large numbers of leaves and that the final number of interpretable segments is manageable.

Optimal partitions and binary splits are determined by a statistical test based on Fisher's F-test for equality in regression models (Chow 1960; Lebart et al. 1979), adapted for two structural models calibrated from different segments (successors of a node). This test, in model comparisons at the overall structural level, compares parent node path coefficients with child node path coefficients as defined by the binary splits. The parent node model reflects homogeneity, whereas the child node models reflect heterogeneity.

The main limitation of pathmox analysis is related to its goal, which is to identify the most significant models differing in terms of available sources of heterogeneity. However, no a posteriori test can verify whether significant differences, if any, exist between the path coefficients for the identified models.

Pathmox analysis can be implemented using the `genpathmox` package in R (Lamberti 2014) using the `pls.pathmox` function. A technical sheet is available on the website describing the necessary steps and R code for definition of the hybrid segmentation variable.

3. Hybrid multigroup PLS-SEM

To overcome the limitations of MGA and pathmox analysis we propose our hybrid multigroup approach, which involves three steps that can be summarized as follow:

- Step 1.** Determine the hybrid segmentation variable as follows:
- a. Perform pathmox analysis to identify the most significant subgroups
 - b. Define the hybrid segmentation variable according to the resulting final tree nodes
 - c. Interpret the levels of the hybrid segmentation variable according to the segmentation variables identified as responsible for the tree partitions.
- Step 2** Estimate and analyse the PLS-SEM models associated with the hybrid segmentation variable levels.
- Step 3** Run MGA comparisons of path coefficients for each group using one of the following tests: parametric test, permutation test, PLS-MGA test or omnibus test.

The hybrid segmentation variable is first built as a categorical variable where levels identify more different PLS-SEM models associated with more different subgroups of individuals. Next, using pathmox analysis, subgroups are identified by partitioning data according to the available segmentation variables. The resulting subgroups are then compared to check if they are significantly different and, if so, the most significant subgroups are identified. Each identified subgroup is therefore associated with one level of the hybrid segmentation variable. Ultimately, the final tree nodes identified by pathmox analysis define the hybrid segmentation variable, while each node identifies a different level of the hybrid segmentation variable. The next stage is to determine the meaning of the levels of the hybrid segmentation variable, which requires a suitable label to be assigned to each variable level. This is done by analysing the segmentation variables that the pathmox analysis identified as responsible for the tree partitions and assigning a suitable label to the level associated with each terminal node. Once the hybrid segmentation variable has been interpreted, the PLS-SEM models associated with each variable level can be estimated and MGA can be implemented in order to identify the statistically different coefficients, using an MGA test selected in accordance with the literature (see Hair et al. 2018).

Before implementing the MGA, it is recommended to test measurement invariance, using, for instance, the measurement invariance of composite models (MICOM) procedure (Henseler et al. 2016).

Measurement invariance is a statistical property of measurement that demonstrates whether or not we are measuring the same constructs in various specified groups. As argued by Hair et al. (2018) and Henseler et al. (2016), the establishment of measurement invariance across groups is a logical prerequisite to conducting group comparisons and necessary to ensure that model estimate differences do not result from distinctive content or meanings for the unobserved variables across groups (Hair et al. 2018).

4. Application to bank employee data

Using work climate data for the financial sector, we illustrate how hybrid multigroup PLS-SEM deals with heterogeneity when there is a large number of segmentation variables. The five segmentation variables used in this case were age, gender, marital status, education, job level (position in the company hierarchy), and antiquity (length of service). The data and model have recently been published (see Lamberti et al. 2020).

In the financial sector, the impact of the work climate on the relationship between strategic human resource (HR) management and organizational performance is especially crucial (Abdolshah et al. 2018; Bartel 2004; Clark 2002; Kaya et al. 2014; Richard and Johnson 2001; Sekaran 1989; Tlaiss 2013; among others). Another issue of relevance is that different groups of employees may respond in different ways to specific HR management practices (Eskildsen et al. 2004; Tlaiss 2013; Valaei and Jiroudi 2016).

Our objectives were (1) to identify key characteristics defining different groups of employees, and (2) to analyse differences in the models of those groups, i.e., differences in the causal relationships between the latent variables. We started by considering a holistic model that related employee satisfaction and loyalty with the five work climate constructs of empowerment, company reputation, pay, work conditions, and leadership (Eskildsen et al. 2010; Eskildsen et al. 2004a, 2004; Gaertner 1999; Kristensen and Westlund 2004; Lane et al. 2010). We then tested all the relationships against the six sociodemographic variables of age, gender, marital status, education, job level, and antiquity.

In this case the number of variables is high, so use of classical MGA would be complex and produce inaccurate results, whereas use of pathmoX analysis could result in significant groups but would not identify the significant differences between them.

4.1. Data

Our data,¹ obtained from an organizational study of an international bank, consisted of 37 variables for 2,000 employees divided into six qualitative variables and 31 indicator (or manifest) variables. The six qualitative variables, reflecting specific employee characteristics that served as potential observed sources of heterogeneity (segmentation variables) were age, gender, marital status, education, job level, and antiquity. The 31 indicator (or manifest) variables were satisfaction (five indicators), loyalty (three indicators), and the five work climate constructs of empowerment (six indicators), company reputation (five indicators), leadership (five indicators), pay (four indicators), and work conditions (three indicators) (see description and summary statistics in Appendix A). Each indicator was measured on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

The employees in the sample were mostly men (69.70%) and were mostly younger than 45 years old (80.55%). Around 50% were married, most had graduate studies (59.65%), and a third (30.20%) had post-graduate studies. Low-level employees followed by high-level employees represented around two third each of the sample (39.45% and 35.10%, respectively), and most employees (72.45%) had been working in the bank for more than five years. Summarized descriptive information for the employees is provided in Table 1.

¹ The data were obtained from an organizational study of a Spanish bank collaborating with a Spanish university statistics department. The data had been collected through a survey (run by the bank's human resource department), consisting of a questionnaire emailed to employees, whose anonymity was guaranteed. The response rate was 94%. The collected data were processed and analyzed by the university statistics department.

Table 1 Descriptive details of 2,000 bank employees

Gender	%	Age	%	Job level	%
Female	30.30%	<31 years	33.45%	Low	39.45%
Male	69.70%	31-45 years	47.10%	Intermediate	25.45%
		>45 years	19.45%	High	35.10%
Antiquity	%	Education	%	Marital status	%
<2000	10.30%	Secondary	9.15%	Not married	49.50%
2000-2004	23.50%	Graduate	59.65%	Married	50.50%
2005-2009	19.75%	Post-graduate	31.20%		
2009-2014	18.90%				
>2014	27.55%				

4.2. Model

Our model, depicted in Figure 1, is similar to/an adaptation of models described elsewhere (Eskildsen et al. 2010, 2004a, 2004; Gaertner 1999; Lane et al. 2010). The model reflects employee satisfaction and loyalty (the endogenous constructs) and their antecedents, namely, empowerment, company reputation, pay, work conditions, and leadership (the exogenous constructs), i.e., latent variables assumed to affect the work climate. Table 2 summarizes the definitions of each construct, their hypothesized effect on other constructs, and the corresponding literature references, while a full explanation of the causal relationships can be found in Lamberti et al. (2020).

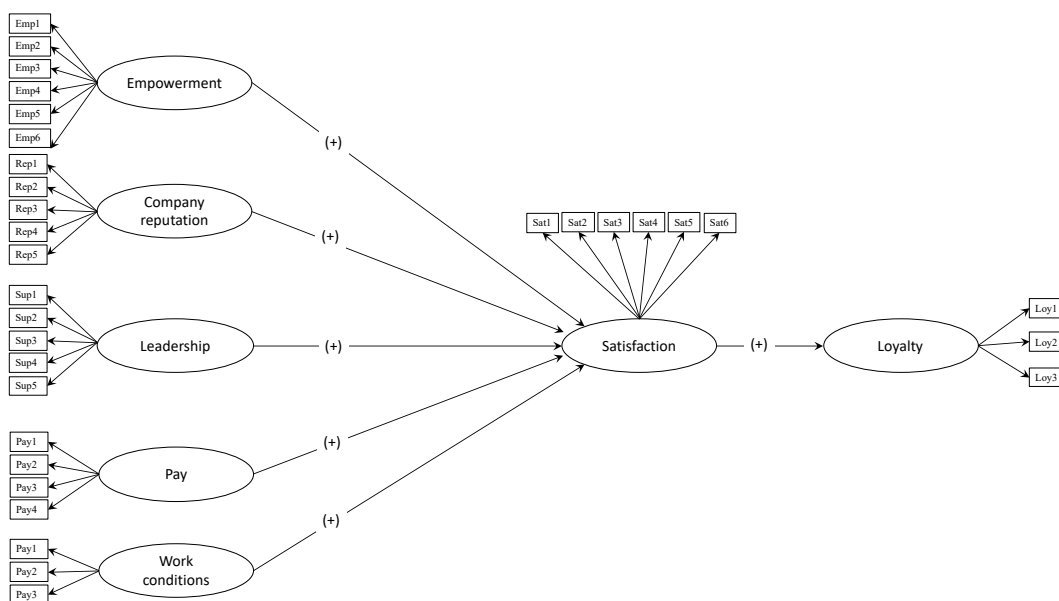


Fig. 1 Work climate model

Table 2 Constructs, hypothesized effects and literature references

Construct	Definition	Hypothesis	References
Empowerment	Employee perceptions of management vision (vertical or horizontal)	Positive effect on satisfaction	Mahmood et al. 2019; Ullah and Yasmin 2013
Reputation	Employee perceptions of the company's reputation	Positive effect on satisfaction	Herrbach et al. 2004; Helm 2011
Leadership	Employee perceptions of leadership practices such as feedback and appraisal	Positive effect on satisfaction	Ugboro and Obeng 2000; Puni et al. 2018
Pay	Employee perceptions of pay and of its equity	Positive effect on satisfaction	Guan et al. 2014; Mahmood et al. 2019
Conditions	Employee perceptions regarding the conditions in which they perform their work	Positive effect on satisfaction	Glaveli and Karassavidou 2011; Kaya et al. 2014
Satisfaction	Employee contentedness with their job and aspects such as kind of work, supervision, etc	Positive effect on loyalty	Eskildsen et al. 2004, 2004a; Dhir et al. 2020
Loyalty	Employee's level of attachment or commitment to their company		Guillon and Cezanne 2014

4.3. Softwares

The SmartPLS 3.2 software (Ringle et al. 2015) and the *genpathmox* package in R software version 0.4 (Lamberti 2014) provided the model estimates (global and multigroup analysis) and the pathmox analysis, respectively.

4.4. Results

4.4.1. Global PLS-SEM

All the latent variables in the model were modelled assuming that each was an antecedent of its own indicators (i.e., a reflective approach). It means that it was assumed that manifest variables were considered to be caused by latent variables. This approach differs from the formative approach, in which the latent construct is assumed to be formed by its indicators (Cheah et al. 2019).

The model was validated, following indications by Ringle et al. (2018) for PLS-SEM model use in the HR management field, and common reliability indexes were computed for each latent variable. The results are presented in Table 3. In line with Henseler et al. (2016), the reliability of a group was considered to be satisfactory if composite reliability (CR), Cronbach's α and ρ_A were greater than 0.7. CR values were between 0.808 and 0.932 for all the constructs. Cronbach's α values were higher than 0.70, except for work conditions (0.692) and loyalty (0.659), while ρ_A (Dijkstra and Henseler 2015) values were all higher than 0.70, except (again) for work conditions (0.695) and loyalty (0.698). We retained those constructs, however, as the values were close to the 0.7 threshold and loadings were higher than 0.7. Internal consistency for all the groups was supported by average variance extracted (AVE). All item loadings were close to or higher than the 0.7 threshold except few cases. Again, we maintained these as reliability and internal consistency was good enough and it seemed important to avoid reducing richness in the construct indicators. All loadings were significant according to the t-test. To assess discriminant validity, we calculated the heterotrait-monotrait (HTMT) ratio of correlations (Henseler et al. 2015), which can be interpreted as the correlation between two constructs; its maximum value, which should be as low as possible, was 0.839 in our case, i.e., lower than the conservative threshold of 0.850.

Common method bias (CMB) was tested following the approach described by Valaei and Jiroudi (2016) for a job satisfaction model applied to a HR management context and validated using Harman's one-factor test. We entered all principal constructs into a single principal component factor analysis, setting to one the number of factors. Results show that only one factor emerged, explaining 40% of the variance, which, at under 50%, indicated the absence of CMB. We also ran the full collinearity test (Kock 2015), consisting of calculating all variance inflation factors (VIFs) for the latent variables and confirming that they were lower than 3.3. Table 4 report the full collinearity test results.

Table 3 Outer model results for 31 indicators corresponding to five constructs

Constructs and indicators	Loadings	t*	Cronbach's α	ρ_A	CR	AVE
<i>Empowerment</i>						
Empo1	0.793	114.175				
Empo2	0.799	128.854				
Empo3	0.752	92.468				
Empo4	0.612	41.841				
Empo5	0.657	50.403				
Empo6	0.636	44.846				
<i>Reputation</i>						
			0.803	0.924	0.841	0.515
Rep1	0.751	49.700				
Rep2	0.665	38.920				
Rep3	0.739	47.317				
Rep4	0.690	37.068				
Rep5	0.739	47.800				
<i>Leadership</i>						
			0.909	0.914	0.932	0.733
Lead1	0.849	156.140				
Lead2	0.826	113.878				
Lead3	0.897	238.507				
Lead4	0.850	156.956				
Lead5	0.858	158.617				
<i>Pay</i>						
			0.766	0.770	0.851	0.590
Pay1	0.818	103.657				
Pay2	0.659	44.722				
Pay3	0.749	60.790				
Pay4	0.835	106.675				
<i>Conditions</i>						
			0.692	0.695	0.830	0.620
Work1	0.798	91.887				
Work2	0.828	112.363				
Work3	0.734	59.499				
<i>Satisfaction</i>						
			0.808	0.808	0.867	0.566
Sat1	0.725	76.799				
Sat2	0.781	98.479				
Sat3	0.736	77.744				
Sat4	0.781	101.329				
Sat5	0.737	72.471				
<i>Loyalty</i>						
			0.659	0.698	0.808	0.584
Loy1	0.825	129.954				
Loy2	0.723	58.706				
Loy3	0.742	61.590				

* t-test significance values.

Table 4 Full collinearity test results for common method bias

Construct	Empowerment	Image	Pay	Work Condition	Leadership	Satisfaction	Loyalty
Empowerment		1.445	1.218	1.363	1.244	2.220	2.062
Image	2.667		1.205	1.393	1.460	2.698	1.901
Pay	2.666	1.430		1.368	1.460	2.665	2.134
Work Condition	2.606	1.442	1.194		1.460	2.671	2.125
Leadership	2.272	1.444	1.217	1.395		2.716	2.135
Satisfaction	2.172	1.430	1.190	1.367	1.455		1.920
Loyalty	2.574	1.285	1.216	1.388	1.459	2.451	

VIF<3.3 indicates no common method bias.

The structural model results are reported in Table 5 and Figure 2. The main drivers for satisfaction were empowerment ($\beta=0.540$), followed by company reputation ($\beta=0.168$), work conditions ($\beta=0.130$), and pay ($\beta=0.118$), with the value for leadership ($\beta=0.044$) indicating it to be the least important driver. All coefficients were significant according to the t-test. Concerning the predictability of the model, in line with Ringle et al. (2018), we calculated R^2 and the Stone-Geisser test of predictive relevance (Q^2), obtaining R^2 values of 0.591 and 0.437 and Q^2 values of 0.537 and 0.341 for satisfaction and loyalty, respectively. We also ran PLSpredict², as indicated in Shmueli et al. (2016, 2019) (see Table 6) to compare PLS-SEM prediction for key dependent constructs and indicators with the prediction obtained using a naïve linear model (LM) approach that did not consider a causal relationship. According to Shmueli (2019), if the distribution of prediction errors can be assumed to be normal, then comparing the square root of the average of the squared differences between the predictions and the actual root mean square error (RMSE) observations between PLS-SEM and LM, we should find the RMSE for PLS-SEM to be lower than the RMSE for LM. In the best-case scenario, if the RMSE for PLS-SEM is lower in 100% of the indicators, then we can conclude that the model

² We fixed to one (i.e., $r=1$) the number of repetitions. According to Shmueli et al. (2019), this is adequate when the prediction is based on a single model. We maintained to 10 the number subsample (i.e., $k=10$) as this value respects the minimum required training sample size.

has high predictive power, while if no lower than 50% of the indicators, then we can assume that that the model has moderate predictive power. As confirmed by the results in Table 6, we our model has moderate predictive power.

Finally, concerning total effects on loyalty, the most important drivers were empowerment ($\beta=0.351$), followed by company reputation ($\beta=0.111$), work conditions ($\beta=0.086$), pay ($\beta=0.078$), and leadership ($\beta=0.029$).

Table 5 Structural model results for specific path coefficients

Path coefficients	Path	t*	p-value	Total effects
Empowerment→Satisfaction	0.54	38.943	<0.001	0.357
Reputation→Satisfaction	0.168	15.326	<0.001	0.111
Leadership→Satisfaction	0.044	3.351	<0.001	0.029
Pay→Satisfaction	0.118	9.051	<0.001	0.078
Conditions→Satisfaction	0.13	69.653	<0.001	0.086
Satisfaction→Loyalty	0.661	9.592	<0.001	
	R²	Q²		
Satisfaction	0.591	0.537		
Loyalty	0.437	0.341		

* t-test significance values.

Table 6 PLSpredict results for satisfaction and loyalty

Constructs	PLS-RMSE	Q ² predict	LM-RMSE	Difference
Loy1	0.728	0.357	0.729	-0.001
Loy4	0.726	0.153	0.696	0.030
Loy5	0.738	0.201	0.739	-0.001
Sat1	0.559	0.360	0.560	-0.001
Sat2	0.658	0.347	0.663	-0.005
Sat3	0.609	0.372	0.610	-0.001
Sat4	0.689	0.326	0.678	0.011
Sat5	0.600	0.248	0.590	0.010

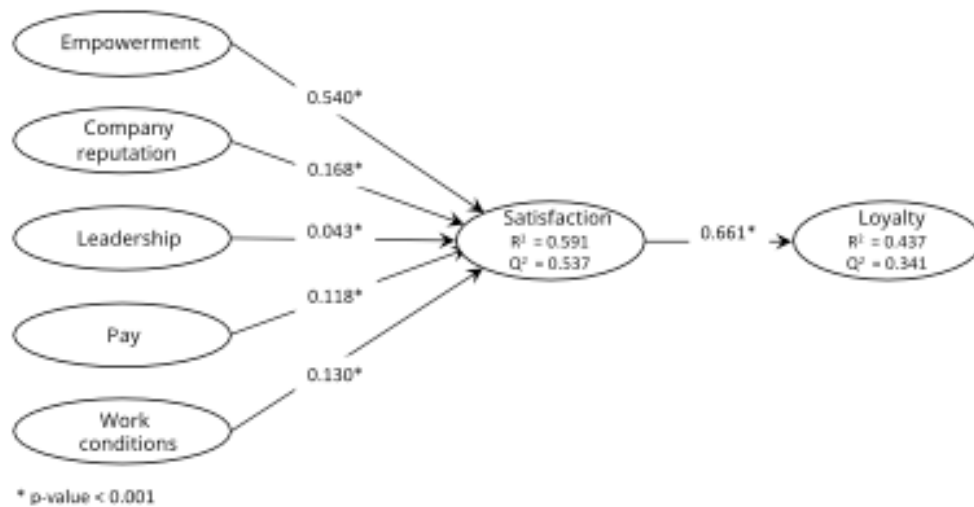


Fig. 2 Structural model results showing R^2 and Q^2 values

4.4.2. Hybrid multigroup PLS-SEM results

Step 1: Determining the hybrid segmentation variable

Pathmax analysis was implemented using the variables listed in Table 1. In terms of investigating possible differences in how job satisfaction, and thus loyalty, were perceived by the bank employees, the hybrid segmentation variable allowed us to identify the most important segments for comparison purposes. Given our objective of identifying only the main groups of employees, we limited the segmentation tree to a maximum depth of two levels and bounded the final number of segments to a maximum of four. The minimum admissible size for a node was set to 10% of the total sample, while the threshold significance for the partitioning algorithm was $p=0.05$. The resulting pathmax tree, depicted in Figure 3, had a total of three terminal nodes.

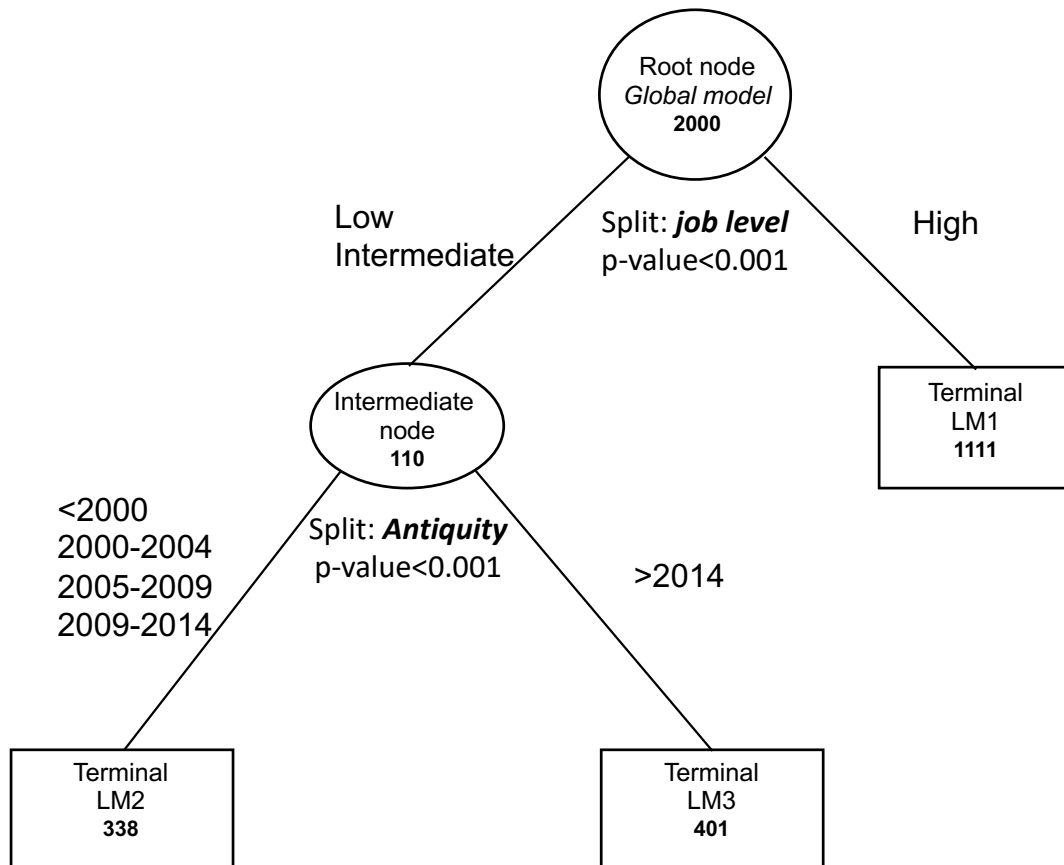


Fig. 3 Pathmox tree

The algorithm selected job level as the variable with the greatest power ($p < 0.001$) to distinguish between higher- and lower-level employees. The lower-level employees were further differentiated according to antiquity ($p < 0.001$). Identified, thus, were three segments as follows: higher-level employees (managers), lower-level employees with greater antiquity (seniors), and lower-level employees with lesser antiquity (juniors). These three segments corresponded to the three hybrid segmentation variable levels used subsequently in the MGA to compare path coefficients.

Step 2: Estimating and analysing the PLS-SEM models associated with the hybrid segmentation variable levels

Table 7 shows the path coefficients and the t-test statistics calculated for the groups according to the hybrid segmentation variable levels. The most important driver for satisfaction overall was empowerment, followed by leadership and pay for junior employees, and company reputation and work conditions for managers. As for seniors, relative to the other two groups, the most important driver for satisfaction was

empowerment. Finally, satisfaction was important for loyalty in all three groups; nonetheless, although differences could be observed between the three groups, we needed to verify whether those differences were statistically significant, which was done using MGA.

Table 7 PLS-SEM model with hybrid segmentation variable levels

Path coefficients	Juniors			Seniors			Managers		
	Path	t*	p	Path	t*	p	Path	t*	p
Empowerment →Satisfaction	0.513	3.307	<0.001	0.65	4.666	<0.001	0.477	15.663	<0.001
Reputation →Satisfaction	0.088	4.484	<0.001	0.113	3.973	<0.001	0.21	10.385	<0.001
Leadership →Satisfaction	0.155	4.767	<0.001	0.025	0.871 ^{NS}	0.324	0.037	2.165	0.019
Pay →Satisfaction	0.143	4.79	<0.001	0.101	4.223	<0.001	0.072	4.159	<0.001
Conditions →Satisfaction	0.121	33.476	<0.001	0.103	33.628	<0.001	0.187	46.391	<0.001
Satisfaction →Loyalty	0.671	15.841	<0.001	0.658	22.568	<0.001	0.633	25.734	<0.001

* t-test significance values. NS indicates non-significant according to the t-test.

Step 3: Multigroup comparisons of path coefficients

Measurement invariance test. Before making the group comparison, we checked measurement invariance following the three-step measurement invariance of composite models (MICOM) procedure (Henseler et al. 2016), which ensures that model estimate differences do not result from distinctive content or meanings for the latent variables across groups (Hair et al. 2018). The MICOM procedure involves three steps: (1) configural invariance, (2) compositional invariance, and (3) equality of composite mean values and variance values. Following Hair et al. (2018), to apply the MGA, we needed to establish both configural invariance and compositional invariance so as to establish partial measurement invariance.

Configural invariance, which ensures that each latent variable has been specified the same for all the groups, exists when constructs are equally parameterized and estimated

across groups. To establish configural invariance, each latent variable in the PLS-SEM model must therefore be specified equally for all the groups, ensuring, in our case, identical indicators for each measurement model, identical data treatment, and identical algorithm settings across groups. Compositional invariance means that composite scores measure exactly the same construct across groups.

The MICOM procedure tests for compositional invariance by calculating the latent score correlations between groups and comparing them with the reference distribution of correlations obtained by permutation of the two groups. If the observed correlation falls in the upper 95% of the distribution, then the null hypothesis that the theoretical correlation is 1 is accepted, and composite invariance of the specific construct is established. If, additionally, the constructs have equal mean values and variances across the groups, full measurement invariance is confirmed, allowing the data of different groups to be pooled.

Tables 8 and 9 show MICOM step 2 and step 3 results for the observed correlations with the lowest 5% quantile of the reference distribution, the observed differences in means between the two compared groups, and the log ratio of variances for both groups with their corresponding 95% confidence intervals (obtained by group permutation). Compositional invariance was verified in most cases; in the very few exceptions, since the observed deviation was minimal we could globally assume the compositional invariance of the constructs. Regarding the equality of means of the composite scores, the fact that we could appreciate differences in most cases confirmed the existence of different segments according to our sources of heterogeneity. However, in terms of the variability of the composite scores for the segments, in many cases the variance between groups was similar. Since our results confirmed partial measurement invariance, we could proceed to MGA.

Table 8 MICON step 2 results

Construct	Group	Compositional invariance			
		Score correlation	5% quantile	Verified	
Empowerment			0.999	YES	
Reputation			0.988	YES	
Leadership	Juniors vs managers		1.000	YES	
Pay			0.996	YES	
Conditions			0.999	0.997	YES
Satisfaction			1.000	0.999	YES
Loyalty			0.995	0.998	NO
Empowerment			0.999	YES	
Reputation			0.993	0.983	YES
Leadership			1.000	1.000	YES
Pay	Juniors vs seniors		0.990	0.994	NO
Conditions			1.000	0.997	YES
Satisfaction			1.000	1.000	YES
Loyalty			0.998	0.996	YES
Empowerment			1.000	0.999	YES
Reputation			0.991	0.992	NO
Leadership			1.000	1.000	YES
Pay	Seniors vs managers		0.997	0.995	YES
Conditions			0.999	0.998	YES
Satisfaction			1.000	1.000	YES
Loyalty			0.993	0.998	NO

Table 9 MICON step 3 results

Construct	Equality of composite values					
	Mean values			Variance values		
	Mean diff.	2.5%-97.5%	Verified	Log. var ratio	2.5%-97.5%	Verified
<i>Junior vs manager</i>						
Empowerment	0.198	[-0.080 - 0.078]	NO	0.076	[-0.132 - 0.132]	YES
Reputation	0.045	[-0.080 - 0.084]	YES	-0.019	[-0.126 - 0.122]	YES
Leadership	0.242	[-0.078 - 0.079]	NO	-0.054	[-0.110 - 0.109]	YES
Pay	-0.458	[-0.079 - 0.077]	NO	0.153	[-0.124 - 0.113]	NO
Conditions	0.381	[-0.079 - 0.080]	NO	-0.138	[-0.116 - 0.112]	NO
Satisfaction	-0.139	[-0.078 - 0.079]	NO	0.548	[-0.157 - 0.154]	NO
Loyalty	0.055	[-0.080 - 0.085]	YES	0.054	[-0.144 - 0.136]	YES
<i>Junior vs senior</i>						
Empowerment	0.563	[-0.100 - 0.100]	NO	-0.153	[-0.146 - 0.148]	NO
Reputation	0.587	[-0.100 - 0.100]	NO	-0.055	[-0.133 - 0.134]	YES
Leadership	0.320	[-0.098 - 0.100]	NO	-0.190	[-0.136 - 0.139]	NO
Pay	-0.225	[-0.099 - 0.101]	NO	-0.088	[-0.133 - 0.133]	YES
Conditions	0.603	[-0.098 - 0.100]	NO	-0.216	[-0.144 - 0.142]	NO
Satisfaction	0.494	[-0.098 - 0.098]	NO	-0.037	[-0.148 - 0.152]	YES
Loyalty	0.715	[-0.101 - 0.099]	NO	-0.068	[-0.142 - 0.141]	YES
<i>Senior vs manager</i>						
Empowerment	-0.411	[-0.078 - 0.080]	NO	0.237	[-0.125 - 0.125]	NO
Reputation	-0.543	[-0.082 - 0.081]	NO	0.045	[-0.118 - 0.112]	YES
Leadership	-0.086	[-0.079 - 0.080]	NO	0.130	[-0.115 - 0.112]	NO
Pay	-0.210	[-0.082 - 0.080]	NO	0.213	[-0.123 - 0.116]	NO
Conditions	-0.243	[-0.080 - 0.084]	NO	0.080	[-0.114- 0.106]	YES
Satisfaction	-0.717	[-0.080 - 0.082]	NO	0.573	[-0.149 - 0.141]	NO
Loyalty	-0.681	[-0.079 - 0.081]	NO	0.122	[-0.125 - 0.124]	YES

PLS-MGA test. For the model associated with the hybrid segmentation variable levels, we compared path coefficients for each group using the PLS-MGA test. Pairwise group comparisons allowed us to investigate specific differences between groups of employees, obtaining the results shown in Table 9. Since the PLS-MGA is a one-tailed test according to Hair et al. (2018), the p-values indicated whether a path coefficient was significantly larger for one group than for another group, while the 1-p values indicated whether there was a significant difference in the opposite direction. We found that empowerment was significantly more important for seniors than for juniors or managers ($p=0.999$ and $p<0.001$, respectively), that company reputation was significantly more important for managers than for juniors or seniors ($p=1.00$ in both cases), that leadership was significantly more important for juniors than for seniors or managers ($p=0.001$ and $p<0.001$, respectively), that pay was significantly more important for juniors than for managers ($p=0.020$) and, finally, that work conditions were more important for managers than for juniors or seniors ($p=0.978$ and $p=0.996$, respectively).

Table 9 PLS-MGA results according to the hybrid segmentation variable

Path coefficient	Path			p		
	JNR	SNR	MNG	JNR vs SNR	JNR vs MNG	SNR vs MNG
Empowerment→Satisfaction	0.513	0.650	0.477	0.999*	0.173	0.000*
Reputation→Satisfaction	0.088	0.113	0.210	0.763	1.000*	1.000*
Leadership→Satisfaction	0.155	0.025	0.037	0.001*	0.000*	0.639
Pay→Satisfaction	0.143	0.101	0.072	0.140	0.020*	0.161
Conditions→Satisfaction	0.121	0.103	0.187	0.311	0.978	0.996*
Satisfaction→Loyalty	0.671	0.658	0.633	0.326	0.063	0.146

*Significant differences. JNR, juniors; SNR, seniors; MNG, managers.

5. Discussion

Our article describes a practical approach to the problem of dealing with heterogeneity in PLS-SEM when the number of segmentation variables is high, making it difficult to determine the most significant differences. In our approach we identify the

most heterogenous segments according to the segmentation variables available and then verify significant differences between the segmented groups.

While MGA is a well-established method for analysing heterogeneity in PLS-SEM models, its usefulness becomes limited when the number of comparisons starts to increase, making it difficult to suitably interpret results. This problem arises when we are dealing with either/both a large number of segmentation variables and segmentation variable levels. Pathmox analysis is a more suitable approach for a large number of variables, as the algorithm is capable of detecting the most significantly different groups; however, it has the drawback that there is no a posteriori significance test for the path coefficients of the estimated models. Our hybrid multigroup PLS-SEM merges the pathmox algorithm with classical MGA in such a way as to overcome the main limitations of both techniques. This hybrid approach also makes it easier, using the pathmox tree results, to define the hybrid segmentation variable and consider interactions between categorical variables.

Since both MGA and pathmox analysis are well established techniques, the method does not represent any substantial theoretical innovation, yet its contribution is eminently practical in terms of offering a way to test differences in path coefficients when dealing with a large number of categorical variables.

We demonstrated application of our hybrid approach using a practical example of job satisfaction and loyalty in a banking context, where it is important to design effective HR management practices according to different employee profiles. Applying our hybrid method to the six segmentation variables of gender, age, marital status, education, job level, and antiquity, we were able to define three specific employee groups, namely, junior employees, senior employees, and managers. For these groups, furthermore, we determined which of empowerment, company reputation, pay, work conditions, and leadership were especially important drivers of satisfaction and loyalty, finding that leadership and pay were significantly more important for juniors, empowerment for seniors, and company reputation and work conditions for managers. Furthermore, job level and antiquity were found to be the most important factors in segmenting employees. This kind of knowledge is valuable in that it can be used to shape more specific, and therefore, more effective, HR management policies.

This approach to segmentation has applications that go beyond HR management, as it could potentially be applied in many areas of business and marketing where

identifying different groups and understanding why they are different is crucial to delimiting target groups and defining good strategies. It also has recently been applied in the psychology field (Signore et al. 2019; Macchitella et al. 2020).

As a practical example in relation to psychology, researchers could apply the hybrid multigroup PLS-SEM to the employability model proposed by Signore et al. (2019) and use it to test whether significant differences exist according to the hybrid segmentation variables based on the categorical variables of age, educational level, disciplinary area, and working conditions.

Regarding when the hybrid multigroup PLS-SEM approach can be used (see Table 10), researchers should consider aspects such as the application context, the source of heterogeneity, the complexity of the research, and the type of hypothesis to be tested. The hybrid multigroup PLS-SEM approach is suitable when the primary goal of the research is exploratory and heterogeneity is observed, i.e., when, in analysing different potential sources of heterogeneity, hypotheses cannot be defined a priori regarding differences between groups and when the hybrid segmentation variables are defined after the pathmax analysis. This approach is also indicated when the complexity of the analysis is high, as, due to either/both a large number of segmentation variables and segmentation variable levels, between-group comparisons and interpretation of differences are complex. The approach is not indicated for confirmatory research, unobserved heterogeneity and testing a specific hypothesis.

Another limitation concerns the way in which the hybrid segmentation variables are defined. The method uses pathmax partitioning criteria based on the classical parametric F-statistic, and so assumes the normality of perturbation terms with equal variance in all the endogenous constructs, yet this is an assumption rarely met in practice. Nonetheless, although violations of this assumption may lead to an invalid F-statistic, Lamberti et al. (2016, 2017) demonstrated the robustness of the test.

From a computational point of view the hybrid approach is more efficient than the classical multigroup approaches, because the multigroup test and MICOM procedure are applied only to the most significant partitions determined according to the pathmax analysis, which means that fewer comparisons are necessary. The importance of this aspect is directly related to the number of observations and the number of comparisons.

While an objective of this research was to provide practical guidelines on dealing with heterogeneity in PLS-SEM when the number of segmentation variables is high (and

so overcome the limitations of pathmox and MGA analyses), a limitation of our research is that we did not perform a simulation study to confirm the potential of this hybrid procedure.

A final point to underline is that results will vary according to the depth of the pathmox tree. However, in a PLS-SEM context, any increase in tree depth will be directly correlated with increased results complexity and greater difficulty in their interpretation. While more complex solutions can be considered according to the interpretability of the results, we recommend a depth of two, corresponding to a maximum of four terminal nodes and six multigroup comparisons.

Table 10 When to use hybrid multigroup PLS-SEM

Criteria	Yes	No
Research context	Exploratory	Confirmatory
Heterogeneity source	Observed	Not observed
Analytical complexity	≥ 2 segmentation variables, many levels	< 2 segmentation variables
Test hypothesis	Open research question	Specific theoretical hypothesis

6. Conclusion

We have described a practical approach to dealing with heterogeneity when the number of categorical variables is high and the research context is exploratory. Our hybrid multigroup PLS-SEM overcomes several limitations associated with classical MGA (incapacity to deal with a large number of segmentation variables or levels and mixed effects of the segmentation variables) and with pathmox analysis (the lack of an a posteriori test of significant differences in path coefficients). As well as providing guidance on how to apply the procedure, we describe an application to real data for an international bank. More specifically, we investigated the most significant drivers of satisfaction and loyalty in employees (empowerment, company reputation, pay, work conditions, and leadership) according to heterogeneity factors (age, gender, marital status, education, job level, and antiquity) that potentially affect perceptions of the work climate. The three distinct groups of employees identified, namely, juniors, seniors, and managers, were found to perceive the work climate in different ways: leadership and pay were more motivational for junior employees, empowerment for seniors, and company reputation

and work conditions for managers. This information is potentially useful for HR managers in profiling employee groups and designing specific HR policies.

7. Declarations

7.1. Funding

Not applicable

7.2. Conflicts of interest/Competing interests

The author declares that he has no conflict of interest

7.3. Availability of data and material (data transparency)

The data that support the findings of this study are not available due to legal policy restrictions.

7.4. Code availability

The SmartPLS 3.2 software and the genpathmox package in R software version 0.4 provided the model estimates (global and multigroup analysis) and the pathmox analysis, respectively.

7.5. Authors' contributions

Not applicable

References

- Abdolshah, M., Khatibi, S.A.M., Moghimi, M.: Factors influencing job satisfaction of banking sector employees. *J. Cent. Bank. Theory Pract.* (2018). <https://doi.org/10.2478/jcbtp-2018-0009>
- Akter, S., Wamba, S.F., Dewan, S.: Why PLS-SEM is suitable for complex modelling? An empirical illustration in big data analytics quality. *Prod. Plan. Control* (2017). <https://doi.org/10.1080/09537287.2016.1267411>
- Ali, M., Lei, S., Wei, X.: The mediating role of the employee relations climate in the relationship between strategic HRM and organizational performance in Chinese banks. *J. Innov. Know.* (2017). <https://doi.org/10.1016/j.jik.2016.12.003>
- Bartel, A.P.: Human resource management and organizational performance: evidence from retail banking. *Ind. Labor Relat. Rev.* (2004). <https://doi.org/10.2307/4126616>
- Becker, J.M., Rai, A., Ringle, C.M., Volckner, F.: Discovering unobserved heterogeneity in structural equation models to avert validity threats. *Manag. Inf. Syst. Q.* (2013). <https://doi.org/10.25300/MISQ/2013/37.3.01>
- Cheah, J. H., Waller, D., Thaichon, P., Ting, H., & Lim, X. J.: Price image and the sugrophobia effect on luxury retail purchase intention. *J. Retail. Consum. Serv.* (2020). <https://doi.org/10.1016/j.jretconser.2020.102188>
- Chih-Hung Wanga, M., Tangb Y.: Examining the antecedents of sport team brand equity: a dual-identification perspective. *Sport Manage. Rev.* (2018). <https://doi.org/10.1016/j.smr.2017.07.010>
- Chin, W.W.: Multi-group analysis with PLS. Retrieved from <http://disc-nt.cba.uh.edu/chin/plsfaq/multigroup.htm> (2000) Accessed 15 May 2020
- Chin, W.W.: A permutation procedure for multi-group comparison of PLS models. In: M. Vilarés, M. Tenenhaus, P. Coelho, V. Esposito Vinzi, A. Morineau (eds.), *PLS and Related Methods: Proceedings of the International Symposium PLS'03* pp. 33-43. Lisbon: Decisia (2003a)
- Chin, W., Cheah, J.-H., Liu, Y., Ting, H., Lim, X.J., Cham, T.H.: Demystifying the role of causal-predictive modeling using partial least squares structural equation modeling in information systems research. *Ind. Manag. Data Syst.* (2020). <https://doi.org/10.1108/IMDS-10-2019-0529>

- Chow, G.C.: Test of equality between sets of coefficients in two linear regressions. *Econometrica* (1960). <https://doi.org/10.2307/1913018>
- Ciavolino, E., Salvatore, S., Mossi, P., Lagetto, G.: High-order PLS path model for multi-group analysis: the prosumership service quality model. *Qual. Quant.* (2019). <https://doi.org/10.1007/s11135-018-0747-0>
- Clark, M.: The relationship between employees' perceptions of organizational climate and customer retention rates in a major UK retail bank. *J. Strateg. Mark.* (2002). <https://doi.org/10.1080/09652540210125260>
- Dijkstra, T.K., Henseler, J.: Consistent and asymptotically normal PLS estimators for linear structural equations. *Comput. Stat. Data Anal.* (2015). <https://doi.org/10.1016/j.csda.2014.07.008>
- Dhir, S., Dutta, T. Ghosh, P.: Linking employee loyalty with job satisfaction using PLS–SEM modelling. *Pers. Rev.* (2020). <https://doi.org/10.1108/PR-03-2019-0107>
- Esposito Vinzi, V., Trinchera, L., Squillacciotti, S., Tenenhaus, M.: REBUS-PLS: a response-based procedure for detecting unit segments in PLS path modelling. *Appl. Stoch. Models Bus. Ind.* (2008). <https://doi.org/10.1002/asmb.728>
- Esposito Vinzi V, Fahmy T, Chatelin YM, Tenenhaus M.: PLS Path Modeling: some Recent Methodological Developments, a Software Integrated in XLSTAT and Its Application to Customer Satisfaction Studies. *Proceedings of the Academy of Marketing Science Conference “Marketing Theory and Practice in an Inter-Functional World”, Verona, Italy, 11-14 July.* (2007)
- Eskildsen, J., Kristensen, K., Antvor, H.G.: The relationship between job satisfaction and national culture. *TQM J.* (2010). <https://doi.org/10.1108/17542731011053299>
- Eskildsen, J., Kristensen, K., Westlund, A.: Work motivation and job satisfaction in the Nordic countries. *Empl. Relat.* (2010). <https://doi.org/10.1108/01425450410511043>
- Eskildsen, J., Kristensen, K., Westlund, A.: Measuring employee assets – The Nordic Employee Index™. *Bus. Process Manag. J.* (2004). <https://doi.org/10.1108/14637150410559216>
- Fordellone, F., Vichi, M.: Finding groups in structural equation modeling through the partial least squares algorithm. *Comput. Stat. Data Anal.* (2020). <https://doi.org/10.1016/j.csda.2020.106957>

- Gaertner, S.: Structural determinants of job satisfaction and organizational commitment in turnover models. *Hum. Resour. Manag. Rev.* (1999). [https://doi.org/10.1016/S1053-4822\(99\)00030-3](https://doi.org/10.1016/S1053-4822(99)00030-3)
- Glaveli, N., Karassavidou, E.: Exploring a possible route through which training affects organizational performance: the case of a Greek bank. *Hum. Resour. Manag. J.* (2011). <https://doi.org/110.1080/09585192.2011.606113>
- Guan, Y., Wen, Y., Chen, S.X., Liu, H., Si, W., Liu, Y., Dong, Z.: When do salary and job level predict career satisfaction and turnover intention among Chinese managers? The role of perceived organizational career management and career anchor. *Eur. J. Work Organ. Psychol.* (2014). <https://doi.org/10.1080/1359432X.2013.763403>
- Guillon, O., Cezanne, C.: Employee loyalty and organizational performance: a crucial survey. *J. Organ. Chang. Manag.* (2014). <https://doi.org/10.1108/JOCM-02-2014-0025>
- Hair, J. F., Sarstedt, M., Ringle, C. M., Gudergan, S.P.: *Advanced issues in partial least squares structural equation modeling*. Thousand Oaks: SAGE Publications (2018)
- Hair, J.F., Ringle, C.M. Sarstedt, M.: PLS-SEM: indeed a silver bullet. *J. Mark. Theory Pract.* (2014). <https://doi.org/10.2753/MTP1069-6679190202>
- Hair J.F., Sarstedt, M., Hopkins L., Kuppelwieser, V.G.: Partial least squares structural equation modeling (PLS-SEM): an emerging tool in business research. *Eur. Bus. Rev.* (2014b). <https://doi.org/10.1108/EBR-10-2013-0128>
- Hair, J.F., Sarstedt, M., Ringle, C.M., Mena, J.: An assessment of the use of partial least squares structural equation modeling in marketing research. *J. Acad. Mark. Sci.* (2012). <https://doi.org/10.1007/s11747-011-0261-6>
- Hair, J.F., Ringle, C.M. Sarstedt, M.: The use of partial least squares (PLS) to address marketing management topics. *J. Mark. Theory Pract.* (2011). Available at SSRN: <https://ssrn.com/abstract=2228902>.
- Helm, S.: Employees' awareness of their impact on corporate reputation. *J. Bus. Res.* (2011). <https://doi.org/10.1016/j.jbusres.2010.09.001>
- Henseler, J., Hubona, G. Ray, P.: Using PLS path modeling in new technology research: updated guidelines. *Ind. Manag. Data Syst.* (2016). <https://doi.org/10.1108/IMDS-09-2015-0382>

Henseler, J., Ringle, C.M., Sarstedt, M.: Testing measurement invariance of composites using partial least squares. *Int. Mark. Rev.* (2016). <https://doi.org/10.1108/IMR-09-2014-0304>

Henseler, J., Ringle, C.M., Sarstedt, M.: A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J. Acad. Mark. Sci.* (2015). <https://doi.org/10.1007/s11747-014-0403-8>

Henseler, J., Ringle, C.M., Sinkovics, R.R.: The use of partial least squares path modeling in international marketing. *Adv. Int. Mark.* (2009). [https://doi.org/10.1108/S1474-7979\(2009\)0000020014](https://doi.org/10.1108/S1474-7979(2009)0000020014)

Herrbach, O., Mignonac, K., Gatignon, A.: Exploring the role of perceived external prestige in managers' turnover intentions. *Int. J. Hum. Resour. Manag.* (2004). <https://doi.org/10.1080/0958519042000257995>

Howat, G., Assaker, G.: The hierarchical effects of perceived quality on perceived value, satisfaction, and loyalty: Empirical results from public, outdoor aquatic centres in Australia. *Sport Manage. Rev.* (2013). <https://doi.org/10.1016/j.smr.2012.10.001>

Huselid, M.A.: The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Acad. Manage. J.* (1995). Available at SSRN: <https://ssrn.com/abstract=1803666>

Hwang, H., Sarstedt, M., Cheah, J. H., Ringle, C. M.: A concept analysis of methodological research on composite-based structural equation modeling: bridging PLSPM and GSCA. *Behaviormetrika.* (2020). <https://doi.org/10.1007/s41237-019-00085-5>

Kaya, N., Koc, E., Topcu, D.: An exploratory analysis of the influence of human resource management activities and organizational climate on job satisfaction in Turkish banks. *Int. J. Hum. Resour. Manag.* (2010). <https://doi.org/10.1080/09585192.2010.505104>

Keil, M., Saarinen, T., Tan, B.C.Y., Tuunainen, V., Wassenaar, A., Wei, K.K.: A cross-cultural study on escalation of commitment behaviour in software projects. *Manag. Inf. Syst. Q.* (2000). <https://doi.org/10.2307/3250940>

Kim, Y., Wang, Q., Roh, T.: Do information and service quality affect perceived privacy protection, satisfaction, and loyalty? Evidence from a Chinese O2O-based mobile shopping application. *Telemat. Inform.* (2020). <https://doi.org/10.1016/j.tele.2020.101483>

- Kock, N.: Common method bias in PLS-SEM: A full collinearity assessment approach. *Int. J. e-Collab.* (2015). <https://doi.org/4018/ijec.2015100101>
- Kristensen, K., Westlund, A.H.: Performance Measurement and Business Results. *Total. Qual. Manag. Bus. Excell.* (2004). <https://doi.org/10.1080/14783360410001680206>
- Lamberti, G., Lopez-Sintas, J., Sukphan, J.: The social process of internet appropriation: Living in a digitally advanced country benefits less well-educated Europeans. *Telecomm. Policy.* (2020). <https://doi.org/10.1016/j.telpol.2020.102055>.
- Lamberti, G., Banet-Aluja, T., Rialp, J.: Work climate drivers and employee heterogeneity. *Int. J. Hum. Resour. Manag.* (2020). <https://doi.org/10.1080/09585192.2020.1711798>
- Lamberti G., Banet-Aluja, T., Sanchez, G.: The Pathmox approach for PLS path modeling: discovering which constructs differentiate segments. *Appl. Stoch. Models Bus. Ind.* (2017). <https://doi.org/10.1002/asmb.2270>
- Lamberti G., Banet Aluja, T., Sanchez, G.: The Pathmox approach for PLS path modeling segmentation. *Appl. Stoch. Models Bus. Ind.* (2016). <https://doi.org/10.1002/asmb.2168>
- Lamberti, G.: R package "genpathmox". Available at: <http://cran.r-project.org/web/packages/genpathmox> (2014)
- Lane, K., Esser, J., Holte, B., Anne, M. M.: A study of nurse faculty job satisfaction in community colleges in Florida. *Teach. Learn. Nurs.* (2010). <https://doi.org/10.1016/j.teln.2009.05.001>
- Lebart, L., Morineau, A., Fenelon J.P.: *Traitement des donnees statistiques*. Paris: Dunod (1979)
- Lim, X. J., Cheah, J. H., Waller, D. S., Ting, H., Ng, S. I. What s-commerce implies? Repurchase intention and its antecedents. *Mark. Intell. Plan.* (2019). <https://doi.org/10.1108/MIP-03-2019-0145>
- Macchitella, L., Marinelli, C. V., Signore, F., Ciavolino, E., Angelelli, P.: Sleepiness, Neuropsychological Skills, and Scholastic Learning in Children. *J. Brain Sci.* (2020). <https://doi.org/10.3390/brainsci10080529>
- Matthews, L.M.: Applying Multi-group Analysis in PLS-SEM: A Step-by-Step Process. In H. Latan and R. Noonan (eds.), *Partial Least Squares Structural Equation Modeling - Basic Concepts, Methodological Issues and Applications*. Berlin: Springer International Publishing AG (2017)

Mahmood, A., Akhtar, M.N., Talat, U., Shuai, C., Hyatt, J.C.: Specific HR practices and employee commitment: the mediating role of job satisfaction. *Empl. Relat.* (2019). <https://doi.org/10.1108/ER-03-2018-0074>

Memon, M. A., Salleh, R., Mirza, M. Z., Cheah, J. H., Ting, H., Ahmad, M. S., Tariq, A.: Satisfaction matters: the relationships between HRM practices, work engagement and turnover intention. *Int. J. Manpow.* (2020). <https://doi.org/10.1108/IJM-04-2018-0127>

Osakwe, C. N., Ruiz, B., Amegbe, H., Chinje, N. B., Cheah, J. H., Ramayah, T.: A multi-country study of bank reputation among customers in Africa: Key antecedents and consequences. *J. Retail. Consum. Serv.* (2020). <https://doi.org/10.1016/j.jretconser.2020.102182>

Puni, A., Mohammed, I., and Asamoah, E.: Transformational leadership and job satisfaction: the moderating effect of contingent reward. *Leadersh. Organ. Dev. J.* (2018). <https://doi.org/10.1108/ER-03-2018-0074>

R Development Core Team R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: <http://www.R-project.org/> (2012)

Reinartz, W., Haenlein, M., Henseler J.: An empirical comparison of the efficacy of covariance-based and variance-based SEM. *Int. J. Res. Mark.* (2009). <https://doi.org/10.1016/j.ijresmar.2009.08.001>

Richard, O.C., Johnson, N.B.: Strategic human resource management effectiveness and firm performance. *Int. J. Hum. Resour. Manag.* (2001). <https://doi.org/10.1080/09585190121674>

Ringle, C.M., Sarstedt, M., Mitchell, R., Gudergan, S.P.: Partial least squares structural equation modeling in HRM research. *Int. J. Hum. Resour. Manag.* Available at SSRN: <https://ssrn.com/abstract=3128200> (2018)

Ringle, C. M., Sarstedt, M., Mooi, E. A.: Response-based segmentation using finite mixture partial least squares: Theoretical foundations and an application to American Customer Satisfaction Index data. *Ann. Inf. Syst.* (2010). https://doi.org/10.1007/978-1-4419-1280-0_2

Ringle, C.M., Schlittgen, R.: A Genetic Algorithm Segmentation Approach for Uncovering and Separating Groups of Data in PLS Path Modeling. In: H. Martens and T.

- Naes. (eds.), Proceedings of the PLS'07 International Symposium, pp. 75-78. Matforsk, As, Norway, (2007)
- Ringle, C. M., Wende, S., Will, A.: SmartPLS 2.0 (Beta). Hamburg: SmartPLS. Retrieved from www.smartpls.de (2005)
- Sanchez, G., Trinchera, L., Russolillo, G. : R package "plsplm". Available at: <http://cran.r-project.org/web/packages/plspm/index.html> (2013)
- Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., Gudergan, S. P.: Estimation issues with PLS and CBSEM: Where the bias lies! *J. Bus. Res.* (2016). <https://doi.org/10.1016/j.jbusres.2016.06.007>
- Sarstedt, M., Henseler, J., Ringle, C. M., Hair, J.F.: On the Emancipation of PLS-SEM: A Commentary on Rigdon. *Long Range Plann.* (2014). <https://doi.org/10.1016/j.lrp.2014.02.007>
- Sarstedt, M., Henseler, J., Ringle, C.M.: Multi-group analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results. *Adv. Int. Mark.* (2011). [https://doi.org/10.1108/S1474-7979\(2011\)0000022012](https://doi.org/10.1108/S1474-7979(2011)0000022012)
- Sarstedt, M., Cheah, J.: Partial least squares structural equation modeling using SmartPLS: a software review. *J. Mark. Anal.* (2019). <https://doi.org/10.1057/s41270-019-00058-3>
- Schlittgen, R., Ringle, C. M., Sarstedt, M., Becker, J.M.: Segmentation of PLS path models by iterative reweighed regressions. *J. Bus. Res.* (2016). <https://doi.org/10.1016/j.jbusres.2016.04.009>
- Sekaran, U.: Paths to the job satisfaction of bank employees. *J. Organ. Behav.* (1989). <https://doi.org/10.1002/job.4030100405>
- Shiau, W.L, Sarstedt, M., Hair, J.: Internet research using partial least squares structural equation modeling (PLS-SEM). *Internet Res.* (2019). <https://doi.org/10.1108/IntR-10-2018-0447>
- Shmueli, G., Ray, S., Estrada, J. M. V., Chatla, S. B.: The elephant in the room: Predictive performance of PLS models. *J. Bus. Res.* (2016). <https://doi.org/10.1016/j.jbusres.2016.03.049>

Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J. H., Ting, H., Vaithilingam, S., Ringle, C. M.: Predictive model assessment in PLS-SEM: guidelines for using PLSpredict. *Eur. J. Mark.* (2019). <https://doi.org/10.1108/EJM-02-2019-0189>

Signore, F., Catalano, A., De Carlo, E., Madaro, A., Ingusci, E.: The role of employability in students during academic experience: a preliminary study through PLS-PM technique. *Electron. J. Appl. Stat. Anal.* (2019). <https://doi.org/10.1285/i20705948v12n4p720>

Tenenhaus, M., Mauger, E., Guinot, C.: Test of a group effect in a regression model relating two blocks of binary variables with ULS-SEM and SEM-PLS. In: *Electronic Proceedings of the Workshop on Knowledge Extraction and Modeling (KNEMO)*, V. Esposito Vinzi, C. Lauro, A. Braverman, H.A.L. Kiers, M.G. Schimek (eds) (2006)

Tlaiss, H.A.: Determinants of job satisfaction in the banking sector: the case of Lebanese managers. *Empl. Relat.* (2013). <https://doi.org/10.1108/ER-10-2011-0064>

Ugboro I.O., Obeng K.: Top management leadership, employee empowerment, job satisfaction, and customer satisfaction in TQM organizations: an empirical study. *J. Qual. Manag.* (2000). [https://doi.org/10.1016/S1084-8568\(01\)00023-2](https://doi.org/10.1016/S1084-8568(01)00023-2)

Ullah, I., Yasmin, R.: The influence of human resource practices on internal customer satisfaction and organizational effectiveness. *J. Internet Bank. Commer.* (2013). <https://doi.org/10.1155/2015/835194>

Valaei, N., Jiroudi, S.: Job satisfaction and job performance in the media industry: a synergistic application of partial least squares path modelling. *Asia Pacific J. Mark. Logist.* (2016). <https://doi.org/10.1108/APJML-10-2015-0160>

Vinzi, V. E., Trinchera, L., & Amato, S. (2010). PLS path modeling: from foundations to recent developments and open issues for model assessment and improvement. In *Handbook of partial least squares* (pp. 47-82). Springer, Berlin, Heidelberg.

Appendix A

Constructs/indicators	Mean	SD	Skewness	Kurtosis
<i>Empowerment</i>				
Emp1 - Recognition of work performed	3.459	0.894	-0.615	0.067
Emp2 - Employee treatment as responsible	3.515	0.832	-0.555	0.23
Emp3 - Teamwork is empowered	3.291	0.92	-0.502	-0.172
Emp4 - Autonomy is favoured	3.239	0.863	-0.348	-0.098
Emp5 - Confidence in performed tasks	3.48	0.853	-0.435	0.168
Emp6 - Creativity and initiative are endorsed	3.231	0.821	-0.373	0.31
<i>Reputation</i>				
Rep1 - Organization's reputation	4.215	0.686	-0.51	0.135
Rep 2 - Organization's values	4.165	0.743	-0.72	0.791
Rep 3 - Organization's customer relationships	3.949	0.718	-0.613	1.386
Rep 4 - Organization's internal relationships	3.914	0.716	-0.371	0.429
Rep 5 - Organization's external projection	3.959	0.738	-0.54	0.91
<i>Pay</i>				
Pay1 - Salary	3.823	0.748	-0.817	1.22
Pay2 - Social benefits	3.672	0.768	-0.49	0.347
Pay3 - My salary corresponds to my duties	3.12	0.915	-0.278	-0.171
Pay4 - My salary corresponds to my effort	3.111	0.932	-0.326	-0.525
<i>Conditions</i>				
Cond1 - Enough personnel in the office	3.242	1.016	-0.463	-0.597
Cond2 - Enough time to perform the tasks	2.924	0.961	-0.151	-0.901
Cond3 - Conditions and tools for work	3.693	0.769	-0.922	1.188
<i>Leadership</i>				
Sup1 - Agenda and planning	3.489	1.042	-0.594	-0.253
Sup2 - Receptiveness	3.914	1.055	-0.966	0.382
Sup3 - Encouraging	3.607	1.054	-0.567	-0.248
Sup4 - Communication	3.326	1.101	-0.376	-0.573
Sup5 - Celebrating success	3.761	1.132	-0.803	-0.102
<i>Satisfaction</i>				
Sat1 - Overall rating of satisfaction	4.037	0.699	-1.043	2.646
Sat2 - Tasks in accordance with capabilities	3.537	0.855	-0.927	0.506
Sat3 - Possibility to know efficiency	3.595	0.768	-0.762	0.696
Sat4 - Possibility to learn new things	3.624	0.839	-0.754	0.512
Sat5 - Usefulness of performed job	3.86	0.658	-0.867	2.157
Sat6 - Fulfilment of expectations	3.489	0.858	-0.857	0.374
<i>Loyalty</i>				
Loy1 – I am unwilling to leave in case of not finding alternative	3.586	0.908	-0.74	0.467
Loy2 - I am committed to the institution	4.134	0.789	-1.03	1.796
Loy3 - I trust in the proper direction of the management	3.84	0.825	-0.651	0.779

Supplementary material

Analysis steps and R code

We recommend consulting `genpathmox` R documentation and Lamberti et al. (2016) for more details on pathmox analysis.

Step 1: Define the global PLS-SEM model

Use smartPLS software

Step 2: Perform the pathmox analysis and define the hybrid segmentation variable

2.1 Install and load the `genpathmox` R package load data

2.1.1 `install.packages(genpathmox)`

2.1.2 `library(genpathmox)`

2.2.2 `Data <- data(jobsat)`

2.2 Define `pls.pathmox` parameters

2.2.1 Define **inner model** (inner parameters)

Empowerment = `c(0,0,0,0,0,0,0)`

Reputation = `c(0,0,0,0,0,0,0)`

Pay = `c(0,0,0,0,0,0,0)`

Condition = `c(0,0,0,0,0,0,0)`

Leadership = `c(0,0,0,0,0,0,0)`

Satisfaction = `c(1,1,1,1,1,0,0)`

Loyalty = `c(0,0,0,0,0,1,0)`

`inner.model = rbind(Empowerment, Reputation, Pay, Condition, Leadership, Satisfaction, Loyalty)`

`colnames(inner.model) = rownames(inner.model)`

`inner.model`

2.2.2 Define **outer model** (list that contains the distinctive blocks of items for each latent variable, with each block identified by the position of the indicators in the data-set)

`outer.model = list(1:6,7:11,12:15,16:18,19:23,24:28,29:31)`

2.2.3 Define **mode** (reflective A or formative B)

`mode.model = c("A","A","A","A","A","A","A")`

- **2.2.4** Other parameters:
- **scheme** is the path scheme by default but can be modified by using the parameter `scheme=factor` or `centroid`.
- **signif** is set to 0.05 (common criterion for the significance of the F-test used in the partition process)
- **deep** is fixed to 2 as we want to keep the maximum tree depth to 2 (i.e., a maximum number of terminal nodes of 4)
- **size** represents the minimum admissible % of observations in a node.

2.2.5 Define the data frame of segmentation variables

`segvar.model= data[,1:5]`

2.3 Run pathmox analysis

`pathmox = pls.pathmox(data, inner=inner.model, outer=outer.model, modes=mode.model, SVAR=segvar.model, signif=0.05, size=0.1, deep =2)`

2.4 Interpret hybrid segmentation variable according to pathmox splits (use the summary function or pathmox\$Fg.r (Pathmox F-global test output)

```
pathmox$Fg.r
```

node	fg.statistic	fg.pvalue	variable	g1.mod	g2.mod
1	21.19	<0.001	jobLevel	Low/Intermediate	High
2	9.79	<0.001	Antiquity	<2000/2000-2004/2005-2009/2009-2014	>2014

2.5 Define the hybrid segmentation variable

```
hybrid.var=pathmox$hybrid
```

3 Export original data and hybrid segmentation variable

```
new.data = data.frame(data, hybrid.var)
```

```
write.csv2(new.data, new.data.csv, row.names = FALSE)
```

4 Run invariance MICOM and multigroup analysis on the hybrid segmentation variable

```
Use smartPLS
```
