European Management Journal 34 (2016) 606–617

Contents lists available at ScienceDirect

European Management Journal

journal homepage: www.elsevier.com/locate/emj

Accounting for sampling weights in PLS path modeling: Simulations and empirical examples

Jan-Michael Becker [a, *], Ida Rosnita Ismail [b, c]

[a] University of Cologne, Department of Marketing and Brand Management, Albertus-Magnus-Platz, 50923 Cologne, Germany
[b] Universiti Kebangsaan Malaysia, Graduate School of Business, Bangi, Malaysia
[c] Graduate School of Business and Hamburg University of Technology (TUHH), Institute of Human Resource Management and Organizations, Am-Schwarzenberg-Campus 4 (D), 21073 Hamburg, Germany

ARTICLE INFO

Article history:
Received 31 January 2016
Received in revised form
20 June 2016
Accepted 21 June 2016
Available online 5 July 2016

Keywords:
PLS path modeling
Weighted PLS (WPLS)
Simulation
Sampling weights
Post-stratification weights
Job satisfaction
Organizational commitment

ABSTRACT

Applications of partial least squares (PLS) path modeling usually focus on survey responses in management, social science, and market research studies, with researchers using their collected samples to estimate population parameters. For this purpose, the sample must represent the population. However, population members are often not equally likely to be included in the sample, which indicates that sampling units have different probabilities of being selected. Hence, sampling (post-stratification) weights should be used to obtain consistent estimates when estimating population parameters. We discuss alterations to the basic PLS path modeling algorithm to consider sampling weights in order to achieve better average population estimates in situations where researchers have a set of appropriate weights. We illustrate the effectiveness and usefulness of the approach with simulations and an empirical example of a job attitude model, using data from Ireland.

1. Introduction

The partial least squares (PLS) path modeling method combines econometric prediction with the psychometric modeling of unobserved (latent) conceptual variables, which are indirectly observed by means of multiple manifest variables (Lohmöller, 1989; Wold, 1982). The method can be applied to different observation forms, such as survey responses and transactional data, although most applications focus on survey responses to psychological variables like job attitude or satisfaction (Rigdon, 2012). PLS path modeling is increasingly applied across behavioral science disciplines, such as business, psychology, sociology, education, and economics (Lu, Kwan, Thomas, & Cedzynski, 2011), with a special focus on strategic management (Hair, Sarstedt, Pieper, & Ringle, 2012), information systems (Ringle, Sarstedt, & Straub, 2012), international business (Richter, Sinkovics, Ringle, & Schlägel, 2016) and marketing (Hair et al., 2012). These researchers usually apply the method to the data as is, aiming to estimate population parameters from the sample they have collected. By doing so, they assume that the sample represents the population very well, but this presumption might not always be fulfilled.

Researchers who work with survey data sometimes conduct complicated sampling designs to obtain a representative sample of the population of interest, or use a simple random sample (Magee, Robb, & Burbidge, 1998). Others might just use a convenience sample without a specific sampling strategy. Whatever the case, population members may not be equally likely to be included in the sample, which means the sampling units (observations) have different probabilities of selection compared with their occurrence in the population (Winship & Radbill, 1994). If the analysis does not incorporate the unequal probability of selection, a substantial bias may arise in the parameter estimates (Asparouhov, 2005; Pfeffermann, 1993). For instance, the relevance of this issue has been shown in regression analyses (e.g., Korn & Graubard, 1995) and in latent variable structural equation model (SEM) analyses (e.g., Kaplan & Ferguson, 1999). The use of sampling weights is a possible solution to correct the results with, for example, weighted means or weighted variances, when estimating population parameters (DuMouchel & Duncan, 1983). Not only can imperfections

* Corresponding author.
E-mail addresses: j.becker@wiso.uni-koeln.de (J.-M. Becker), idarosnita@ukm.edu.my (I.R. Ismail).

http://dx.doi.org/10.1016/j.emj.2016.06.009
0263-2373/

in the sample due to unequal probabilities of selection be corrected by applying appropriate weights, but also imperfections in terms of unit nonresponse and noncoverage (Groves, Dillman, Eltinge, & Little, 2002).

In regression analysis, for example, weighted least squares (WLS) is used to account for sampling weights to obtain consistent population parameters (DuMouchel & Duncan, 1983). In the context of sampling weights, this regression estimator only differs from the usual application of WLS for heteroskedastic errors by its motivation for and choice of weights (Magee et al., 1998). In terms of covariance-based structural equation modeling (CB-SEM), Asparouhov (2005) shows how to incorporate sampling weights via the WLS estimator into SEMs with mixed outcomes (Muthén, 1984).[1] In PLS path modeling, to the best of our knowledge no similar technique has been proposed to take sampling weights into account.

The purpose of this study is to close this research gap by (1) developing and testing a new weighted partial least squares (WPLS) path modeling algorithm that provides consistent population estimates for situations where sampling or post-stratification weights are available, and (2) familiarizing PLS users with the concept of sample and post-stratification weights. This article is organized as follows: First, it discusses the use of sample and post-stratification weights and a few basic guidelines on how to generate such weights. Second, it briefly explains the steps required to alter the basic PLS path modeling algorithm to consider sampling and post-stratification weights. Third, it tests the new method in an illustrative example and a small simulation study. We show the appropriateness of the new method and contrast it with other simple (naive) weighting strategies. In order to illustrate the application of the sampling weights in management research, we provide an empirical example by comparing weighted and non-weighted data in a job attitude model using the SmartPLS 3 software (Ringle, Wende, & Becker, 2015), which implements our suggested WPLS approach. This study concludes with a summary of the results and suggestions for future research directions.

2. Sampling and post-stratification weights

Having a representative sample of the population is of paramount importance when conducting an empirical survey study, if correct inference regarding the population is the objective. Despite their best efforts, researchers may accidentally or intentionally oversample some observational units and undersample others. In other words, the way a certain characteristic of the population (such as age, education, and gender) is distributed in the sample may differ from the way it is distributed in the population (Gelman, 2007). For example, the sample may consist of 60% women, although they comprise only 40% of the population. This introduces bias in the estimates from the sample, because the statistical procedures will assign more weight to the observational units that are oversampled. Assuming that there are differences between women and men in the sample, these differences affect the parameter estimations and bias the results if average population quantities are estimated (e.g., Asparouhov, 2005; Pfeffermann, 1993).

Using sampling or post-stratification weights is one way to correct these biases (e.g., Gelman, 2007). Sampling weights can be generated before data collection based on the sampling design. The sampling design may feature intentional oversampling or undersampling of certain observational units to reduce the survey cost or the robustness of the data. For example, researchers may oversample certain observational units with a low occurrence in the population to obtain a more stable estimate for them.

Post-stratification weights[2] are a subcategory of sampling weights, and are calculated based on defined strata, that is, specific groups of observational units with the same characteristics. In order to calculate sampling or post-stratification weights, researchers need auxiliary data to which the sample data can be compared (e.g., Little, 1993). Table 1 shows an example of the US American population from age 15 and a sampling distribution to illustrate the mechanism. The combinations of age groups and gender in this example are the strata, the categories of interest where differences in the response style might be possible. Sampling weights adjust for the unequal probabilities of being included in the sample.

In this example, we observe differences between the auxiliary variables of age and gender. In particular, younger age groups are oversampled and older age groups are undersampled. In addition, inequalities in terms of gender appear in some of the age groups. The weights that can be calculated from this information downweight the strata that are oversampled and assign a higher weight to the undersampled strata. The goal is to generate weights such that the distribution of the sampling weights is in agreement with the known auxiliary information, such as the census in this example. In particular, the weights are the inverse of the likelihood of inclusions; that is, the probability of occurrence in the population divided by the probability of occurrence in the sample (Little, 1993). However, more complex sampling weights can be created based on complex survey designs (Cameron & Trivedi, 2005).

It becomes obvious that without precise information about the population, the calculation of weights is an impossible endeavor. However, with appropriate information about the population and the sample, the weights can be easily calculated.

It is quite common for market researchers and social scientist to use sampling weight techniques in survey data analysis (e.g., Biemer & Christ, 2008; Cameron & Trivedi, 2005). Although survey research is common in the business disciplines, one of our observations is that sampling weights are rarely applied in studies published in the academic journals of business disciplines such as management, information systems, and marketing. Slater and Atuahene-Gima (2004), for example, provide a good overview of sampling designs and procedures in strategic management research, but overlook the possible use of sampling weights.

One reason for this neglect could simply be authors' limited knowledge about the use of such weights and/or advanced modeling tools' (like PLS path modeling) limited support regarding incorporating such weights. Another reason could be that researchers struggle to define their population precisely, or that auxiliary population information is often not available, which is a prerequisite for calculating the weights. The use of student or convenience samples, which offer easy access to observations, but are unlikely to represent the population of interest (except for rare cases where the population of interest are of university students), corroborates the assumption that researchers have problems defining their population precisely. Rungtusanatham, Choi, Hollingworth, Wu, and Forza (2003), for example, report that only 53% of the analyzed studies in major operations management journals report the sampling process or the target population sufficiently. Inferences from such samples are therefore always biased

[1] Note that the WLS estimator in structural equation modeling is different from the WLS estimator used in a regression context. The WLS estimation in SEM is more commonly called ADF (asymptotic distribution-free) estimation (Browne, 1982; 1984). In contrast to the problem described in this study (i.e., sample and post-stratification weights that are defined based on sample quantities), the weights in a normal WLS or ADF estimation of an SEM are estimated from the fourth-order moments of the residuals (Browne, 1984).

[2] The remainder of this article will only use the term sampling weights, which includes post-stratification weights.

Table 1
Example of post-stratification weights of US census.

Auxiliary variables						
Age (years)	Gender	Population[a] N	Population %	Sample N	Sample %	Weight
15–24	Male	22,473,687	8.63	400	20	0.432
15–24	Female	21,358,609	8.20	600	30	0.273
25–54	Male	63,838,086	24.52	200	10	2.452
25–54	Female	63,947,036	24.56	300	15	1.637
55–64	Male	19,731,664	7.58	200	10	0.758
55–64	Female	21,172,201	8.13	100	5	1.626
65+	Male	21,129,978	8.12	100	5	1.623
65+	Female	26,700,267	10.26	100	5	2.051
	Σ	260,351,528		2000		

[a] Population information for the US population for age and gender from http://www.cia.gov/library/publications/resources/the-world-factbook/geos/us.html (2015 est.; Retrieved September 04, 2015).

and their conclusions might not be meaningful, if the reference population was not correctly defined, or the sampling procedures were inadequate.

It should also be noted that the use of sampling weights is not without limitation or constraints. The use of sampling weights as a data manipulation strategy should be applied carefully and should not be used to make results more appealing (e.g., "fishing expeditions" or "p hacking"). The construction and use of sampling weights should follow a rigorous process and should fulfill the basic requirements and assumptions (e.g., Cameron & Trivedi, 2005; Gelman, 2007; Shin, 2012). The approach that we suggest is only appropriate if the following four key requirements for the use of sampling weights are met:

1) The auxiliary variables in the sample need to be measured in the same way as those for the population. This requirement is usually easy to fulfill with variables like age and gender, but more complicated when constructs' different definitions or measurement play a role. For example, there are many different definitions of small and medium-sized enterprises (SMEs) and of employment status. Careful assessment of the equivalence of the measured variables and their definitions is necessary when using them for weighting purposes.
2) The auxiliary variables that define the weights should not include the dependent variable. For example, if income is a dependent variable in the model, it should not be used to define the sampling weights that correct underrepresentation in terms of the income distribution. Using the dependent variable for weighting would bias the results.
3) There should be an adequate number of sample elements in all strata to allow the efficient estimation of the respective parameters. For example, if the strata of males aged >65 years in our example were to include only five observations, these might be assigned a very large weight. Single sampling units' idiosyncrasies might therefore be multiplied and the parameter estimates might be worse than when the sampling inequalities are ignored. Hence, the observations within a stratum also need to fulfill the minimum requirements of representativeness. By contrast, oversampling is an often applied strategy to increase the precision of the estimation for strata with only a few population members.
4) The approach requires selecting the sample elements in each stratum at random. Unbiased and efficient estimates cannot be expected from a nonprobability sample, even if it uses sampling weights (Shin, 2012).

Finally, on a more general basis, using sampling weights only makes sense if the (regression) model's assumptions do not fully hold. If the model is correctly specified, the estimates of weighted and unweighted samples should have the same probability limit and both should be consistent (Cameron & Trivedi, 2005). Most empirical models, however, usually include a certain degree of misspecification. For example, it is rare for a model to be specified without any omitted regressors, not have any collinearity issues, *and* be perfectly linear in all relationships. Hence, sample weighting is useful in a broad range of applications and a quite common and established practice in market research. Nevertheless, if the model is correctly specified, the WLS estimator for regressions is less efficient than the OLS estimator (Cameron & Trivedi, 2005; Korn & Graubard, 1995). Thus, when researchers unnecessarily apply sampling weights, they can create an inefficient estimator without reducing bias (Bollen, Biemer, Karr, Tueller, & Berzofsky, 2016). Bollen et al. (2016) provide an overview of different tests and their efficacy in assessing the appropriateness of sampling weights.

3. WPLS – weighted partial least squares

In the following, this study assumes that a researcher has a variable with meaningful and appropriate sampling weights W that he/she wants to incorporate into a PLS path modeling study to consider the sampling process. The WPLS algorithm needs these weights as input data. It is not possible to calculate the weights from the sample characteristics in the WPLS algorithm, but this calculation needs to be done before the PLS analysis.

The intention is to achieve a refined version of the basic PLS path modeling algorithm that is similar to the standard WLS approach in an ordinary least squares (OLS) regression context, where the regression parameters β_w are estimated by

$$\beta_w = \left(X^T W X\right)^{-1}\left(X^T W Y\right), \quad (1)$$

with X an $n \times p$ data matrix of p independent variables with n observations, Y an $n \times 1$ data matrix of a single dependent variable, and W a diagonal $n \times n$ matrix of weights.

With regard to sampling weights, this regression estimator only differs from the usual application of WLS for heteroskedastic errors in its motivation for and choice of weights W (Magee et al., 1998). Hence, the WLS estimator is not limited to heteroskedasticity problems.

PLS path modeling is a system of regressions on standardized indicator data, and it calculates weighted composites as representative of the conceptual latent variables in an iterative algorithm. A naive approach to incorporating sample weighting into PLS path modeling would be to use the data set D, which includes the observed manifest variables used to measure the conceptual latent variables in PLS path modeling and generate a weighted data set with D_w by using the square root of weights W and multiplying

Table 2
Results of the illustrative example for PLS, WPLS and preweighting.

	Original PLS		WPLS		Original PLS with first 100 observations		Pre-weighted PLS on unst. data		Pre-weighted PLS on stand. data	
	Path coeff.	*p* value	Path coeff.	*p* value	Path coeff.	*p* value	Path coeff.	*p* value	Path coeff.	*p* value
Loyalty										
Complaints	0.07	0.26	0.15	0.06	0.15	0.06	0.14	0.13	0.14	0.08
Image	0.20	0.01	0.05	0.69	0.05	0.69	0.23	0.21	0.04	0.71
Satisfaction	0.49	0.00	0.56	0.00	0.56	0.00	0.61	0.00	0.57	0.00
Satisfaction										
Expectation	0.06	0.21	−0.01	0.84	−0.01	0.84	−0.06	0.45	−0.01	0.93
Image	0.18	0.00	0.14	0.20	0.14	0.21	0.10	0.36	0.15	0.18
Quality	0.51	0.00	0.60	0.00	0.60	0.00	0.80	0.00	0.58	0.00
Value	0.20	0.00	0.23	0.03	0.23	0.03	0.16	0.03	0.23	0.03
Value										
Expectation	0.05	0.53	0.14	0.23	0.14	0.24	0.25	0.25	0.14	0.22
Quality	0.56	0.00	0.40	0.01	0.40	0.01	0.71	0.00	0.40	0.01
Quality										
Expectation	0.56	0.00	0.46	0.00	0.46	0.00	0.98	0.00	0.45	0.00
Expectation										
Image	0.51	0.00	0.49	0.00	0.49	0.00	0.98	0.00	0.49	0.00
Complaints										
Satisfaction	0.53	0.00	0.65	0.00	0.65	0.00	0.96	0.00	0.65	0.00

them with every observation. Such an approach can be used to appropriately weight an OLS regression to obtain WLS results:

$$\beta_w = \left(X^T W X\right)^{-1}\left(X^T W Y\right) = \left(X_w^T X_w\right)^{-1}\left(X_w^T Y_w\right). \quad (2)$$

By contrast, such a procedure might cause problems in a PLS setting. The main reason for such problems is that PLS path modeling relies highly on standardized data during its iterative algorithm. The formed composites' scores are always standardized (i.e., they have a zero mean and standard deviation of one). Hence, a precomputation data-weighting strategy might not work, because the resulting data set is used for the calculation of weighted composites as proxies for the conceptual latent variables, which will be standardized during the iterative algorithm. Standardizing a pre-weighted data set will obscure the underlying correlation pattern, because multiplication with a nonconstant value is not commutative. Thus, there is a difference between correlations retrieved from weighted standardized and standardized weighted data.

Nevertheless, the appropriate correction of the PLS path modeling algorithm is not complex. The algorithm only needs to take the weights into account in each calculation of the mean, the variance, and the covariance (correlation) of the components in the iterative algorithm. For example, each time the basic PLS path modeling algorithm performs a standardization of a variable, it should use the weighted mean and the weighted variance of the variable rather than the normal mean and normal variance:

$$\text{Weighted mean}: \ \mu_w(X) = \frac{\sum_{i=0}^{n} w_i x_i}{\sum_{i=0}^{n} w_i} \quad (3)$$

$$\text{Weighted variance}: \ \sigma_w(X) = \frac{\sum_{i=0}^{n} w_i (x_i - \mu_w)^2}{\left(\sum_{i=0}^{n} w_i\right) - 1} \quad (4)$$

$$\text{Weighted covariance}: \ \sigma_w(X,Y) = \frac{\sum_{i=0}^{n} w_i (x_i - \mu_w(X))(y_i - \mu_w(Y))}{\left(\sum_{i=0}^{n} w_i\right) - 1} \quad (5)$$

In addition, correlations (e.g., to calculate the weights in PLS Mode A) should use weighted covariance and weighted standard deviations. Finally, all regressions (i.e., the inner model regression and those for Mode B outer models) should use weighted standardized data (i.e., WLS regressions). Thus, a weighted standardized beta coefficient β_{sw} is given by

$$\beta_{sw} = cor_w(X)^{-1} cor_w(X,Y). \quad (6)$$

The effect of these corrections is that all the calculations during the iterative PLS algorithm (e.g., estimation of outer weights, outer loadings, and inner weights) are appropriately weighted with the sampling weights, while retaining all information from the original data set in the model.

4. Evaluation of the new WPLS method

In order to ensure that the new WPLS method provides correct estimates, we use an illustrative example with a specific set of demonstrative weights to highlight the differences between the WPLS and the strategy of multiplying the data matrix with the square root of weights before running the PLS computations. In addition, we will use a simple simulation study to verify the new method's appropriateness with sampling weights and evaluate its performance compared with that of the basic PLS algorithm.

4.1. Illustrative example

We will use the ECSI data set provided by Tenenhaus, Vinzi, Chatelin, and Lauro (2005) for the illustrative example. It is a classic example for a medium complex PLS path model with seven conceptual latent variables (Image, Expectation, Quality, Value, Satisfaction, Complaints, and Loyalty) and 24 observed variables.

A specific weight variable $W_{1/0}$ will serve the illustrative purpose of showing the differences between (1) the new WPLS algorithm, (2) the original PLS algorithm, and (3) the preweighted PLS path modeling results. The variable $W_{1/0}$ contains values of either one (for the first 100 observations) or zero (for the remaining 250 observations). This weighting variable should be equal to a model where only the first 100 observations are used.

The results of the three approaches (i.e., original PLS algorithm, WPLS algorithm, and preweighted PLS) are reported in Table 2. In addition, the study estimates an original PLS path model with only the first 100 observations.

We preweight the unstandardized data set as well as the standardized data set. Preweighting the unstandardized data merely serves illustrative, cautionary purposes. As the 10-point Likert scale of the observed variables in the ECSI data set does not have a natural zero (ranging from 1 to 10), weighting observations to zero will skew the data and make the results incompatible.

The standard error and the respective *p* values for the path coefficients were estimated using the bootstrapping procedure[3] with 1000 resamples (Henseler, Ringle, & Sinkovics, 2009). The results (Table 2) show a substantial difference between the original PLS and the WPLS algorithm in terms of the path coefficient estimates and their significance. In addition, there is a strong difference between the WPLS and preweighted PLS. This difference shows that the two approaches are not equivalent (unlike in a traditional OLS–WLS setting). The obvious question that arises from this example is: Which method provides the correct results?

A first strong indication of the answer to this question is that the preweighted PLS results are different from a PLS path model's result that comprises only the first 100 observations. By contrast, the WPLS results match the PLS path model's result of a calculation of the first 100 observations exactly. This finding indicates the appropriateness of the WPLS algorithm, which is in contrast with the preweighting, because a weighting with $W_{1/0}$ should result in a model in which only the first 100 observations are used. Hence, the preweighted PLS model's results do not seem to be correct if the weight vector contains zeros.

The difference is not surprising, because the WPLS does not use the information from those observations that are weighted zero in any calculation of the mean, variance, and covariance. By contrast, the preweighting of the standardized data results in many observations that are zero, but still contribute to the means and the variances. The preweighted data are no longer standardized after multiplication with the weights. Consequently, the inclusion of the observations in the subsequent calculations results in biased means and variances. These differences might appear small, but could be substantial in specific settings. The bias also transfers to the calculation of standard errors and subsequent *t* and *p* values, as well as confidence intervals in the bootstrapping procedure.

4.2. Simulation study with sampling weights

In order to evaluate the new WPLS method with sampling weights further, this study conducts a simulation study that uses a population model with two different types of observational units (e.g., male and female). These units have different underlying path model parameters and are included in the sample with different probabilities.[4] In particular, one type of observational unit constitutes 60% of the population and the other type 40%. The path model is a simple three-construct model with two exogenous latent variables (X_1 and X_2) and one endogenous variable (Y). Each of these has three observed manifest variables. The loading pattern for each construct is {0.70, 0.80, 0.90}. The population path coefficients parameters are 0.20 and 0.60 for the first type of observations and vice versa for the other type. This results in average population parameters of 0.36 and 0.44 (i.e., $0.20 \times 60\% + 0.60*40\% = 0.36$ and $0.60*60\% + 0.20*40\% = 0.44$; Fig. 1).

[3] The bootstrapping is adapted in a way that it draws not only the subsamples from the data matrix but also the corresponding weights for every observation from the weighting variable.

[4] Note: Such situations can also be explained in terms of unobserved heterogeneity (e.g., Becker et al., 2013). Instead of using sampling weights to obtain the true average population parameter, researchers may want to uncover the unobserved differences and use homogenous subsamples (i.e., groups) to estimate valid parameters for the subpopulations.

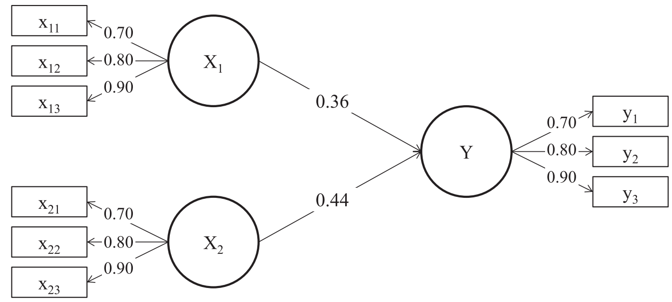

Fig. 1. Average population model for the simulation.

With the prespecified parameters, the simulation study generates a population data set of one million observations.[5] Applying the basic PLS algorithm to the population data results in PLS population parameters of 0.308 and 0.377 because of the known PLS bias (Table 3). By using consistent PLS (PLSc; e.g., Dijkstra & Schermelleh-Engel, 2014; Dijkstra & Henseler, 2015) to avoid the PLS bias, the population data show a perfect fit with the population model.

The study then randomly samples 1000 samples, each with a size of 1000, from the population data set. First, the study samples according to the distribution of the units in the population, so that units of type *one* are chosen with 60% probability and those of type *two* are chosen with 40% probability. The corresponding weights would be 1.0 for all the observations in the sample, because there is no need to correct the sampling procedure. In this situation, the weighted und unweighted PLS results are the same and the average estimates over all the samples are very close to the PLS population data parameters, that is, 0.309 and 0.377 for both models (Table 3).

Then, the study samples differently to the population distribution, so that type *one* units have a 40% probability of being chosen and type *two* units have a 60% probability. The corresponding weights in this situation are 1.5 for type *one* units and 0.666 for type *two* units. In general, as discussed in Section 2, because of the sampling design, the weights represent the inverse of the probability of being included in the sample. For example, if the probability in the population is 60% and in the sample 40%, the weight is simply $60\%/40\% = 1.5$. In this simple weighting design, the weights usually sum up to the sample size. However, more complex corrections for multistage designs or unit nonresponse and noncoverage could also be applied and would alter the sampling weights. In practice, the researcher needs information about the population, as well as about the corresponding auxiliary variables for the sample (e.g., gender, age, and employment status). In our simulation, all the requirements for sampling weights are fulfilled.

The results of the uneven sampling in the simulation study reveal the expected differences between the unweighted and weighted PLS results. The unweighted mean parameters over all the samples are 0.378 and 0.309, while the weighted PLS mean parameters are 0.310 and 0.377. We observe that the WPLS results correspond more to the PLS population data than the unweighted basic PLS results, which deviates largely from the population parameters. The parameter estimates for preweighted original PLS also deviate from the population parameters (0.438 and 0.498).

[5] Data were generated by using the R framework (R Core Team, 2014) by means of the MASS library (Venables & Ripley, 2002). The subsequent calculations use an implementation of the WPLS algorithm in SmartPLS 3 (Ringle et al., 2015).

Table 3
Simulation results of the path coefficients.

	$X_1 \rightarrow Y$ Mean (Sd)	$X_2 \rightarrow Y$ Mean (Sd)
Population model	*0.36*	*0.44*
Population data	0.308	0.377
PLS path model (60/40)		
Correct sampling (60/40) – weighted and unweighted PLS	0.309 (0.025)	0.377 (0.026)
Uneven sampling (40/60) – unweighted (original) PLS	0.378 (0.026)	0.309 (0.027)
Uneven sampling (40/60) – weighted PLS (WPLS)	0.310 (0.028)	0.377 (0.028)
Uneven sampling (40/60) – Pre-weighted original PLS	0.438 (0.022)	0.498 (0.022)
Population data	0.360	0.440
PLSc path model (60/40)		
Uneven sampling (40/60) – unweighted (original) PLSc	0.441 (0.030)	0.361 (0.031)
Uneven sampling (40/60) – weighted PLSc (WPLSc)	0.361 (0.033)	0.440 (0.033)
Uneven sampling (40/60) – Pre-weighted original PLSc	0.457 (0.025)	0.526 (0.025)

Similarly to the WPLS estimates, this procedure retains that X_2 has a larger coefficient than X_1, but, compared with the PLS population data, the coefficients are upward biased. This finding confirms that preweighting is not an appropriate way to consider sampling weights.

Finally, the PLSc results show that the weighted version of PLSc (similar changes as previously discussed are required) recovers population parameters, while the unweighted version produces biased estimates, as does a preweighting of the data (Table 3).

The results show that if sampling weights are used (i.e., in the case that the observations' probability of being included in the sample does not match their occurrence in the population), the WPLS and WPLSc methods provide far better estimates than the unweighted basic PLS and PLSc algorithms, or a preweighted PLS analysis, and are closer to the population parameters.

5. Empirical job attitude example

In order to illustrate the utility of sampling weights in a PLS context in management research, we test the application of WPLS by using a simple job attitude model. This parsimonious job attitude model (Fig. 2) comprises Hult (2005) three main concepts: job features, job satisfaction, and organizational commitment. In his study, Hult (2005) finds that extrinsic features, intrinsic features, and societal features cover three job feature dimensions. Past studies also showed that these three features have direct effects on job satisfaction and organizational commitment (e.g., Hauff & Richter, 2015; Hult, 2005).

In line with this argument, we hypothesize that extrinsic features, intrinsic features, and societal features affect job satisfaction and organizational commitment directly. In addition, we hypothesize that job satisfaction mediates the relationship between job features and organizational commitment. When employees experience positive extrinsic, intrinsic, and societal job features, they are likely to feel more satisfied in their current job. This makes them more committed to the organization by being willing to work harder, to help the organization succeed, feeling proud of working for the organization, and turning down other jobs with relatively higher pay to stay in the organization.

Our model comprises five latent variables. We model extrinsic features (EXTR), intrinsic features (INTR), societal features (SOCI), and organizational commitment (ORGC) as reflective constructs, because these indicators represent the latent variables' manifestations (Hair, Hult, Ringle, & Sarstedt, 2016). This approach is in line with the dimensional analysis by Hult (2005). Using the ISSP 2005 work orientations data and following Hult's job feature dimensions,

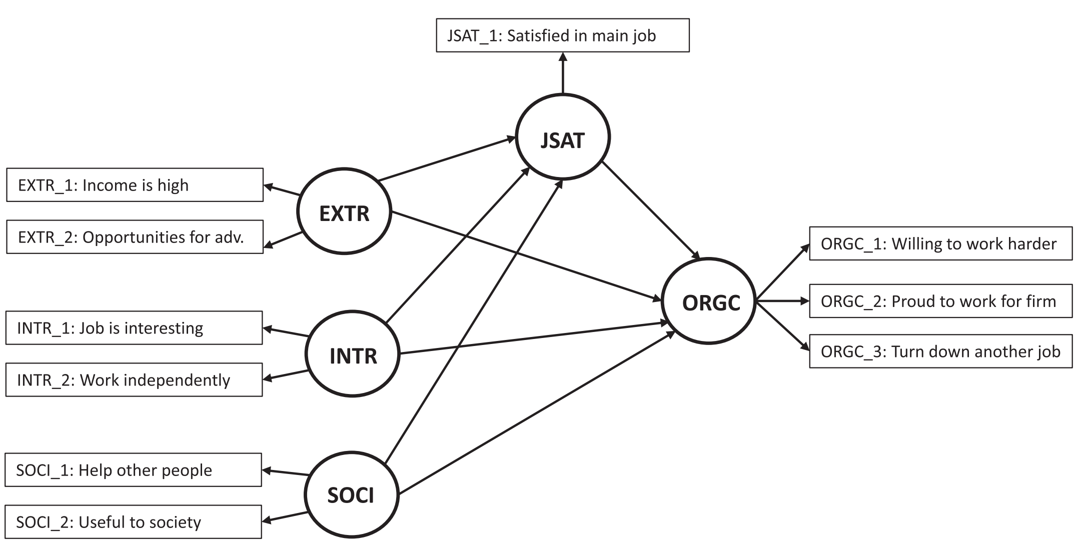

Fig. 2. Parsimonious model of job attitude.

Table 4
Latent variables, indicators, and scale.

Latent variable	Indicators	Scale
Extrinsic features	My income is high. Opportunities for advancement are high.	5-point Likert-type scale 1 (strongly agree) to 5 (strongly disagree)
Intrinsic features	My job is interesting. I can work independently.	
Societal features	In my job, I can help other people. My job is useful to society.	
Job satisfaction	How satisfied are you in your (main) job?	7-point Likert-type scale 1 (completely satisfied) to 7 (completely dissatisfied)
Organizational commitment	I am willing to work harder than I have to in order to help the firm I work for to succeed. I am proud to be working for my firm. I would turn down another job that offered quite a bit more pay in order to stay with this firm.	5-point Likert-type scale 1 (strongly agree) to 5 (strongly disagree)

we assign two to three indicators to each reflective construct. Job satisfaction (JSAT), on the contrary, is modeled as a single-item construct, because there is only one item that measures the overall job satisfaction in the ISSP data that we use. Single indicators work well for concrete constructs that can be asked directly to respondents, for example, their overall satisfaction (Diamantopoulos, Sarstedt, Fuchs, Wilczynski, & Kaiser, 2012; Hair et al., 2016; Wanous, Reichers, & Hudy, 1997).

5.1. Data

The data used in testing our model were obtained from the 2005 Work Orientations module available on the ISSP website. ISSP is an established cross-national collaboration research project that has carried out routine and recurrent surveys of important social science areas since 1984 (GESIS, 2015). The Work Orientations module is one of the 11 ISSP modules, focusing on work attitude, work orientation, and the description of work contents (ISSP Research Group, 2013). Data on the Work Orientations module are available for 1989, 1997, and 2005. 32 countries participated in the 2005 surveys, comprising a total sample size of 44,368 responses (GESIS, 2015).

As shown in Table 4, all the measurement indicators were adopted from the ISSP 2005 Work Orientations module. Following Hult (2005) dimensional analysis, we categorized the job features into extrinsic features, intrinsic features, and societal features. The indicator scores for job features and organizational commitment varied from 1 (strongly agree) to 5 (strongly disagree), whereas the indicator scores for job satisfaction varied from 1 (completely satisfied) to 7 (completely dissatisfied).

The ISSP data also include a country-level weighting variable to correct the imbalance in the country subsample, which can be either overrepresentative or underrepresentative regarding certain respondent characteristics. The sampling procedure is classified as a “multistage stratified random sample” (ISSP Research Group, 2013). The weighting variable combines all the corrections of the relevant sample characteristics into one weight per respondent. In the remainder of the manuscript, we assume that the assumptions and prerequisites of sampling weights have been fulfilled, although we have limited information on the detailed sampling procedure and weight calculation. We note this as a limitation of our empirical example.

The weighting in the ISSP data set makes international comparison problematic, as the weights are calculated differently across countries.[6] We therefore selected data from a single country, Ireland, for this study. Besides the availability of sampling weights for this country, the sample characteristics match the illustrative purpose of our analysis, as the sample is not too large. Having a very large sample could undermine the effect of the sampling weights, because the differences between the respondents become less important.

The total number of observations available in the data set from Ireland is $N = 1001$. We selected respondents working for pay in all types of occupations at the time of the data collection process, because the focal job attitude indicators are only available for those respondents. This selection allowed us to analyze the job attitudes with respect to a certain company for which the respondents work. We removed 41 observations with missing values, using case-wise deletion. Therefore, only $n = 497$ observations were usable to test the proposed model.

Although the weights have to be recalculated after removing the missing values, we were unable to do so, because the exact calculation procedure that would have allowed us to recalculate the weights according to the original process was not available. Hence, we use the existing weights in this empirical example to illustrate the importance of sample weighting, but also note this limitation.

Table 5 shows the descriptive statistics of the sample. The majority of the respondents was female (55%), married (57.1%), and had completed a higher secondary education (27.1%). Their average age was 39.60 years ($SD = 11.77$). A large number of respondents worked for private firms (60.2%) and were full-time employees (76.3%). They also have various occupations, with the majority working as service workers, as well as shop and market sales workers (18.6%).

5.2. Model estimation and results

We applied the basic PLS algorithm procedure, the bootstrapping procedure with 5000 resamples,[7] and the blindfolding

[6] The availability of the weighting factor depends on the participating countries. Some countries, such as Great Britain, France, and Canada, provided weighting factors in their data, but others, such as Australia, Germany, and Japan, only provided unweighted data. The weighting procedures also vary among the countries that provided a weighting factor in the data. Switzerland, for example, used three criteria to calculate the weight, which are the probability of the respondents being selected in their households, age, and gender strata. The Netherlands, on the contrary, used sampling weights based on national benchmark, sampling frame, and household roster information (GESIS, 2015).

[7] We used the “no sign changes” option and checked the histograms of the bootstrap distributions of direct, indirect, and total effects for non-normal or bimodal distributions. We found more or less normally distributed coefficients and, specifically, no bimodal distributions. Hence, interpretation of the t and p values of the bootstrapping procedure is possible.

Table 5
Descriptive statistics.

Gender	45% male, 55% female
Age	Average = 39.60, standard deviation = 11.77, minimum = 18, maximum = 64
Marital status	Single = 33.3%, married = 57.1%, others = 9.5%
Highest education level	No formal qualification = 0.6%, lowest formal qualification = 9.3%, above lowest qualification = 17.4%, higher secondary completed = 27.1%, above higher secondary level = 24.4%, university degree completed = 21.2%
Employment status	Full-time employment = 76.3%, part-time employment = 21.7%, others = 2%
Occupational classifications (ISCO88)	Armed forces = 0.4%, legislators, senior officials and managers = 12.8%, professionals = 18.4%, technicians and associates professionals = 9.1%, clerks = 13.6%, service workers and shop and market sales workers = 18.6%, skilled agricultural and fishery workers = 5.1%, craft and related trades workers = 8.9%, elementary occupations = 5.1%, plant and machine operators and assemblers = 8.1%
Work sector	Government = 26.4%, private firms = 60.2%, others = 13.4%

procedure, with an omission distance of 7 to the job attitude model, using the same data *with* and *without* sampling weights.

The assessment of the model follows a two-step process, including reflective measurement model and structural model assessment (Hair et al., 2016). Four criteria, including indicator reliability, internal consistency reliability, convergent validity, and discriminant validity, are used to assess the reflective measurement model (Hair et al., 2016). Following the recent development in assessing construct discriminant validity in PLS-SEM, we also apply Henseler, Ringle, and Sarstedt's (2015) new discriminant validity criterion, the heterotrait–monotrait ratio of correlations (HTMT) in addition to the established Fornell–Larcker criterion. Henseler, Ringle, and Sarstedt (2015) suggest that $HTMT_{.85}$ and $HTMT_{.90}$ are two suitable thresholds when assessing discriminant validity. We use the more conservative $HTMT_{.85}$, where an HTMT value < 0.85 indicates that discriminant validity has been established.

In the following section, the measurement model assessment results are presented separately regarding the model without weighting (nonweighted model) and the model with sampling weights (weighted model). The structural model results are directly compared between the two models.

Table 6 shows the results of the reflective measurement model regarding the nonweighted and weighted models. In the nonweighted model, all the indicator loadings are above 0.70. Although the composite reliability of all the constructs is above 0.70, Cronbach's alpha for EXTR and INTR is below 0.70. The poor results of the Cronbach's alpha are not surprising. According to Streiner (2003), the length of a measurement scale affects Cronbach's alpha values. Hence, a scale's lack of sufficient length is a possible explanation for the low Cronbach's alpha in our model, because only two indicators are assigned to measure EXTR and INTR. However, as shown in Table 3, the lowest and highest values of AVE are 0.652 (ORGC) and 0.815 (SOCI), which are satisfactory for the reflective measurement model. In addition, discriminant validity has been established, as evidenced by the cross-loadings, Fornell–Larcker criterion, and the new HTMT assessment (Appendix A, Tables A.1–A.3).

In the weighted data model, one indicator (ORGC_3) has a loading lower than 0.70. We retained the indicator, because the construct's composite reliability and Cronbach's alpha are above the recommended threshold of 0.70 (Hair, Ringle, & Sarstedt, 2011). All the composite reliability values are again above 0.70, but two constructs have Cronbach's alpha values below the suggested threshold. Similar to the nonweighted model, the poor EXTR and INTR results can be attributed to the measurement scale's lack of length, as it only has two indicators. In terms of convergent validity, the weighted model shows sufficient levels of AVE, with ORGC having the lowest AVE (0.635) and SOCI having the highest AVE (0.784). Similar to the nonweighted model, all three discriminant validity criteria are satisfied in the weighted model (Appendix A, Tables A.1–A.3).

Table 6
Reflective measurement model results.

Model	Construct	Indicator	Indicator loadings (≥0.70)	Indicator reliability (≥0.50)	Cronbach's alpha (≥0.70)	Composite reliability (≥0.70)	AVE (≥0.50)	Discriminant validity (Yes/No)
Nonweighted model	EXTR	EXTR_1	0.802	0.643	0.559	0.818	0.693	Yes
		EXTR_2	0.862	0.743				
	INTR	INTR_1	0.888	0.789	0.487	0.789	0.654	Yes
		INTR_2	0.721	0.520				
	SOCI	SOCI_1	0.873	0.762	0.778	0.898	0.815	Yes
		SOCI_2	0.932	0.869				
	ORGC	ORGC_1	0.837	0.701	0.731	0.848	0.652	Yes
		ORGC_2	0.867	0.752				
		ORGC_3	0.709	0.503				
Weighted model	EXTR	EXTR_1	0.787	0.619	0.555	0.816	0.690	Yes
		EXTR_2	0.872	0.760				
	INTR	INTR_1	0.889	0.790	0.465	0.781	0.644	Yes
		INTR_2	0.705	0.497				
	SOCI	SOCI_1	0.844	0.712	0.732	0.879	0.784	Yes
		SOCI_2	0.925	0.856				
	ORGC	ORGC_1	0.836	0.699	0.710	0.838	0.635	Yes
		ORGC_2	0.860	0.740				
		ORGC_3	0.683	0.466				

Note. EXTR = extrinsic job feature, INTR = intrinsic job feature, SOCI = societal job feature, JSAT = job satisfaction, ORGC = organizational commitment.

Table 7
Structural model results.

Model		Nonweighted model			Weighted model		
		Direct effect	Indirect effect	Total effect	Direct effect	Indirect effect	Total effect
Path	EXTR → JSAT	0.211, $p = 0.000$	n/a	0.211, $p = 0.000$	0.245, $p = 0.000$	n/a	0.245, $p = 0.000$
	INTR → JSAT	0.380, $p = 0.000$	n/a	0.380, $p = 0.000$	0.266, $p = 0.004$	n/a	0.266, $p = 0.004$
	SOCI → JSAT	0.007, $p = 0.884$	n/a	0.007, $p = 0.884$	0.077, $p = 0.303$	n/a	0.077, $p = 0.303$
	EXTR → ORGC	0.080, $p = 0.046$	0.067, $p = 0.000$	0.147, $p = 0.000$	0.083, $p = 0.177$	0.066, $p = 0.000$	0.149, $p = 0.013$
	INTR → ORGC	0.285, $p = 0.000$	0.121, $p = 0.000$	0.406, $p = 0.000$	0.249, $p = 0.000$	0.072, $p = 0.034$	0.320, $p = 0.000$
	SOCI → ORGC	0.120, $p = 0.003$	0.002, $p = 0.884$	0.122, $p = 0.004$	0.140, $p = 0.007$	0.021, $p = 0.355$	0.160, $p = 0.002$
	JSAT → ORGC	0.319, $p = 0.000$	n/a	0.319, $p = 0.000$	0.269, $p = 0.000$	n/a	0.269, $p = 0.000$
R^2	JSAT		0.231			0.181	
	ORGC		0.356			0.275	
f^2	EXTR → JSAT		0.054			0.071	
	INTR → JSAT		0.144			0.069	
	SOCI → JSAT		0.000			0.006	
	EXTR → ORGC		0.009			0.009	
	INTR → ORGC		0.085			0.064	
	SOCI → ORGC		0.018			0.022	
	JSAT → ORGC		0.121			0.082	
Q^2	JSAT		0.219			0.156	
	ORGC		0.233			0.161	

Note. EXTR = extrinsic job feature, INTR = intrinsic job feature, SOCI = societal job feature, JSAT = job satisfaction, ORGC = organizational commitment, n/a = not available.

Having established a reliable and valid measurement model for the nonweighted and weighted models, we continued with the structural model assessment. Although no major difference was found between the nonweighted and weighted models during the measurement model assessment, we found several differences between the two models when assessing the structural model. Table 7 shows the structural results of the model, comprising the nonweighted model and the weighted model's estimations, the coefficient of determination (R^2 values), the effect sizes (f^2), and the predictive relevance (Q^2 values).

Comparing the two models, we find three path coefficients differing more than 0.05. First, in the nonweighted model, the path coefficient for INTR → JSAT is 0.380 and, in the weighted model, the path coefficient is 0.266. The difference is 0.114. Second, the path coefficient for SOCI → JSAT differs by 0.070 between the two models (0.007 vs. 0.077). However, the path coefficients in both models are not significant at the 5% level. Third, the path coefficient of JSAT → ORGC is differing by 0.05 (0.319 vs. 0.269). In addition, we find that the path from EXTR → ORGC is significant in the nonweighted model, but not in the weighted model at a 5% level, although the parameter estimates differ only marginally (0.080 vs. 0.083).

The indirect effect's results show that the indirect paths between the nonweighted and weighted models do not change their significance and differ only marginally. A notable exception is observed in the magnitude of the indirect path INTR → ORGC, whose difference is 0.049 (0.121 vs. 0.072). This difference results in a much larger total INTR → ORGC effect in the nonweighted model (0.406) than in the weighted model (0.320).

Applying the sampling weights to the model also affects the endogenous constructs' explained variance (R^2 values). In the nonweighted model, the R^2 value of JSAT is 0.231, whereas the R^2 value of ORGC is 0.356. In the weighted model, the R^2 values of JSAT and ORGC are 0.181 and 0.275, respectively. These findings indicate that when comparing the two models, the R^2 values of both endogenous constructs decrease by 0.05 or more. This decrease does not mean that the higher explained variance (e.g., better model fit) makes the nonweighted model preferable. Weighting accounts for the unrepresentativeness of the sample, which is due to the unequal probabilities of the sample members being included in the sample when compared to the population. Thereby, weighting provides better population estimates of the relevant coefficients (including the R^2). In consequence, a higher R^2 in the nonweighted data could imply an overfitting of the model compared to the true population model because of the unrepresentativeness.

Although the R^2 values differ between the two models, the f^2 effect sizes do not differ much in magnitude. As shown in Table 7, the f^2 effect sizes of the exogenous latent variables in the nonweighted model and the weighted model have small to medium effects.

The blindfolding procedure allows determining the predictive relevance Q^2 of PLS path models (Hair et al., 2016). For this evaluation criterion, we also find differences of more than 0.05 in the two models. In the nonweighted and weighted models, the Q^2 values of JSAT are respectively 0.219 and 0.156, resulting in a difference of 0.063. Similarly, the difference between the two model's Q^2 values of ORGC is 0.072. That is, in the nonweighted model, the Q^2 value of ORGC is 0.233, whereas in the weighted model the Q^2 value of ORGC is 0.161.

In order to test the mediation analysis, we apply the procedure for PLS path modeling suggested by Hair et al. (2016), which follows the general recommendations for mediation analysis by Zhao, Lynch, and Chen (2010), Preacher and Hayes (2008) using bootstrapping. The results show that one of the three mediating paths in the nonweighted and weighted models differ (Appendix B, Tables B.1 and B.2); that is, JSAT partially mediates the path EXTR → ORGC in the nonweighted model, but fully mediates the same path in the weighted model. By contrast, JSAT partially mediates the relationship INTR → ORGC in both models. The mediation type is a complementary mediation, because both effects point in the same direction and are significant at a 5% level (Hair et al., 2016; Zhao et al. 2010). SOCI, on the contrary, exerts its effect on ORGC directly; hence, we do not find a mediation for this effect in the two models.

Overall, the results show that substantial differences can occur when the sampling weights are not taken into account in the estimation of PLS path model parameters. This neglect can affect the significance, as well as the magnitude of the path coefficient, and affects quality criteria like the explained variance (R^2) and the model's predictive relevance (Q^2). Interpretations of the effects, as in the mediation of

EXTR → ORGC, can be distorted if the sampling inequalities are not appropriately addressed.

The effect that the explained variance (R^2) and predictive relevance of the model (Q^2) of the weighted model decreases compared with the unweighted model should not concern researchers interested in the population parameters. If the weights are correctly computed and the assumptions hold, the weighted estimates should represent the population parameters better than the unweighted (unrepresentative) sample estimates. A better fit with the data in terms of explained variance does not imply that the population model is a good representation.

6. Implications and future research directions

This study proposes a new modified version of the original PLS path modeling approach, namely the WPLS algorithm that incorporates sampling weights. It shows the new approach's appropriateness with an illustrative example and simulated data. The results show that the new modified version takes the specified weights correctly into account. In addition, this algorithm provides better average population model parameter estimates than the basic PLS algorithm when sampling weights are available. In particular, correcting the estimates for deviations in the sampling procedure provides less biased results that are closer to the population parameters. If researchers are interested in inference to the population, they should ensure that they correct the sampling deviations of their data set, as well as ensure that they use sampling procedures that allow them to draw these conclusions.

The empirical examples' results also show the importance of applying sampling weights in model estimations. For example, applying the sampling weights available in the ISSP Work Orientations 2005 to a simple job attitude model shows that drawing conclusions could be misleading when weights are not included in the model estimation. In particular, the results show that although applying the weighting does not alter the measurement model evaluation's results, the structural model results are substantially different in the weighted and unweighted models. Not only was a significant path in the nonweighted model found to be nonsignificant in the weighted model, but the magnitude of the path coefficients may also change substantially. This deviation can have consequences for the theoretical and managerial implications drawn from the analysis.

The new WPLS method is also applicable to other PLS algorithms that the research community developed in recent years. With minor additional corrections, this approach is easily transferable to consistent PLS (Dijkstra & Henseler, 2015; Dijkstra & Schermelleh-Engel, 2014), CTA-PLS (Gudergan, Ringle, Wende, & Will, 2008), blindfolding (Chin, 1998; Wold, 1982), multigroup analysis (MGA; Henseler et al., 2009; Sarstedt, Henseler, & Ringle, 2011), and measurement invariance of composite models (MICOM; Henseler, Ringle, & Sarstedt, 2016). All these algorithms are essentially based on correct PLS path modeling results as input for further calculations. The new method is therefore not limited to the estimation and interpretation of path coefficients, outer model loadings, and outer model weights, but can be applied in various analysis contexts. We showed the use of the weighted blindfolding procedure in the empirical example of job attitudes and the use of weighted consistent PLS in the simulation study.

The use of sampling weights also raises questions for future research. Their usefulness is high if respondents differ in their response pattern, that is, if there are different subgroups in the population. In this case, sampling weights correct deviations in the sampling procedure to yield a better estimate of the average population effect. Nevertheless, the average population effect might not be of high value for all researchers. Differences in response patterns, which can be described as response heterogeneity (i.e., the differences between respondent subgroups), may also be uncovered and the parameters for the all subgroups estimated (Rigdon, Ringle, Sarstedt, & Gudergan, 2011; Sarstedt & Ringle, 2010). Unobserved response heterogeneity can threaten the results' validity, while uncovering and appropriately treating heterogeneity might be a more advisable strategy for theory development (Becker, Rai, Ringle, & Völckner, 2013) than using corrected average population estimates. Future research might elaborate further on these different perspectives on unobserved heterogeneity and their implications for research results.

From the regression context, it is also known that there is a trade-off between the larger bias of unweighted estimators and the larger variability of weighted estimators (Korn & Graubard, 1995). Estimating weighted parameters when weighting is not necessary, inflates the variance of the parameters, and results in inefficient estimators. Hence, researchers should test whether weighting is appropriate or not. Bollen et al. (2016) provide an overview of different tests and their efficacy in assessing the appropriateness of sampling weights in a regression-based context. Future research could also evaluate if these findings hold in a PLS path modeling context.

In addition, in its simulations and illustrative examples, this study focused on parameter bias only. Future research should also try to investigate parameter variability in the PLS path modeling context, where standard errors are not directly estimated but inferred, using bootstrapping to further understand the differences between weighted and unweighted estimators in real empirical applications. If the new WPLS method's statistical power declines, it might be less useful to those researchers who focus less on unbiased parameter estimates and more on the statistical significance of parameters (avoiding type I and type II errors).

The new WPLS method is not limited to situations in which researchers work with sampling weights. The method can also be used for other types of weights. An example could be to estimate group-specific path models with the results emerging from a fuzzy (probabilistic) FIMIX-PLS partitioning (Hahn, Johnson, Herrmann, & Huber, 2002; Sarstedt et al., 2011). The posterior probabilities of FIMIX-PLS segment membership could then be used to estimate the correct PLS path model based on these segmentation results and assess the significance of the parameters by means of bootstrapping. To date, users of these methods have to assign observations to specific segments (discrete partitioning) based on the probability of segment membership and thereafter estimate the group-specific path model. This procedure suffers from a loss of information, because it forces observations in either one of the segments, instead of taking the correct probability of their segment membership into account. Future research should also investigate the usefulness of the WPLS method regarding facilitating such segment-specific estimations based on the probabilities of class membership in higher depth.

Acknowledgements

Jan-Michael Becker has financial interest in the SmartPLS software for modeling and estimating PLS path models. We like to thank the anonymous reviewers, the participants of the 2nd International Symposium on Partial Least Squares Path Modeling — The Conference for PLS Users, and Christian M. Ringle for feedback on prior versions of this manuscript.

Appendix A. Discriminant validity assessment.

Table A.1
Cross-loadings for nonweighted model and weighted model.

Indicator	Nonweighted model					Weighted model				
	EXTR	INTR	JSAT	ORGC	SOCI	EXTR	INTR	JSAT	ORGC	SOCI
EXTR_1	0.802	0.180	0.211	0.222	0.080	0.787	0.116	0.187	0.204	0.066
EXTR_2	0.862	0.228	0.292	0.211	0.079	0.872	0.184	0.301	0.169	0.063
INTR_1	0.236	0.888	0.443	0.444	0.412	0.188	0.889	0.343	0.382	0.422
INTR_2	0.153	0.721	0.229	0.352	0.275	0.092	0.705	0.184	0.276	0.234
SOCI_1	0.068	0.404	0.115	0.256	0.873	0.037	0.418	0.111	0.248	0.844
SOCI_2	0.099	0.388	0.219	0.305	0.932	0.091	0.353	0.239	0.294	0.925
JSAT_1	0.306	0.435	1.000	0.490	0.193	0.300	0.344	1.000	0.409	0.209
ORGC_1	0.212	0.425	0.386	0.837	0.234	0.192	0.344	0.362	0.836	0.227
ORGC_2	0.214	0.439	0.453	0.867	0.335	0.191	0.377	0.323	0.860	0.330
ORGC_3	0.205	0.328	0.339	0.709	0.169	0.140	0.264	0.294	0.683	0.163

Note. EXTR = extrinsic job feature, INTR = intrinsic job feature, SOCI = societal job feature, JSAT = job satisfaction, ORGC = organizational commitment.

Table A.2
Fornell–Larcker criterion for nonweighted model and weighted model.

Latent variable	Nonweighted model					Weighted model				
	EXTR	INTR	JSAT	ORGC	SOCI	EXTR	INTR	JSAT	ORGC	SOCI
EXTR	0.832					0.830				
INTR	0.247	0.809				0.184	0.803			
JSAT	0.306	0.435	1.000			0.300	0.344	1.000		
ORGC	0.259	0.496	0.490	0.807		0.221	0.416	0.409	0.797	
SOCI	0.095	0.435	0.193	0.313	0.903	0.077	0.426	0.209	0.309	0.885

Note. EXTR = extrinsic job feature, INTR = intrinsic job feature, SOCI = societal job feature, JSAT = job satisfaction, ORGC = organizational commitment. The diagonal shows the square root of the AVE and below the diagonal are the construct correlations.

Table A.3
Heterotrait–monotrait ratio (HTMT) for nonweighted model and weighted model.

Latent variable	Nonweighted model				Weighted model			
	EXTR	INTR	JSAT	ORGC	EXTR	INTR	JSAT	ORGC
INTR	0.456				0.335			
JSAT	0.404	0.592			0.394	0.479		
ORGC	0.409	0.817	0.570		0.355	0.706	0.486	
SOCI	0.141	0.692	0.210	0.400	0.114	0.711	0.231	0.414

Note. EXTR = extrinsic job feature, INTR = intrinsic job feature, SOCI = societal job feature, JSAT = job satisfaction, ORGC = organizational commitment. All results satisfy the $HTMT_{.85}$ criterion.

Appendix B. Mediation analysis of nonweighted and weighted models.

Table B.1
Significance analysis of direct and indirect effects of nonweighted model.

Paths	Direct effect	95% CI of the direct effect	*t* value	Sig. ($p < 0.05$)?	Indirect effect	95% CI of the indirect effect	*t* value	Sig. ($p < 0.05$)?	Mediation type
EXTR → ORGC	0.080	[0.003, 0.161]	1.992	Yes	0.067	[0.066, 0.067]	4.198	Yes	Complementary (partial mediation)
INTR → OCRG	0.285	[0.183, 0.379]	5.863	Yes	0.121	[0.120, 0.121]	4.718	Yes	Complementary (partial mediation)
SOCI → ORGC	0.120	[0.043, 0.200]	2.988	Yes	0.002	[0.001, 0.002]	0.138	No	Direct only (no mediation)

Note. EXTR = extrinsic job feature, INTR = intrinsic job feature, SOCI = societal job feature, JSAT = job satisfaction, ORGC = organizational commitment. The confidence interval (CI) was estimated using bias-corrected and accelerated (BCa) bootstrap.

Table B.2
Significance analysis of direct and indirect effects of weighted model.

Paths	Direct effect	95% CI of the direct effect	*t* value	Sig. ($p < 0.05$)?	Indirect effect	95% CI of the indirect effect	*t* value	Sig. ($p < 0.05$)?	Mediation type
EXTR → ORGC	0.083	[−0.042, 0.202]	1.351	No	0.066	[0.065, 0.066]	3.577	Yes	Indirect only (full mediation)
INTR → OCRG	0.249	[0.111, 0.373]	3.702	Yes	0.072	[0.071, 0.072]	2.118	Yes	Complementary (partial mediation)
SOCI → ORGC	0.140	[0.042, 0.245]	2.710	Yes	0.021	[0.020, 0.021]	0.937	No	Direct only (no mediation)

Note. EXTR = extrinsic job feature, INTR = intrinsic job feature, SOCI = societal job feature, JSAT = job satisfaction, ORGC = organizational commitment. The confidence interval (CI) was estimated using bias-corrected and accelerated (BCa) bootstrap.

References

Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Structural Equation Modeling, 12*(3), 411–434.

Becker, J.-M., Rai, A., Ringle, C. M., & Völckner, F. (2013). Discovering unobserved heterogeneity in structural equation models to avert validity threats. *MIS Quarterly, 37*(3), 665–694.

Biemer, P. P., & Christ, S. L. (2008). Weighting survey data. In E. D. de Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 317–341). London: Routledge.

Bollen, K. A., Biemer, P. P., Karr, A. F., Tueller, S., & Berzofsky, M. E. (2016). Are survey weights needed? A review of diagnostic tests in regression analysis. *Annual Review of Statistics and Its Application, 3*, 375–392.

Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (Ed.), *Topics in applied multivariate analysis* (pp. 72–141). Cambridge: Cambridge University.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology, 37*, 1–21.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. New York: Cambridge University Press.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern business research methods* (pp. 295–336). Mahwah, NJ: Lawrence Erlbaum Associates.

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science, 40*, 434–449.

Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly, 39*(2), 297–316.

Dijkstra, T. K., & Schermelleh-Engel, K. (2014). Consistent partial least squares for nonlinear structural equation models. *Psychometrika, 79*(4), 585–604.

DuMouchel, W. H., & Duncan, G. J. (1983). Using sample survey weights in multiple regression analyses of stratified samples. *Journal of the American Statistical Association, 78*(383), 535–543.

Gelman, A. (2007). Struggles with survey weighting and regression modeling. *Statistical Science, 22*(2), 153–164.

GESIS. (2015). *The ISSP*. Retrieved from http://www.gesis.org/en/issp/home/.

Groves, R. M., Dillman, D. A., Eltinge, J. L., & Little, R. J. A. (2002). *Survey nonresponse*. New York: John Wiley & Sons.

Gudergan, S. P., Ringle, C. M., Wende, S., & Will, A. (2008). Confirmatory tetrad analysis in PLS path modeling. *Journal of Business Research, 61*(12), 1238–1249.

Hahn, C., Johnson, M. D., Herrmann, A., & Huber, F. (2002). Capturing customer heterogeneity using a finite mixture PLS approach. *Schmalenbach Business Review, 54*(3), 243–269.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). Sage Publications.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice, 19*(2), 139–151.

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012). 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 Planning, 45*(5), 320–334.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science, 40*(3), 414–433.

Hauff, S., & Richter, N. (2015). Power distance and its moderating role in the relationship between situational job characteristics and job satisfaction. *Cross Cultural Management, 22*(1), 68–89.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2016). Testing measurement invariance of composites using partial least squares. *International Marketing Review, 33*(3), 405–431.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. In R. R. Sinkovics, & P. N. Ghauri (Eds.), *Advances in international marketing* (pp. 277–319). Bingley: Emerald.

Hult, C. (2005). Organizational commitment and person-environment fit in six western countries. *Organization Studies, 26*(2), 249–270.

ISSP Research Group. (2013). *International social survey Programme: Work orientations III ISSP 2005*. http://dx.doi.org/10.4232/1.11648. GESIS Data Archive, Cologne. ZA4350 Data file Version 2.0.0.

Kaplan, D., & Ferguson, A. J. (1999). On the utilization of sample weights in latent variable models. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(4), 305–321.

Korn, E. L., & Graubard, B. I. (1995). Examples of differing weighted and unweighted estimates from a sample survey. *The American Statistician, 49*(3), 291–295.

Little, R. J. (1993). Post-stratification: A modeler's perspective. *Journal of the American Statistical Association, 88*(423), 1001–1012.

Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*. Heidelberg: Physica.

Lu, I. R. R., Kwan, E., Thomas, D. R., & Cedzynski, M. (2011). Two new methods for estimating structural equation models: An illustration and a comparison with two established methods. *International Journal of Research in Marketing, 28*(3), 258–268.

Magee, L., Robb, A. L., & Burbidge, J. B. (1998). On the use of sampling weights when estimating regression models with survey data. *Journal of Econometrics, 84*(2), 251–271.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika, 49*(1), 115–132.

Pfeffermann, D. (1993). The role of sampling weights when modeling survey data. *International Statistical Review/Revue Internationale de Statistique, 61*(2), 317–337.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879–891.

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/.

Richter, N. F., Sinkovics, R. R., Ringle, C. M., & Schlägel, C. (2016). A critical look at the use of SEM in international business research. *International Marketing Review, 33*(3), 376–404.

Rigdon, E. E. (2012). Rethinking partial least squares path modeling: In praise of simple methods. *Long Range Planning, 45*(5/6), 341–358.

Rigdon, E. E., Ringle, C. M., Sarstedt, M., & Gudergan, S. P. (2011). Assessing heterogeneity in customer satisfaction studies: Across industry similarities and within industry differences. *Advances in International Marketing, 22*, 169–194.

Ringle, C. M., Sarstedt, M., & Straub, D. W. (2012). Editor's comments: A critical look at the use of PLS-SEM in MIS quarterly. *MIS Quarterly, 36*(1), iii–xiv.

Ringle, C. M., Wende, S., & Becker, J.-M. (2015). *SmartPLS 3. SmartPLS GmbH, Bönningstedt*. http://www.smartpls.com.

Rungtusanatham, M. J., Choi, T. Y., Hollingworth, D. G., Wu, Z., & Forza, C. (2003). Survey research in operations management: Historical analyses. *Journal of Operations Management, 21*(4), 475–488.

Sarstedt, M., Becker, J.-M., Ringle, C. M., & Schwaiger, M. (2011). Uncovering and treating unobserved heterogeneity with FIMIX-PLS: Which model selection criterion provides an appropriate number of segments? *Schmalenbach Business Review, 63*(1), 34–62.

Sarstedt, M., Henseler, J., & Ringle, C. M. (2011). Multigroup analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results. *Advances in International Marketing, 22*, 195–218.

Sarstedt, M., & Ringle, C. M. (2010). Treating unobserved heterogeneity in PLS path modeling: A comparison of FIMIX-PLS with different data analysis strategies. *Journal of Applied Statistics, 37*(8), 1299–1318.

Shin, H. C. (2012). A cautionary note on post-stratification adjustment. In *Section on survey research methods – Joint statistical meeting (JSM), San Diego*.

Slater, S. F., & Atuahene-Gima, K. (2004). Conducting survey research in strategic management. *Research Methodology in Strategy and Management, 1*, 227–250.

Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment, 80*(1), 99–103.

Tenenhaus, M., Vinzi, V. E., Chatelin, Y. M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics & Data Analysis, 48*(1), 159–205.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). New York: Springer.

Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: How good are single-item measures? *Journal of Applied Psychology, 82*(2), 247–252.

Winship, C., & Radbill, L. (1994). Sampling weights and regression analysis. *Sociological Methods & Research, 23*(2), 230–257.

Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. G. Jöreskog, & H. Wold (Eds.), *Systems under indirect observations: Part I* (pp. 1–54). Amsterdam: North-Holland.

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37*(2), 197–206.