Using partial least squares in archival accounting research: an application to earnings quality measuring

Ana Licerán-Gutiérrez & Manuel Cano-Rodríguez

Published online: 09 May 2019.

Submit your article to this journal

Article views: 275

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Routledge
Taylor & Francis Group

Check for updates

ARTICLE

Using partial least squares in archival accounting research: an application to earnings quality measuring

Ana Licerán-Gutiérrez and Manuel Cano-Rodríguez

Department of Financial Economics and Accounting, University of Jaén, Jaén, Spain

ABSTRACT

Despite the advantages of Structural Equation Modelling (SEM) over regression models that have contributed to its popularisation in several fields of research in social sciences, it has not been broadly applied in archival accounting research. In this paper, we present a guidance for the application of SEM – and, particularly, the Partial Least Squares (PLS) method – to the (arguably) most recurrent topic on empirical archival accounting research: earnings quality. We highlight several problems that arise in earnings quality measuring, indicating how PLS can help to overcome them. We also run a simulation process whose results show that PLS method outperforms the other approaches even in situations of limited information.

Utilización de Mínimos Cuadrados Parciales en la investigación contable de archivo: Aplicación a la medición de la calidad del resultado.

RESUMEN

A pesar de las ventajas de los Modelos de Ecuaciones Estructurales (SEM) respecto a los modelos de regresión que han contribuido a su popularización en diversos campos de invenstigación en Ciencias Sociales, no ha sido ampliamente aplicada en la investigación contable de archivo. En este estudio, presentamos una guía para la aplicación del SEM - y, en particular, el método de Mínimos Cuadrados Parciales (PLS) - al tema (posiblemente) más recurrente en investigación empírica contable de archivo: la calidad del resultado. Destacamos diversos problemas que surgen en la medición de la calidad del resultado, indicando cómo PLS puede ayudar a solventarlos. Asimismo desarrollamos un proceso de simulación cuyos resultados muestran cómo el método PLS supera a otros métodos incluso en situaciones de información limitada.

1. Introduction

Some recent papers argue that empirical archival research in accounting can benefit from the use of more advanced techniques, in particular, structural equation modelling (Gow, Larcker, & Reiss, 2016; Hinson & Utke, 2018; Larcker & Rusticus, 2010; Leuz,

CONTACT Ana Licerán-Gutiérrez ✉ aliceran@ujaen.es ⬤ Department of Financial Economics and Accounting, University of Jaén, Paraje Las Lagunillas, s/n. E-23071, Jaén, Spain

Nanda, & Wysocki, 2003). Following this idea, in this paper, we discuss the advantages of Structural equation modelling (SEM) and, more specifically, Partial least squares (PLS) technique over the traditional first-generation regression methods typically used in archival accounting research.

SEM is a set of statistical techniques used to study the relationships between one or more independent variables and one or more dependent variables, both of which can be directly observed or latent variables[1] (Ullman, 2006). SEM is less restrictive than regression models, as they allow the use of multiple predictors and criterion variables, latent (unobservable) variables, model error in measurement for observed variables, as well as mediation and moderation relationships in a single model (Nitzl, 2016). Despite these advantages make SEM a suitable technique for empirical research in archival accounting (Leuz & Wysocki, 2016), the number of papers that have used this technique is negligible (Hinson & Utke, 2018; Lee, Petter, Fayard, & Robinson, 2011), especially compared to other fields of social sciences, such as psychology, strategic management, management information systems or marketing (Hair, Ringle, & Sarstedt, 2013; Lee et al., 2011; Sarstedt, Ringle, & Hair, 2014), or even other fields of accounting research, such as management accounting and behavioural accounting (Hampton, 2015; Lee et al., 2011).

Blanthorne, Jones-Farmer, and Almer (2006) suggest that the underutilisation of SEM may be attributed to the fact that it is a relatively recent complex technique, and that guidance on conducting research using SEM is distributed among a large number of papers published in multiple and varied sources, being many researchers unaware of the benefits of SEM over traditional methods. In the particular case of empirical accounting research, these reasons may be reinforced by the lack of papers that use SEM as an empirical method. The aim of this paper is to contribute to help archival accounting researchers to enhance their knowledge on SEM by offering a guidance for the application of one of SEM techniques – Partial Least Squares (PLS) method – to the (arguably) most recurrent topic on archival accounting research: earnings quality. In this guidance, we highlight several problems associated to earnings quality measuring, indicating how PLS can overcome them. Additionally, for a better assessment of the advantages of using PLS over the methods traditionally used for measuring earnings quality, we run a simulation process where the performance of PLS and the three more common approaches for measuring earnings quality (single indicator, equally weighted index, and common factor scores) are compared. The results show that PLS typically outperforms the other approaches, even in scenarios with incomplete information.

The contribution of our paper is twofold. Firstly, our paper contributes to increase archival accounting researchers' knowledge about the advantages of SEM in general and PLS in particular, as well as how to adopt this methodology in a rigorous manner. Thus, although there are some previous accounting papers that provide guides on the application of SEM (Lee et al., 2011; Nitzl, 2016), they are mainly focused on management accounting research, being our paper, to our knowledge, the first guide on PLS oriented towards archival financial accounting. Secondly, we illustrate the advantages of PLS method over the traditional regression models using an example that corresponds to the (arguably) most recurrent topic in archival financial accounting: earnings quality measuring. We show how previous papers on earnings quality are based on untested implicit assumptions that cannot be tested using first-generation regression models.

Such assumptions, however, can be tested using SEM. Moreover, we provide empirical evidence of measuring earnings quality using PLS over the traditional measures of earnings quality in our simulation process.

The rest of the paper is structured as follows: Section 2 presents an overview of SEM, comparing its two most common methods (covariance-based and variance-based SEM or PLS). In section 3, we expose a framework for measuring earnings quality, highlighting the problems that arise, analysing how the empirical extant research has dealt with them and the solutions that the PLS method offers. In section 4, we design a simulation process to compare the estimation errors of the PLS approach with the traditional methods employed in empirical research on earnings quality. Section 5 concludes.

2. An overview of structural equation modelling techniques

SEM is a methodology that brings together psychometric and econometric analyses, exploiting the best features of both (Fornell & Larcker, 1981). Thus, the analysis of behaviour and relationships of variables is not restricted to observed variables (econometric analysis) as in regression, but also to non-directly observed (latent) ones (psychometric analysis), which are not directly observed but inferred from directly observable indicators in datasets (Becker, Rai, & Rigdon, 2013; Fornell & Larcker, 1981; Gefen, Rigdon, & Straub, 2011; Henri, 2007; Nitzl, 2016; Rigdon, 2016; Tenenhaus, Vinzi, Chatelin, & Lauro, 2005; Ullman, 2006). Intentions, subjective norms for specific decision-making process, or attitudes are typical examples of these latent variables (Nitzl, 2016). All these things together, the use of latent variables makes SEM suitable for many topics in empirical accounting research that deal with complex, theoretical concepts that cannot be directly observed, as it would be the case of earnings quality.

SEM evaluates in a single and comprehensive analysis two levels of relationships: The measurement model – relationship between the latent variables and their empirical indicators – and the structural model – relationships among the different latent variables (Gefen et al., 2011; Roldán & Sánchez-Franco, 2012; Tenenhaus et al., 2005). Figure 1 shows a graphical example of these two levels.

The main feature of SEM, and what makes it advantageous against other techniques, is the integration of measurement and hypothesised causal paths into a simultaneous assessment (Dijkstra, 2015; Gefen et al., 2011; Hinson & Utke, 2018; Rigdon, 2012; Ullman, 2006). Moreover, it allows the modelling of complex, complete theories with multiple relationships (Davick, 2014; Hair et al., 2013; Ringle, Sarstedt, & Schlittgen, 2014), that can be defined both in the measurement model and the structural model (Henri, 2007; Henseler & Sarstedt, 2013; Hinson & Utke, 2018; Ronkko & Evermann, 2013; Tenenhaus et al., 2005).

There are two main SEM techniques: covariance-based method (CB-SEM), and variance-based method or Partial Least Squares (PLS). These two methods differ in the objective of analysis (see, for example, Gefen et al., 2011; Gefen, Straub, & Boudreau, 2000; Hair, Sarstedt, Ringle, & Gudergan, 2017; Hair, Tomas, Hult, Ringle, & Sarstedt, 2016; Henseler, Ringle, & Sinkovics, 2009), analysis approach (Becker et al., 2013; Davick, 2014; Goodhue, Lewis, & Thompson, 2012; Wold, 1985), statistical supporting (Mateos-Aparicio, 2011; Tenenhaus et al., 2005), consequences in estimation that are caused by model misspecification (Henseler et al., 2014; Henseler, Ringle, &
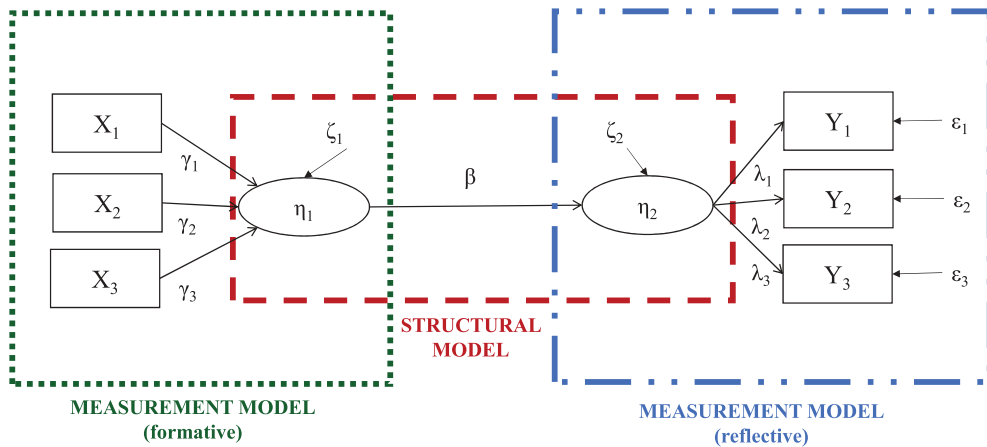
Figure 1. SEM design.

Figure 1 shows a general design of SEM. The circles represent latent variables whereas the rectangles represent observable indicators. The presented model is composed by a structural model and two measurement models. In the structural model, a latent variable (η_1) is related to another latent variable (η_2). In the first measurement model, latent variable η_1 presents a formative relationship with indicators X_1, X_2 and X_3, being γ_1, γ_2, and γ_3 their respective weights and ζ_1 the error term of the latent variable. Latent variable η_2 is measured reflectively using indicators Y_1, Y_2, and Y_3, whose loadings are λ_1, λ_2, and λ_3, and whose estimation errors are ε_1, ε_2, and ε_3 (respectively), existing a measurement error of the latent variable of ζ_2.

Sarstedt, 2012; Reinartz, Haenlein, & Henseler, 2009), type of relationships among variables (Gefen et al., 2011, 2000; Henseler et al., 2012, 2009), and the nature of statistics for goodness of fit (Dijkstra & Henseler, 2015; Fornell & Larcker, 1981; Gefen et al., 2011, 2000; Henseler & Sarstedt, 2013; Rigdon, 2012, 2014). A sum of the differences comparing CB-SEM and PLS is presented in Table 1.

Table 1. Differences between CB-SEM and PLS.

	CB-SEM	PLS
Objective of analysis	Confirm a given theory by explaining the covariance matrix among the items	Maximise the total variance of dependent variables which is explains by the independent ones
Analysis approach	Suitable only for theory confirmation. Limited for predictive analysis.	More suitable for predictive purposes (higher predictive power than CB-SEM) but also suitable for theory confirmation.
Statistical supporting	Strong requirements about theory development, distribution of the variables, sample size or model complexity.	Less restrictive in data requirements. It outperforms CB-SEM in complex models.
Model misspecifications	Not allowed. Misspecification in any subpart of the model implies inability of CB-SEM to estimate.	Allowed. It can correctly estimate even if the model is misspecified in any of its subparts.
Structure of measurement error	It must be exactly known to be modelled.	The exact behaviour and structure of the measurement error need not be known.
Relations between variables	Only reflective relationships.	Both reflective and formative relationships.
Analysis of model fit	Several statistics for the analysis of global goodness of fit.	Lack of a global measure for goodness of fit (alternative tests to evaluate the goodness of the estimated coefficients).

This table presents the differences between the two families of Structural Equation Models: Covariance-Based Structural Equation Modelling (CB-SEM) and Partial Least Squares (PLS) according to the objective of analysis, analysis approach, statistical supporting, problems of estimation derived from model misspecifications, structure of the measurement error, relationships between the variables and analysis of model fit.

Therefore, considering aforementioned differences CB-SEM and PLS are complementary rather than competitive (Gefen et al., 2000; Hair, Ringle, & Sarstedt, 2011; Hair, Sarstedt, Pieper, & Ringle, 2012; Henseler et al., 2012; Sarstedt et al., 2014).

In summary, the selection of CB-SEM or PLS should depend on the suitability of assumptions and objectives of SEM approach with research (Roldán & Sánchez-Franco, 2012). In this paper, we focus on the application of PLS to archival empirical accounting research because of its softer requirements and the possibility of using formative approaches. For a better understanding of the advantages of using PLS, we describe its application to the research on a specific research topic (earnings quality), showing how this technique can solve several problems for the measuring of this concept.

3. A guidance on the use of PLS for measuring earnings quality

In the previous section, we have presented an overview of the advantages of SEM and PLS over traditional regression models. In this section, we discuss how these advantages would benefit the research on the perhaps most common topic in archival accounting research: earnings quality. For doing so, we will review the process for empirical research on Social Sciences, indicating the different steps that researchers should follow to measure earnings quality. For each step, we will compare the typical solutions adopted by the extant literature on earnings quality with the solutions that PLS method provides, showing that this method requires less strong assumptions than the traditional ones.

Despite this vast research on the topic, the term 'earnings quality' is in a current state of 'ambiguity' (Dechow, Ge, & Schrand, 2010) because authors have employed a wide range of empirical measures that are expected to represent earnings quality.[2] This ambiguity can justify the existence of the mixed evidence on the causes and consequences of earnings quality.

3.1. A review of the process for empirical research on social sciences

On broad terms, the goal of theory-based empirical research on Social Sciences is to test the adequacy of a theoretical model to the real world. To achieve this goal, empirical researchers follow a process that covers two different levels (Babbie, 2017; Bisbe, Batista-Foguet, & Chenhall, 2007): the conceptual specification level and the operational level.

Researchers must define the theoretical model in the conceptual specification level. Theoretical models are sets of relationships among different conceptual variables that formalise the key elements of a theory (Bollen, 2002). This level starts with the conceptualisation process, in which the researcher identifies and specifies the exact meaning of the conceptual variables of interest.[3] This process also involves the description of the different aspects of the concept, known as dimensions, as well as the indicators that will be used to measure the concept (Babbie, 2017; Bisbe et al., 2007; Edwards & Bagozzi, 2000). Subsequently, the set of relationships among the conceptual variables that forms the model will be determined (Bisbe et al., 2007).

In the operational level, the researcher develops the specific procedures that will result in empirical observations that represent the conceptual variables in the real world (Babbie, 2017). Finally, by analysing the relationships among the empirical observations

that represent the conceptual variables, the researcher indirectly tests the extent to which the theoretical model is consistent with the real world (Bisbe et al., 2007).

This process can be represented using Libby et al.'s predictive validity framework (Libby, Bloomfield, & Nelson, 2002) shown in Figure 2. At the conceptual level, the researcher defines the conceptual variables (boxes A and B) and develops the model that relates them (link 1). At the operational level, the researcher specifies how the conceptual variables are to be operationalised (boxes C and D, and links 2 and 3). Finally, the relationship between the operationalised variables is assessed (link 4), as a representation of the theoretical relationship between the conceptual variables (link 1).

Next, we will discuss how to apply this process to the research on earnings quality.

3.2. The selection of earnings quality characteristics

As stated before, the conceptual specification level starts with the conceptualisation process, in which researchers specify the exact meaning of the conceptual variables, as well as they describe its dimensions (Babbie, 2017).[4] Regarding earnings quality, and despite the vast literature on the topic, there is no unique, generally accepted definition (Chaney, Cooil, & Jeter, 2008; Hermanns, 2006). Neither the Financial Accounting Standards Board (FASB) nor the International Accounting Standard Board (IASB) gives a formal definition of earnings quality, although they provide a list of qualitative characteristics that are expected to increase the utility of financial information, such as relevance, faithful representation, comparability, verifiability, timeliness, and understandability (IASB, 2010). Consistent with this approach, empirical researchers have typically used different earnings characteristics that are expected to increase the usefulness of accounting information for decision-making, thereby being associated to earnings quality (Dechow et al., 2010; Ewert & Wagenhofer, 2011; Hermanns, 2006; Perotti & Wagenhofer, 2014; Schipper & Vincent, 2003). There is no consensus either, however, on the composition of the list of earnings quality characteristics. Table 2 shows the
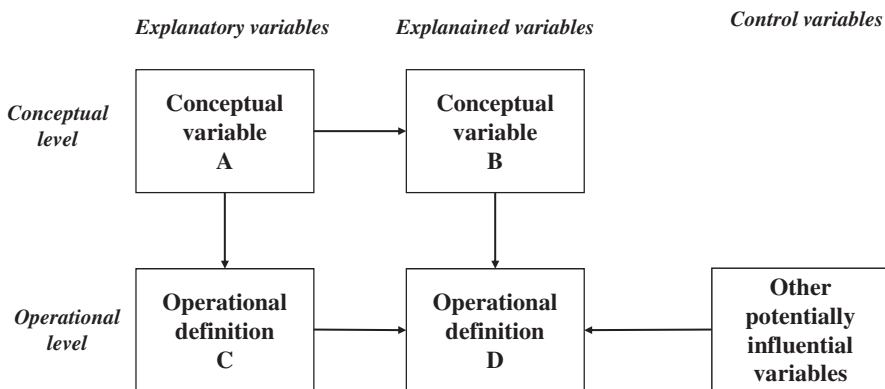
Figure 2. Predictive validity framework (Libby et al., 2002).

This figure represents the process followed to test the adequacy of a theoretical model in reflecting the real world. At the conceptual level, the researcher defines the conceptual variables (boxes A and B) and their relationship (link 1). At the operational level, the conceptual variables are specified (boxes C and D and links 2 and 3). The relationship among the operationalised variables (link 4) is considered a representation of the relationship among the conceptual variables (link 1).

Table 2. The dimensions of earnings quality in prior literature.

Articles	Dimensions of earnings quality
Schipper and Vincent (2003)	Persistence, predictability, variability, abnormal or discretionary accruals, association of accruals and cash flows, comparability, decision usefulness (relevance), estimates, and judgements of experts.
Dechow and Schrand (2004)	Persistence, association of accruals and cash flows, earnings management, conservative accounting, investor response to earnings (ERC), relevance, audit opinion, voluntary disclosure, forecast accuracy.
Francis, LaFond, Olsson, and Schipper (2004)	Accruals quality, smoothness, persistence, predictability, conservatism, timeliness, value relevance
Barth et al. (2008)	Earnings management (smoothing and target beating), timely loss recognition, value relevance
Chaney et al. (2008)	Persistence, predictability, smoothness, accruals variability, conservatism
Laksmana and Yang (2009)	Persistence, predictability, smoothness, accruals quality
Dechow et al. (2010)	Accruals quality, smoothness, persistence, conservatism, investor responsiveness, other indicators of earnings misstatements
Ewert and Wagenhofer (2011)	Persistence, predictability, smoothness, accruals quality, value relevance
Gaio and Raposo (2011)	Accruals quality, smoothness, persistence, predictability, conservatism, timeliness, value relevance
Demerjian, Lewis, Lev, and McVay (2013)	Accruals quality, association of accruals and cash flows, persistence, restatements
Ferrer and Lainez (2013)	Persistence, predictability, variability, smoothness, earnings management, accruals quality, discretionary accruals.
Perotti and Wagenhofer (2014)	Accruals quality, smoothness, persistence, predictability, value relevance
Hermanns (2006)	Persistence, sustainability, predictability, variability, informativeness, association of accruals and cash flows, expertise of auditors

This table presents the different dimensions of earnings quality considered by several papers that account for the multidimensional nature of earnings quality.

diversity of earnings quality characteristics by presenting some papers that have reported a list of the earnings quality dimensions.

Despite the great variety of earnings quality characteristics, the most prevalent method for measuring earnings quality in the previous research has been the selection of just one of those characteristics, represented by a single empirical proxy (Licerán-Gutiérrez & Cano-Rodríguez, 2018).[5] This approach, however, can be valid only if it is assumed that the chosen characteristic is the only one relevant for earnings quality or, alternatively, that the other relevant earnings characteristics remain constant for the specific setting of the research. Given this strong and untested assumption, it is not surprising that the results on earnings quality from this approach have rendered mixed results, showing that some expected causes or consequences of earnings quality are related to some of these proxies, but unrelated to others (Dechow et al., 2010). Additionally, the huge diversity of metrics makes very difficult to interpret the empirical results, because it is unclear whether these indicators are measuring different concepts, different facets of a single concept, or the same facet of a concept (Ewert & Wagenhofer, 2011).

A few papers, on the other hand, have tried to combine different earnings characteristics, computing an index variable built upon the aggregation of the ranking scores of a set of empirical indicators (Bhattacharya, Daouk, & Welker, 2003; Biddle & Hilary, 2006; Boulton, Smart, & Zutter, 2011; Burgstahler, Hail, & Leuz, 2006; Doupnik, 2008; Gaio & Raposo, 2011; Leuz et al., 2003; VanTendeloo & Vanstraelen, 2008). Although those index variables can be considered multidimensional measures of earnings quality, they also present some important limitations. The first one is that the earnings characteristics included in the different indices are selected basing on the subjective judge of the researcher, as no analysis on the validity of those earnings characteristics is

undertaken in the papers that have used these earnings quality index measures. Therefore, this approach is also under the strong assumption that the characteristics included in the index are all relevant and that they are the only relevant ones.

Finally, some papers have developed measures of earnings quality based on the common factor score from a factor analysis of various empirical indicators (Bhattacharya, Ecker, Olsson, & Schipper, 2012; Francis, Nanda, & Olsson, 2008). These measures of earnings quality are based upon the assumption that the different empirical indicators included in the analysis are manifestations of the same construct (earnings quality). These papers, then, do not consider earnings quality dimensions, given that they link all the empirical indicators to one single construct, which is earnings quality. This assumption, however, contrasts with Dechow et al. (2010), who analysed the convergent validity of several earnings quality metrics, concluding that they are not representing the same construct, but measuring different dimensions.

In summary, for measuring earnings quality, researchers must specify the set of earnings characteristics that form the earnings quality construct. In the previous literature, researchers selected these characteristics basing on their subjective judgment, without testing if the selected characteristics were or not relevant for the earnings quality construct. The results of these previous studies on earnings quality, therefore, depend heavily on these untested assumptions.

The use of SEM may help researchers to avoid the subjectivity in the selection of earnings characteristics, as this technique evaluates the measurement model. Thus, the researcher can initially define the set of earnings characteristics that are expected to form earnings quality. The application of the SEM technique will inform the researcher about if the chosen characteristics are or not relevant for defining the earnings quality construct.

3.3. The relationships between earnings quality and its dimensions

Conceptualisation process also requires to determine the nature of the relationships between the construct (earnings quality in our case) and their dimensions (earnings characteristics) (Edwards & Bagozzi, 2000). There are two broad ways in which a construct is related to its indicators or dimensions: reflective and formative measurement (Sarstedt, Hair, Ringle, Thiele, & Gudergan, 2016).

In reflective measurement, indicators are considered as error-prone manifestations of the construct, in the sense that the presence or absence of that construct produces variations in the value of the indicators (Edwards, 2001), being the variations in the indicators' considered *consequences* of the construct. Beliefs, intentions, opinions, perceptions or judgements are typical examples of constructs reflectively related to their indicators (Rodgers & Guiral, 2011). In formative measurement, indicators are seen as the constitutive facets of the construct and, therefore, the indicators as a group jointly determine the conceptual meaning of the construct (Bisbe et al., 2007). The construct is then modelled as a combination of the indicators (Bollen & Bauldry, 2011). Concepts such as liquidity, leverage or profitability, or new non-financial metrics can be represented by formative constructs, combining several pieces of accounting information into a single construct (Rodgers & Guiral, 2011).

The difference between reflective and formative measurement presents important implications for the operationalisation process. Thus, in the reflective measurement, all the indicators of the same construct are expected to be highly correlated (Edwards & Bagozzi, 2000). Additionally, they are interchangeable, and removing a specific indicator would not alter the conceptual domain of the construct (Bisbe et al., 2007; Jarvis, MacKenzie, & Podsakoff, 2003).[6] Consequently, the researcher does not need to use all the available indicators, but a sample of indicators would be enough for measuring the concept as far as the convergent and discriminant validity tests support its consistency.

The indicators in a formative model are not necessarily highly correlated, though, because they do not share the same causes.[7] More importantly, they are not interchangeable: The omission of an indicator would imply that one of the constitutive facets of the construct is left out, thereby changing the definition of the construct (Jarvis et al., 2003). Therefore, a full census of indicators is required (Bisbe et al., 2007). Besides, the validity of the construct must be assessed using nomological and/or criterion-related validity, as internal consistency reliability is an inappropriate criterion for formative measurement models (Jarvis et al., 2003).

The determination of the reflective or formative nature of the relationships between the conceptual variable and its indicators or dimensions is a key feature of the conceptualisation, because the misspecification of these relationships may have serious consequences for the drawn conclusions. Various papers have demonstrated that the use of a reflective (formative) measurement model to a truly formative (reflective) construct lead to inaccurate conclusions about the structural relationships between constructs (Chang, Franke, & Lee, 2016; Jarvis et al., 2003; MacKenzie, Podsakoff, & Jarvis, 2005; Rodgers & Guiral, 2011). In order to help researchers to select the proper model, Jarvis et al. (2003) compile a set of decision rules for determining whether a construct is formative or reflective. Table 3 reports these rules.

The question about the type of relationship between the earnings quality and its empirical proxies has been scarcely addressed by previous research. In order to discuss the nature of the relationship between earnings quality and the empirical measures, we apply Jarvis et al.'s (2003) decision rules to Dechow et al. (2010) categorisation of earnings quality proxies into three broad groups: earnings properties, investor's responsiveness to earnings, and external indicators of misstatements.

Earnings properties are those earnings characteristics that are expected to increase the usefulness of reported earnings in the decision-making process. Accruals quality, earnings smoothing, earnings predictability or conservatism are typical examples of such characteristics. A common feature of earnings properties is that they are jointly determined by the fundamental performance of the company, the ability of the accounting system to measure such performance, and the manager's decisions on the accounting system (Dechow et al., 2010). In other words, managers make accounting choices that affect these properties, thereby increasing or decreasing earnings quality level. Applying Jarvis et al.'s conditions, it can be concluded that; (1) The direction of the relationship is from the earnings properties to the earnings quality construct (changes in the properties cause changes in earnings quality); (2) earnings properties are not substitute nor interchangeable, and they are not highly correlated (Dechow et al., 2010); (3) the different properties do not share the same nomological net nor have the same antecedents and/or consequences. Consequently, according to Jarvis et al.'s

Table 3. **Decision rules for differentiating between formative and reflective models** (Jarvis et al., 2003).

	Formative model	Reflective model
Criterion 1. Direction of causality	• From items to construct • Indicators define characteristics of the construct • Changes in the indicators produce changes in the construct • Changes in the construct should not produce changes in the indicators	• From construct to item • Indicators are manifestations of the construct • Changes in the indicators should not produce changes in the construct • Changes in the construct produce changes in the indicators
Criterion 2. Interchangeability of indicators	• Indicators are not interchangeable • Indicators need not have the same or similar content or share a common theme • Dropping one indicator alters the conceptual domain of the construct	• Indicators are interchangeable • Indicators should have the same or similar content and share a common theme • Dropping one indicator does not affect the conceptual domain of the construct
Criterion 3. Covariation among the indicators	• It is not necessary for indicators to covary with each other • Changes in one indicator are not necessarily associated with changes in the other indicators	• Indicators are expected to covary with each other • Changes in one indicator are associated with changes in the other indicators
Criterion 4. Nomological network of construct indicators	• Nomological network may differ across indicators • Indicators are not required to have the same antecedents and consequences	• Same nomological network for all the indicators • Indicators are required to have the same antecedents and consequences

This table summarises the rules proposed by Jarvis et al. (2003) to determine if the relationship between a construct and its indicators is formative or reflective.

conditions, earnings properties can be expected to be related to earnings quality in a formative way.

The other two categories of earnings quality proxies defined by Dechow et al. (2010) are the investor's responsiveness to earnings and the external indicators of misstatements. Investor's responsiveness to earnings is the measures that analyse the influence of earnings on equity investors' decisions, typically by analysing the relationship between accounting earnings and market returns.[8] The underlying assumption is that higher quality earnings are more relevant for equity investors', thereby increasing the relationship between reported earnings and the results of investors' decisions (market returns). Regarding the third category, the external indicators of earnings misstatements are indicators of the existence of problems with the quality of earnings issued by an external party, such as SEC Accounting and Auditing Enforces Releases, accounting restatements, or reported internal control procedure deficiencies (Dechow et al., 2010).

The relationship between the earnings quality proxies of these two categories and earnings quality can be considered reflective, as these proxies are the result of the observation of the earnings quality level by an external party (investors, SEC, or auditors). Then, it is the construct (earnings quality) which causes the indicator (empirical proxies of investors' reactions or of external indicators of misstatements). Besides, these measures represent knowledge from auditors, managers or SEC, which can be better captured using reflective relationships (Rodgers & Guiral, 2011).

It is unclear, however, that investor responsiveness measures and external indicators represent the same construct, as both categories are not likely to have the same

nomological net nor the same antecedents and consequences. The reason would be that the two groups correspond to the reaction to earnings quality of different groups of users of the accounting information (investors, SEC, the management team, and auditors) and, earnings quality depends on the specific group of users or the specific decision-making setting (Dechow et al., 2010).

In conclusion, researchers have two alternative ways of measuring earnings quality. On the one hand, they can use a formative model in which several earnings properties are combined to define the earnings quality construct. On the other hand, they can use a reflective model using indicators that reflect the existence or absence of earnings quality for a given decision-making setting. Both approaches have their pros and cons. If the formative model is followed, it will be necessary to define and measure all the earnings properties that configure earnings quality, as well as to estimate the weights of each property in the earnings quality construct. The main advantage of this method would be that it can be applied for different decision-making settings, simply by estimating the specific weights of each characteristic for that specific setting. The reflective measurement model, on the other hand, presents the advantage that a sample of indicators of earnings quality would be enough if the convergent and discriminant validity tests support its consistency. Its main con is that, as the definition of earnings quality varies with the decision-making setting, only indicators for that specific setting should be included.

Despite the importance of the correct specification of the nature of the relationships between earnings quality and its characteristics, to our knowledge, previous research does not address this issue. The main reason could be that most of the papers on earnings quality relay on a single characteristic model, for which the specification of the relationship between earnings quality and its characteristic would be irrelevant. Additionally, papers that use aggregated indices assume implicitly that the relationship between the earnings quality construct and the different characteristics included in the index is formative. This implied assumptions, however, may lead to combine in the same index earnings characteristics that, according to our former analysis, are related in a formative way to earnings quality (earnings properties) with characteristics whose relationship would be reflective (as investors' reactions to earnings).[9] On the other hand, those papers that have developed a composite measure based on common factor scores (Bhattacharya et al., 2012; Francis et al., 2008) assume implicitly the existence of a reflective relationship between earnings quality and its metrics.

In summary, the nature of the relationship between earnings quality and its metrics has not been addressed by previous research,[10] despite the potential bias effects of the mis-specification of these relationships. Following Jarvis et al.'s (2003) decision rules, we have concluded that accounting earnings properties can be considered as formatively related to earnings quality, whereas investors' reaction to earnings or other external indicators of earnings quality are more likely to be reflectively related to earnings quality.

For the rest of the paper, we will focus on the formative measurement model of earnings quality (that is, based on earnings properties) because of two reasons: First, because as we have indicated, the reflective methods (using investor' reactions to earnings quality or external indicators of misstatements) would render an earnings quality measurement for a specific group of users of the financial information (investors in the first case; SEC, managers or auditors in the second), whereas using the formative model we can measure earnings quality for any group of users or decision-making setting simply by changing the

weights of the earnings properties. Second, because the great majority of studies on earnings quality have used accounting properties as earnings quality measures (Licerán-Gutiérrez & Cano-Rodríguez, 2018). Additionally, we will consider the variance-based SEM (or PLS) as the technique used to measure earnings quality, as this technique can be applied to both reflective and formative-related variables.

3.4. The estimation of the weights of each dimension

After defining the list of earnings properties, the researcher needs to estimate how those characteristics are combined to configure the earnings quality construct. Previous papers that have tried to measure earnings quality from a formative approach (aggregating the ranking scores of the empirical indicators) have assigned equal weights for all the measures included in the index. This assumption, however, is not likely to be met in practice: As the combination of the earnings quality characteristics that configures the earnings quality construct depends on the trade-offs among the costs and benefits of all the properties for the decision-maker (DeFond, 2010), it can be expected that not all dimensions will be equally relevant. Moreover, this combination will be different for each user of the financial information, as the benefits and costs of a given characteristic will not be the same for all the groups of users of financial information, making the concept of earnings quality dependent on the decision setting (Dechow et al., 2010). One of the advantages of using PLS for measuring earnings quality is that the researcher does not assume any value for the weights of the earnings quality dimensions, as these weights are estimated empirically.

More specifically, the PLS method estimates the weights for each earnings property that minimises the residual variance of the predictive relationships in each latent variable (Mateos-Aparicio, 2011), thereby increasing the weights of those indicators that are more reliable to explain the latent variable (Hair et al., 2012; Henseler et al., 2014; Ronkko & Evermann, 2013). The estimation of these weights allows the researcher to discriminate between relevant (high weights) and irrelevant (low weights) earnings properties for the decision-making setting of the study.

Additionally, PLS is more flexible than traditional regression models with regards to the relationships among the different latent variables, allowing the differentiation between direct, indirect and mediated relationships.

In addition to the estimation of the weights, researchers should check that the included characteristics are really different dimensions. If we consider the dimensions reported in Table 1, it is not clear that they are all different characteristics, as several of them may be just different ways of assessing the same characteristic. For instance, discretionary accruals, accruals quality, the association of accruals and cash flows, or accruals variability are all referred to the quality of the accruals component of earnings. These variables may be then not different properties of earnings quality, but simply different measures of a single property that would be the quality of the accruals portion of earnings. Similarly, persistence, earnings predictability, earnings variability, and earnings smoothing are also closely related,[11] and they could be different ways of assessing the information content of current earnings about future earnings rather than different earnings characteristics. The use of PLS techniques allows to test the discriminant validity of the different dimensions, that is to say, if some apparently

related earnings properties are really different (albeit related) properties or if they are just different measures of one single property.

This discriminant validity checks if a construct is unique and captures phenomena of interest that is not represented by other different construct in the model (Hair et al., 2016; Henseler, Ringle, & Sarstedt, 2015). To test the discriminant validity, the researcher can use the Fornell-Larcker (1981) criterion[12] or the heterotrait-monotrait ratio (HTMT) (Henseler et al., 2015).[13]

3.5. Determination of the empirical proxies

After completing the conceptual specification level, the next step would be to determine how to measure the conceptual variables in the real world (Babbie, 2017). That is, the researcher needs to specify the observed variables that will serve as indicators of the different constructs defined in the conceptual specification level, and how those indicators will be related to the construct. For the specific case of earnings quality, researchers should specify the different proxies that will represent each of the earnings characteristics that are related to earnings quality, and how those proxies will be related to the correspondent characteristic.

Regarding the form of the relationship (reflective or formative) between the empirical indicators and the earnings characteristics, we see the indicators as manifestations of the construct they represent, being therefore the relationship between the constructs and the indicators reflective. Thus, for example, accruals quality is typically measured using discretionary accruals. The underlying idea is that earnings manipulation makes total accruals to deviate from the expected value, thereby generating discretionary accruals. The measure (discretionary accruals) is then caused by the construct (manipulation). Another example would be earnings smoothing proxies: by smoothing earnings, managers reduce the variability of the earnings more than the variability of the comparing variable (sales or cash-flows); additionally, if managers use accruals for smoothing earnings, they will make the correlation between accruals and cash-flows more negative. Similarly, a high value of ERC would be the consequence of a high investors' responsiveness to earnings. In all these three examples, the indicator (variability of earnings compared to sales or cash-flows, correlation between accruals and cash-flows, and ERC) is determined by the latent variable (income smoothing or high investors' responsiveness to earnings).

In summary, we will consider that the relationship between the empirical indicators and their associated earnings properties is reflective, because those indicators are usually caused by the properties.

Apart from defining the relationship between each construct and its empirical indicators, researchers will also have to choose which indicators she/he will use for measuring the different earnings characteristics. To this respect, empirical research presents a wide range of different empirical proxies for each of the above-mentioned earnings quality characteristics, what makes unclear which one of those empirical proxies provides a more accurate measuring of the represented characteristic. A clear example of this problem is the case of accruals quality. Accruals quality has been measured using different approaches, such as the total value (or the absolute value) of accruals; the variability of accruals; the standard deviation of the errors from a residual regression of working capital accruals on lagged, current, and forwarded cash-flows (Dechow & Dichev, 2002); or the residuals (or the absolute value of the residuals) from

a model for estimating non-discretionary accruals. All these measures, however, have been heavily criticised in the literature (Christodoulou, Ma, & Vasnev, 2018; Dechow, Sloan, & Sweeney, 1995; Jackson, 2018; Jones, Krishnan, & Melendrez, 2008; McNichols & Stubben, 2018), so it can be argued that these empirical indicators are measuring their corresponding earnings characteristic with error.

Previous literature has not addressed the potential measurement error of the different earnings quality proxies, apart from those papers that have used several different proxies as robustness checks. PLS method, on the other side, has the advantage that the researcher may incorporate as many indicators as needed for representing each latent variable (Gefen et al., 2011; Lee et al., 2011; Reinartz et al., 2009). Moreover, it is convenient to provide a high number of indicators for each property, because the consistency of PLS increases with the number of indicators (Gefen et al., 2011; Reinartz et al., 2009; Ringle et al., 2014; Wold, 1980).

Additionally, PLS method also allows testing the extent to which a given empirical indicator is representing accurately the underlying construct by assessing the individual reliability of the indicator. This analysis is made by checking if most of the variance of the indicator is explained by its associated latent variable (Fornell & Larcker, 1981; Hair et al., 2016; Henseler et al., 2009; Mackenzie, Podsakoff, & Podakoff, 2011; Roldán & Sánchez-Franco, 2012). PLS analyses the strength of such association observing the indicator loadings as an absolute correlation between the construct and each indicator (Hair et al., 2011, 2012, 2016; Henseler et al., 2009).[14]

Apart from testing the individual reliability of each indicator, the researcher needs to test if the indicators that are expected to represent one given earnings characteristic are really representing the same unobservable variable. For testing this, PLS method allows different tests: the internal consistency of the construct and the convergent validity of the construct. The internal consistency of the construct is assessed by testing how the different indicators vary together (Gefen et al., 2000; Mackenzie et al., 2011; Ringle et al., 2014; Roldán & Sánchez-Franco, 2012). To assess the internal consistency reliability, researchers can use the composite reliability index by Fornell and Larcker (1981) criterion (Hair et al., 2012; Nitzl, 2016). This index provides an estimate of the reliability of a construct that is based on the intercorrelations of the observed indicator variables (Henseler et al., 2009; Mackenzie et al., 2011) according to their loadings (Hair et al., 2011, 2012, 2016; Henseler et al., 2009; Nitzl, 2016; Roldán & Sánchez-Franco, 2012).[15]

The convergent validity test analyses the average proportion of variance of the set of indicators that the latent variable is able to explain. Then, if the empirical proxies are really manifestations of their associated earnings property, a high proportion of the variance of those empirical proxies will be explained by the earnings property. The measure of this average proportion of variance is called Average Variance Extracted (AVE) and was developed by Fornell and Larcker (1981).[16]

Finally, another problem with the selection of empirical indicators is that, in some cases, the same indicator may be associated to different earnings characteristics. For instance, earnings variability can be an indicator of the absence of earnings smoothing (Barth, Landsman, & Lang, 2008) or an indicator of low earnings predictability (Dichev & Tang, 2009). Similarly, the existence of a low negative correlation between accruals and cash-flows can indicate the absence of earnings smoothing, but also a higher level of conditional conservatism (Ball & Shivakumar, 2005). Another example would be the accumulation of

negative accruals (Givoly & Hayn, 2000) or the existence of hidden reserves (Penman & Zhang, 2002), which have been used both as indicators of unconditional conservatism and as indicators of income-decreasing manipulation. To this respect, we have previously indicated that PLS allows to test the discriminant validity of the different constructs, which can be useful for checking if two apparently related constructs are really two different constructs or they are just the same one. Additionally, using the discriminant validity analysis, the researcher can assess if a specific indicator that was expected to be associated to one specific earnings property is a better indicator of another different earnings property, thereby showing which property the indicator represents more accurately. This can be assessed by analysing the cross-loadings from the indicators. The criterion, suggested by Barclay, Higgins, and Thompson (1995) and Chin (1998), states that each indicator must load more highly on their own construct than on any other construct and, consequently, that all constructs must share more variance with their indicators than with any other constructs (Chin, 2010; Hair et al., 2011; Henseler et al., 2015, 2009; Roldán & Sánchez-Franco, 2012).

In summary, researchers may test the quality of their measure model by using PLS technique, thereby assessing if the chosen empirical indicators are truly representing accurately their associated dimension, if the set of indicators that are intended to represent one specific dimension are really representing one and the same dimension, and to check which dimension is better represented by some ambiguous indicators. In Table 4, we present a sum-up of these tests and the threshold values for each one to be considered as acceptable.

3.6. A proposal of measurement model for earnings quality using PLS

In the previous sections, we have shown how the PLS method can overcome several problems in the measuring of earnings quality that, in the previous research, have been overcome relying on untested assumptions. A summary of these problems and how the PLS and the traditional methods deal with them are presented in Table 5.

Table 4. Sum-up of the measurement validation tests in PLS for proxy selection in earnings quality measurement.

Questions to be solved	Tests in PLS to answer the questions	Criterion for validation	Rules of thumb for acceptable values
Are each of the proxies measuring accurately the specific aspect they aim to measure?	Individual indicator reliability	Magnitude of the loadings	>0.7 It can be lower but always >0.4
What is the theoretical concept reflected by the proxies? Are all they measuring the same concept?	Internal consistency reliability	Composite reliability	>0.7
	Convergent validity	Average Variance Extracted (AVE)	>0.5
Which earnings characteristic is most associated to the empirical proxy?	Discriminant validity	Cross loadings	Loading in its construct > Loading in the rest of constructs

This table summarises the main problems highlighted earnings quality literature regarding its measurement and the consequent questions to be answered for an improvement of earnings quality measurement. In the table, it is stated how PLS is able to solve each of these questions with the tests from measurement model valuation. The table is divided into five columns: (1) Problems from earnings quality proxies according to prior literature. (2) Questions (concerns) to be solved for an improvement in earnings quality measurement. (3) Tests in each of the steps of measurement model valuation in PLS that respond to the questions. (4) Criteria used in each of the tests in PLS to assess the validity of the analysed issue. (5) Rules of thumb for the acceptance of validity in each of the criteria.

Table 5. Problems in earnings quality measuring.

	How addressed by			
	Individual indicator	Equally weighted index	Factor index	PLS
(1) What characteristics define earnings quality?	It is assumed that there is only one relevant characteristic, which is that measured by the individual indicator	It is assumed that the characteristics included in the index are the only relevant earnings characteristics	There are no different characteristics (reflective relationship)	The author defines an initial set of relevant characteristics. PLS method evaluates if they are or not relevant
(2) What is the type of relationship (reflective or formative) between earnings quality and its characteristics?	It is assumed that there is only one relevant characteristic	Assumed formative	Assumed reflective	The author defines theoretically the expected type of relationship. PLS method allows testing both formative and reflective relationships.
(3) Do the different characteristics have a direct influence on earnings quality, or do they only affect other earnings quality characteristics?	It is assumed that there is only one relevant characteristic	It is not tested	There are no different characteristics (reflective relationship)	PLS method allows testing indirect and mediating effects
(4) What is the relative importance of each characteristic in the definition of earnings quality for each decision making setting?	It is assumed that there is only one relevant characteristic	It is assumed (untested) that all the characteristics have similar importance (equally-weighted)	There are no different characteristics (reflective relationship)	PLS method estimates the weights for each characteristic
(5) How accurate is the indicator in representing the earnings characteristic to be measured?	It is assumed (not tested) that the indicator reflects accurately the earnings characteristic	It is assumed (not tested) that the indicators reflect accurately the earnings characteristics	It is assumed (not tested) that the indicators reflect accurately earnings quality	PLS method tests indicator individual reliability
(6) Are the empirical indicators measuring different earnings properties or are they different ways of measuring a single characteristic?	It is assumed that there is only one relevant characteristic	It is implicitly assumed (untested) that they represent different characteristics	It is implicitly assumed (untested) that they represent the same characteristic	PLS method tests if they represent the same or different characteristics (internal consistency, convergent validity, discriminant validity)
(7) Which specific characteristic is exactly representing the empirical indicator?	It is assumed (not tested) that the indicator is measuring its associated characteristic and not other different characteristic	It is assumed (not tested) that the indicators are measuring their associated characteristics and not other different characteristics	It is assumed (not tested) that the indicators are measuring earnings quality	PLS method tests which characteristic is represented more accurately by the indicator

This table summarises the problems that arise in earnings quality measuring and how the different approaches for earnings quality measurement address then. Problems are organised in rows whereas the approaches are organised in columns.

In this section, we propose a model for measuring earnings quality using PLS. This model is depicted in Figure 3. In the figure, we represent a structural model where the latent variable *Earnings Quality* is the explanatory variable of a dependent variable. We consider three alternative measurement models for the *Earnings Quality* construct. First, it can be estimated from a set of *Earnings Properties* that are related to *Earnings Quality* in a formative way (blue arrows). These *Earnings Properties* are also latent variables, so they have to be estimated from a set of empirical indicators that are related in a reflective way to each one of those properties. Alternatively, *Earnings Quality* can also be estimated from a set of indicators that represent the investors' reaction to earnings quality (only if the research setting is related to investors' decision-making) or, alternatively, from a set of external indicators of misstatements (only if the research setting is related to such external parties). In these two cases, the expected relationship between *Earnings Quality* and these indicators is reflective.

Next, we will base on this proposal for comparing the performance of PLS and the traditional methods used to measure earnings quality with a simulation process.

4. A comparison of the performance of the methods employed to measure earnings quality

In this section, we conduct a simulation process to compare the estimates of the different approaches that have been previously used to measure earnings quality with
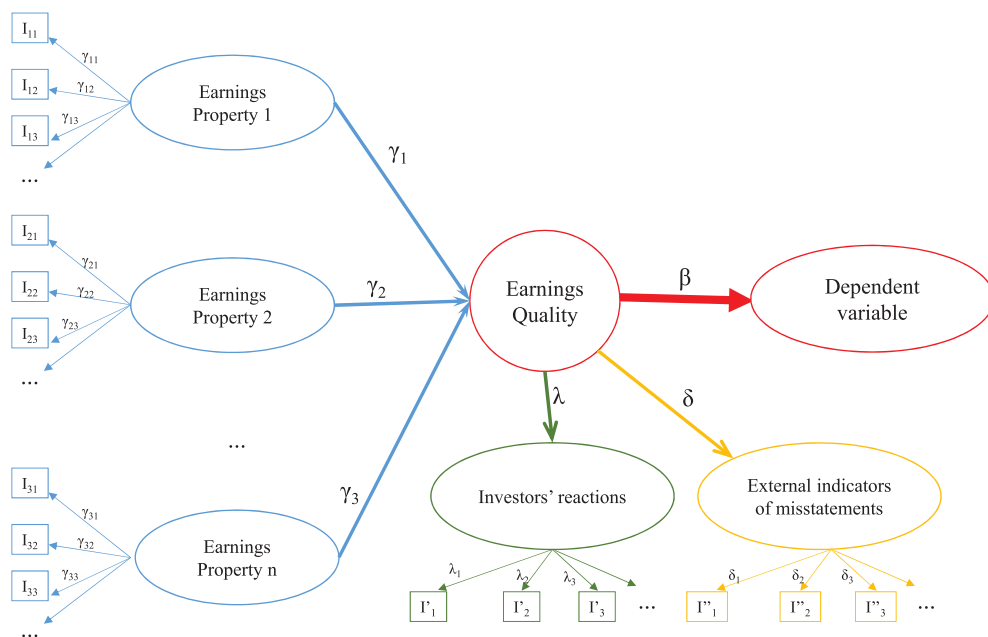

Figure 3. Earnings quality measure model.

This figure depicts the theoretical framework for estimating the influence of the *Earnings Quality* construct on a dependent variable. Circles and ovals represent constructs; squares represent indicators. *Earnings Quality* can be estimated either from a set of earnings properties using a formative model (blue arrows), from the investor's reactions indicators using a reflective model (green arrows), or from the external indicators of misstatements using a reflective model too (yellow arrows).

the PLS method. For doing so, we estimate the influence of a non-observable earnings quality construct (noted by EQ) on a dependent variable (noted as $DEPENDENT$). Earnings quality construct EQ is formed by five dimensions (from EQ_1 to EQ_5), existing a formative relationship between the construct and the dimensions. We consider five available indicators for estimating each dimension (we noted them as eq_{ij}, ranging both i and j from 1 to 5). Figure 4 represents the simulated model.

4.1. Description of the simulation process

The process for the simulation is as follows:

We first simulate a correlation matrix for the five earnings characteristics (EQ_1 to EQ_5). The values of the correlations between two different characteristics are drawn
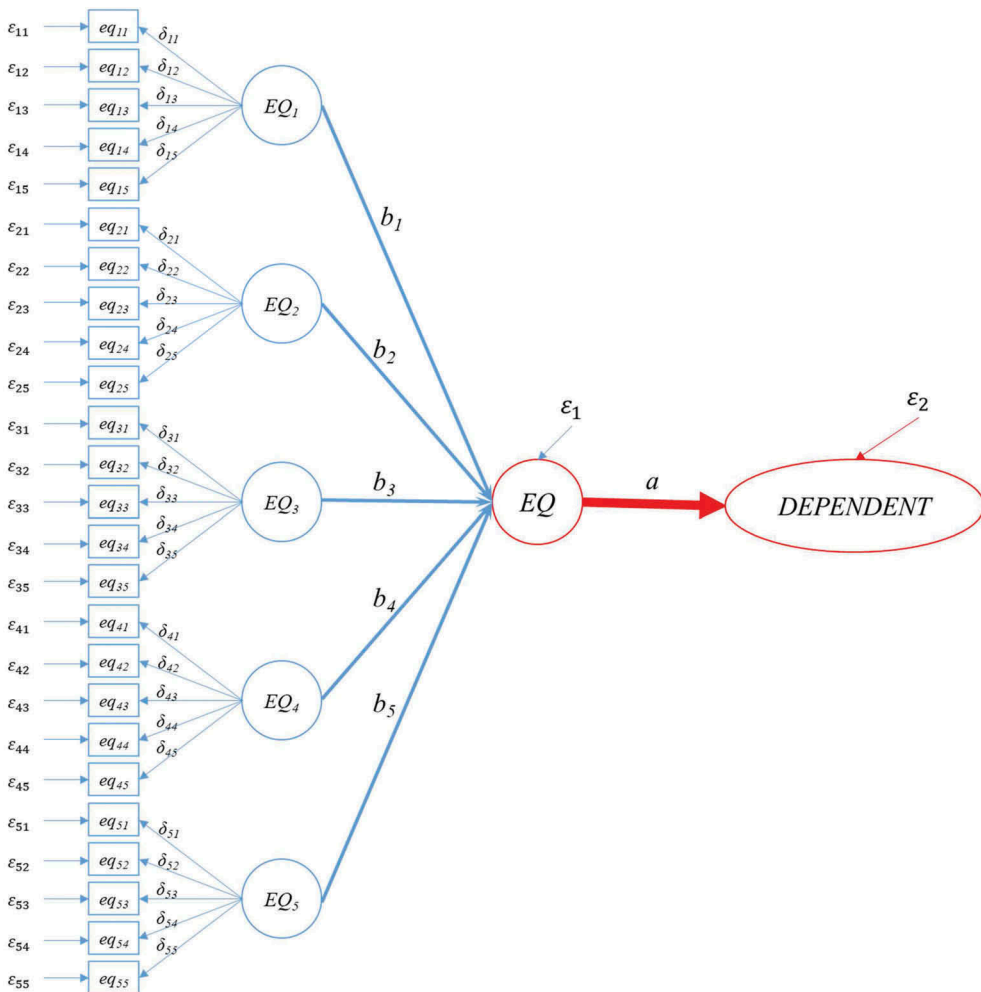

Figure 4. Earnings quality model for the simulation process.

This figure depicts the simulated earnings quality model. Round shapes represent constructs, whereas square shapes represent indicators.

from a uniform distribution with values between – 0.25 and +0.25,[17] ensuring that the resulting matrix is semi-positive definite. Then, we generate the values for the five earnings characteristics from a multivariate normal distribution with 0 mean, standard deviation equal to 1, and using the correlations of the former matrix.

Next, we computed the value of the earnings quality construct according to the following equation:

$$EQ = \sum\nolimits_{i=1}^{5} b_i \cdot EQ_i + \varepsilon_1$$

EQ represents the earnings quality construct, being EQ_i the values generated for the five dimensions that form the construct. b_i are the weights of each dimension. As we assume that these weights are unknown for the researcher, we generated five random values from a uniform distribution between 0 and 1, with the restriction that the sum of the five parameters b_i is equal to 1. The error term of the equation (ε_1) is generated from a normal distribution with zero mean, and uncorrelated with all the other random variables. As the standard deviation of this error is not observable, we computed it as a proportion (generated from a uniform random variable between 0.1 and 0.4) of the standard deviation of the construct. The error is then generated from a normal variable with zero mean and the resulting standard deviation. After computing the values of EQ, we standardised this variable to get a variable with null mean and standard deviation equal to 1.

Next, we computed the values of the dependent variable according to the next equation:

$$DEPENDENT = a \cdot EQ_{st} + \varepsilon_2$$

$DEPENDENT$ is the dependent variable; a is the coefficient of the linear relationship between the dependent variable and the standardised values of the earnings quality construct (EQ_{st}); the error term (ε_2) is simulated from another normal standard variable with 0 mean and independent from any other random variable, whose standard deviation is computed for making the standard deviation of $DEPENDENT$ equal to 1. Additionally, we set the value of parameter a to 0.5.[18]

Then, for each earnings quality dimension (EQ_i), we simulated five indicator variables (from eq_{i1} to eq_{i5}) according to the next equation:

$$eq_{ij} = \delta_{it} \cdot EQ_i + \varepsilon_{ij}$$

where eq_{ij} represents each indicator; EQ_i is the earnings quality dimension represented by that indicator; parameter δ_{it} represents the relationship between the indicator and the component; and ε_{ij} is the error term. We assume that the researcher does not know the exact relationship between the indicators and the components (δ_{it}), so the values of δ_{it} were randomly generated from a uniform distribution between 0.5 and 1. The error term is generated from a normal distribution with zero mean and uncorrelated with all the other variables, and with a standard deviation computed for making the standard deviation of the indicator equal to 1.

We use this simulation process to compare the following four approaches in estimating earnings quality:

(1) Single proxy: in this approach, we picked just one of the indicators (specifically, eq_{11}) as the first earnings quality proxy.
(2) Equally weighted index: computed as the average decile ranking of the chosen indicators for each observation. To avoid the potential bias effect of the metrics of the latent variables (Aguirre-Urreta and Marakas, 2012; Chang et al., 2016), we standardised the values of the index.
(3) Factor index: computed as the common factor from the factorial analysis of the chosen earnings components.
(4) PLS: estimated considering that earnings quality is a 2nd-level non-observable construct that is formatively related to other 1st-level non-observable constructs (EQ_i), related reflectively to their corresponding indicators.

We generated 20,000 observations per variable and estimated the coefficient of the structural relationship for the four approaches. Then, we computed the error for each variable as the squared value of the difference between the actual value of parameter a (0.5) and the estimated coefficient. We iterated this process 10,000 times and computed the average estimation error for each approach.

Additionally, in order to test the robustness of the multidimensional approaches (average decile ranking index, common factor and PLS estimation) to the use of limited information, we computed them considering only four, three, or two dimensions and using four, three, or two indicators per dimension.

4.2. Simulation results

Results of the simulation process are reported in Table 6.

Panel A of Table 6 reports the average quadratic estimation errors for the four approaches when all the information is used (five dimensions, five indicators per dimension). The highest error (0.1703) is observed for the Factor index. This high error is

Table 6. Simulation process results.

	Single indicator	Equally weighted index	Factor index	PLS
Panel A. Complete information				
Average squared error	0.1502	0.0822	0.1703	0.0613
			Dimensions	
	Indicators	4	3	2
Panel B. Incomplete information				
Decile ranking index	4	0.0966	0.1134	0.1352
	3	0.0972	0.1140	0.1357
	2	0.0983	0.1150	0.1366
Factor analysis index	4	0.1726	0.1741	0.1882
	3	0.1725	0.1739	0.1880
	2	0.1718	0.1731	0.1866
PLS	4	0.0758	0.0944	0.1218
	3	0.0763	0.0949	0.1221
	2	0.0773	0.0958	0.1230

This table reports the average squared estimation errors for the different approaches (single indicator, equally weighted index, factor index, and PLS). Panel A reports the results for complete information (five dimensions, five indicators per dimension). Panel B reports the results of the multidimensional approaches when the information is incomplete (using from four to two dimensions and from four to two indicators).

consistent with those studies that highlight the misspecification problems that arise when a reflective measurement model to formative relationships (Chang et al., 2016; Jarvis et al., 2003; MacKenzie et al., 2005; Rodgers & Guiral, 2011). The single indicator method exhibits the second highest error (0.1502), showing that the most common measurement method used in the previous research is likely to lead to erroneous inferences about the relationship between earnings quality and its causes or consequences.

The decile ranking indices and the PLS methods exhibit the lower estimation errors. This result is not surprising, given that these two methods consider that the relationship between the earnings properties and the earnings quality construct is formative. The difference between these two methods, however, is that, whereas the deciles ranking index considers that all the empirical indicators have the same influence on the earnings quality construct, the PLS method estimates first the factor scores for each earnings property, and then estimates the actual weights of these properties in earnings quality. The result is that the PLS method outperforms all the other methods in all the different settings.

One potential problem that cannot be solved with PLS is that, for properly estimating a formative measurement model, a full census of the dimensions is required. It may be difficult for researchers, though, to make sure that all the possible causes related to the construct are accounted for (Davick, 2014; Ringle, Sarstedt, & Straub, 2012). Additionally, as we indicated before, the consistency of PLS increases with the number of indicators (Gefen et al., 2011; Reinartz et al., 2009; Ringle et al., 2014; Wold, 1980), so a low number of indicators would increase the estimation error of this method. To assess the performance of the PLS when the information is not complete (that is, when not all the dimensions are or all the indicators are used), we computed the average square errors using less than five indicators per earnings property and less than five earnings properties. Results are reported in Panel B of Table 6.

These results show that, as the information becomes more incomplete, all the methods suffer an increase in their estimation errors.[19] Despite this increase in the errors, the PLS method still outperforms the alternative methods.

In conclusion, these simulation results show that the PLS method produces smaller estimation errors than the methods previously used for measuring earnings quality, even in situations if the researcher has not identified all the earnings properties or has not used all the different empirical indicators available for each earnings property.

5. Concluding remarks

Despite its advantages, SEM methods – and, particularly, PLS – have not been broadly applied in archival accounting research. In this paper we present an overview of SEM techniques, focusing on PLS. For highlighting its advantages, we discuss how to use it in the research on earnings quality. Thus, we review the problems associated to earnings quality measuring, comparing how the traditional methods and the PLS deal with these problems: Whereas the traditional methods try to overcome the problems making strong (and untested) implicit assumptions, PLS method allows the testing of the measurement model, giving researchers a more flexible tool for testing their hypotheses. Additionally, the results of our simulation process show that PLS typically outperforms the other approaches, even in scenarios with poor information, proving the advantages of this method.

Our paper is consistent with those papers that demand the use of more advanced techniques in empirical archival research in accounting (Gow et al., 2016; Hinson & Utke, 2018; Larcker & Rusticus, 2010; Leuz et al., 2003). Although we have used earnings quality measuring for illustrating the advantages of SEM in general, and PLS in particular, these methods can be applied to many other accounting and auditing topics where the variables of interest are non-directly observable. Concepts such as firm's performance, degree of competition, managers' incentives, managerial attitudes, investors' protection level, audit quality, and auditor's independence or competence are non-directly observable concepts that, to date, have been represented using several different proxies. The use of SEM methods can help researchers to develop measures of these concepts that can represent them more accurately.

In addition to the assessment of the validity of the measurement model, PLS presents other additional advantages over traditional methods that can also be exploited by researchers. Thus, the flexible path modelling of the relationships among the different latent variables can help to get a better understanding of the observed relationships, differentiating between direct, indirect, and mediated relationships.

Notwithstanding these advantages, PLS has also some limitations that should be considered by researchers. Thus, consistency of PLS estimations depends on the number of indicators and sample size (Dijkstra & Henseler, 2015; Goodhue et al., 2012; Hair et al., 2013; Henseler et al., 2012; Ronkko & Evermann, 2013), producing upwards-biased estimates and exhibiting low out-of-sample predictive ability when the number of indicators or the sample size is low (Davick, 2014; Dijkstra & Henseler, 2015; Gefen et al., 2011; Hair et al., 2013, 2012; Henseler et al., 2012; Reinartz et al., 2009; Rigdon, 2014, 2016; Ringle et al., 2014).

To conclude with, for the application of PLS to earnings quality measurement, a sum of the considerations to be adopted in the research design could be the following: i) Identify the different properties of accounting information that explain the behaviour of the variable earnings quality; ii) Analyse the relationship between these properties and earnings quality. In this regard, given that the properties of accounting information are jointly explaining the meaning of earnings quality, we propose a formative relationship as a more suitable relationship; iii) Measure empirically the properties that define earnings quality. To do so, the PLS allows for the inclusion of as many indicators as needed (Reinartz et al., 2009). These indicators are related to the property that they are representing in a reflective way, for they are a reflection of the higher or lower extent of that property; iv) Determine the extent of importance (weight) of the different properties to explain earnings quality. This is calculated automatically in PLS, estimating the optimal weights according to the importance of each variable in the explanation of the behaviour of earnings quality (Ullman, 2006); v) Once earnings quality is appropriately measured (measurement model), connect this variable with the explained one (structural model).

Notes

1. In this paper, we will use the terms 'latent variables' and 'constructs' indistinctly.
2. Some examples of these measures are (the absence of) abnormal accruals, the absence of discontinuities in the cross-sectional distribution of earnings, the predictability or the smoothness of reported earnings, the value relevance of earnings or book values, the

degree of accounting conservatism, investors' reactions to reported earnings, or the opinion of external parties.

3. We define conceptual variables as the representation of ideas or abstract concepts that researchers establish to design the models (Sarstedt et al., 2016).

4. For simplicity, we focus on the specific conceptualisation of earnings quality, not discussing the conceptualisation of the dependent variable.

5. Licerán-Gutiérrez and Cano Rodríguez (2018) document that more than 90% of empirical papers on earnings quality used a measure based on a single proxy (70.8% used just one proxy, 22.7% used several proxies in separate models).

6. The fewer indicators included in the model, however, the lower the reliability of the set of indicators.

7. Given that all the formative indicators influence the same construct, it can be expected some correlation among them, but the model does not assume nor require it (Jarvis et al., 2003).

8. The typical proxies for investor responsiveness are the earnings response coefficient (ERC) and the R^2 from the earnings-return model (Dechow et al., 2010). Other proxies that could be classified in this category are the value relevance of earnings or book value (Barth et al., 2008; Francis et al., 2004).

9. For instance, Gaio and Raposo (Gaio & Raposo, 2011) used an index that included both accounting-based earnings attributes (formatively related to earnings quality according to our analysis) and market-based earnings attributes (reflectively related to earnings quality according to our analysis).

10. This lack of discussion keeps being present even in those papers that try to measure earnings quality using SEM. For instance, Hinson and Utke (Hinson & Utke, 2018) use a reflective relationship between the earnings quality construct and its empirical measures (earnings volatility, the absolute value of abnormal accruals based on balance, and the absolute value of abnormal accruals based on cash-flow statement) in their SEM model, but without justifying the reasons for expecting such reflective relationship.

11. Dichev and Tang (2009) show that earnings predictability increases with earnings persistence and decreases with earnings variability, and that low-volatility earnings show greater persistence than high-volatility earnings. Additionally, by smoothing earnings, managers make earnings more predictable and with a lower variability (Chaney et al., 2008; Dechow et al., 2010).

12. This criterion states that any latent construct shares more variance with its assigned indicators than with any other latent variable in the structural model (Hair et al., 2011, 2016; Henseler et al., 2015). The measure is indicative of appropriate discriminant validity whenever the AVE of each construct is greater than its highest squared correlation with any other construct (Hair et al., 2011, 2016; Henseler et al., 2015, 2009; Roldán & Sánchez-Franco, 2012).

13. The HTMT ratio is an estimate of what would be the true correlation between two constructs if they were perfectly measured (Hair et al., 2016), and it is considered a more reliable criterion to assess discriminant validity (Nitzl, 2016) than the Fornell-Larcker criterion. The threshold for this criterion depends on whether latent variables are conceptually very similar (values above 0.90 indicate lack of discriminant validity) or, on the contrary, more distinct (values above 0.85 are unacceptable) (Hair et al., 2016; Henseler et al., 2015).

14. It is considered that the indicator represents accurately the construct when its loading exceeds 0.7 (Carmines & Zeller, 1979), although some authors consider that values between 0.40 and 0.70 are acceptable (Chin, 1998; Hair et al., 2011, 2016; Roldán & Sánchez-Franco, 2012). Values below 0.40 show that the indicator is not representing appropriately the latent variable (Hair et al., 2011; Henseler et al., 2009; Roldán & Sánchez-Franco, 2012).

15. It is considered that a given set of indicators show an acceptable degree of internal consistency when this measure exceeds values of 0.60–0.70 in exploratory research, or 0.70–0.90 in more advanced stages of research (Hair et al., 2011, 2016; Henseler et al., 2009; Numally & Bernstein, 1994; Roldán & Sánchez-Franco, 2012). Values below 0.60

indicate that the set of indicators are not representing the same construct (Hair et al., 2016; Henseler et al., 2009).

16. An AVE value higher than 0.5 is considered as acceptable, as the construct would be explaining more than the half of the variance of its indicators (Hair et al., 2011, 2016; Henseler et al., 2009; Roldán & Sánchez-Franco, 2012).

17. These values are consistent with the empirical correlations observed by Dechow et al. (2010) among the different earnings properties.

18. We repeated the simulation process with different values for parameter a (specifically, 1, 0.25 and 0.1). The results (untabulated) were not qualitatively different from those reported.

19. The estimation errors of the factor analysis index, however, get slightly reduced when the number of indicators is reduced.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

ORCID

Ana Licerán-Gutiérrez ⓘ http://orcid.org/0000-0001-9154-8685
Manuel Cano-Rodríguez ⓘ http://orcid.org/0000-0002-3482-3836

References

Aguirre-Urreta, M. I, & Marakas, G. M. (2012). Revisiting bias due to construct misspecification: different results from considering coefficients in standardized form. *Mis Quarterly*, *36*(1), 123-138.

Babbie, E. (2017). *The basics of social research* (7th ed.). Boston, United States, Boston: Wadsworth Publishing. (Cengage, Ed.)

Ball, R., & Shivakumar, L. (2005). Earnings quality in UK private firms: Comparative loss recognition timeliness. *Journal of Accounting & Economics*, *39*(1), 83–128. Retrieved from http://search.proquest.com/docview/213017798?accountid=14555

Barclay, D. W., Higgins, C. A., & Thompson, R. (1995). The partial least squares approach to causal modeling: Personal computer adoption and use as illustration. *Technology Studies*, *2*(2), 285–309.

Barth, M. E., Landsman, W. R., & Lang, M. H. (2008). International accounting standards and accounting quality. *Journal of Accounting Research*, *46*(3), 467–498. Retrieved from http://search.proquest.com/docview/206722480?accountid=14555

Becker, J. M., Rai, A., & Rigdon, E. E. (2013). Predictive validity and formative measurement in structural equation modeling: Embracing practical relevance. *Thirty Fourth International Conference on Information Systems*, Association for Information Systems, Milan, 1–19.

Bhattacharya, N., Ecker, F., Olsson, P., & Schipper, K. (2012). Direct and mediated associations among earnings quality, information asymmetry, and the cost of equity. *The Accounting Review*, *87*(2), 449–482.

Bhattacharya, U., Daouk, H., & Welker, M. (2003). The world price of earnings opacity. *The Accounting Review*, *78*(3), 641–678.

Biddle, G. C., & Hilary, G. (2006). Accounting quality and firm-level capital investment. *The Accounting Review*, *81*(5), 963–982.

Bisbe, J., Batista-Foguet, J.-M., & Chenhall, R. (2007). Defining management accounting constructs: A methodological note on the risks of conceptual misspecification. *Accounting, Organizations and Society*, *32*(7–8), 789–820.

Blanthorne, C., Jones-Farmer, L. A., & Almer, E. D. (2006). Why you should consider SEM: A guide to getting started. In V. Arnold, B. D. Clinton, P. Luckett, R. Roberts, C. Wolfe, & S. Wright (Eds.), *Advances in accounting behavioral research* (Vol. 9, pp. 179–207). Emerald Group Publishing, Limited. Bingley, United Kingdom.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, *53*, 605–634.

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

Boulton, T. J., Smart, S. B., & Zutter, C. J. (2011). Earnings quality and international IPO underpricing. *The Accounting Review*, *86*(2), 483–505.

Burgstahler, D., Hail, L., & Leuz, C. (2006). The importance of reporting incentives: Earnings management in European private and public firms. *The Accounting Review*, *81*(5), 983–1016.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment. SAGE University papers on quantitative applications in the social sciences 07-17*. Beverly Hills, CA: SAGE.

Chaney, P. K., Cooil, B., & Jeter, D. (2008). *A latent class model of earnings attributes.*(June 12, 2008). Available at SSRN: https://ssrn.com/abstract=916605

Chang, W., Franke, G. R., & Lee, N. (2016). Comparing reflective and formative measures: New insights from relevant simulations. *Journal of Business Research*, *69*(8), 3177–3185.

Chin, W. (1998). The partial least squares approach for estructural equation modeling. *Modern methods for business researchs* (pp. 295–336). Psychology Press. Hove, United Kingdom.

Chin, W. W. (2010). How to write up and report PLS analyses. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of Partial Least Squares: Concepts, methods and applications* (pp. 655–690). Berlin, Germany: Springer-Verlag. Retrieved from http://link.springer.com/10.1007/978-3-540-32827-8

Christodoulou, D., Ma, L., & Vasnev, A. (2018). Inference-in-residuals as an estimation method for earnings management. *Abacus*, *54*(2), 154–180.

Davick, N. S. (2014). The use and misuse of structural equation modeling in management research: A review and critique. *Journal of Advances in Management Research*, *35*(2008), 441–458.

Dechow, P., & Dichev, I. (2002). The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review*, *77*(Supplement), 35–59.

Dechow, P., Ge, W., & Schrand, C. (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting & Economics*, *50*(2/3), 344–401.

Dechow, P., & Schrand, C. (2004). Earnings quality. *The research foundation of CFA institute*. Charlottesville,VA., United States.

Dechow, P., Sloan, R. G., & Sweeney, A. P. (1995). Detecting earnings management. *The Accounting Review*, *70*(2), 193–225.

DeFond, M. L. (2010). Earnings quality research: Advances, challenges and future research. *Journal of Accounting and Economics*, *50*(2–3), 402–409.

Demerjian, P., Lewis, M., Lev, B., & McVay, S. (2013). Managerial ability and earnings quality. *The Accounting Review*, *88*(2), 463–498.

Dichev, I., & Tang, V. W. (2009). Earnings volatility and earnings predictability. *Journal of Accounting and Economics*, *47*(1–2), 160–181.

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

Dijkstra, T. K., & Henseler, J. (2015). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics and Data Analysis*, *81*, 10–23.

Doupnik, T. S. (2008). Influence of culture on earnings management: A note. *Abacus*, *44*(3), 317–340.

Edwards, J. R. (2001). Multidimensional constructs in organizational behavior research: An integrative analytical framework. *Organizational Research Methods*, *4*(2), 144–192.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, *5*(2), 155–174.

Ewert, R., & Wagenhofer, A. (2011). *Earnings quality metrics and what they measure* (Working Paper). University of Graz.

Ferrer, C., & Lainez, J. (2013). Detecting differences on the earnings quality measurement: Empirical evidence on Spanish firms. *Revista de Métodos Cuantitativos Para La Economía y La Empresa*, *16*, 5–28.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50.

Francis, J., LaFond, R., Olsson, P., & Schipper, K. (2004). Costs of equity and earnings attributes. *The Accounting Review*, *79*(4), 967–1010. Retrieved from http://search.proquest.com/docview/218535414?accountid=14555

Francis, J., Nanda, D., & Olsson, P. (2008). Voluntary disclosure, earnings quality and cost of capital. *Journal of Accounting Research*, *46*(1), 53–99.

Gaio, C., & Raposo, C. (2011). Earnings quality and firm valuation: International evidence. *Accounting and Finance*, *51*(2), 467–499.

Gefen, D., Rigdon, E. E., & Straub, D. W. (2011). An update and extension to SEM guidelines for administrative and social science research. *MIS Quarterly*, *35*(2), iii–A7.

Gefen, D., Straub, D. W., & Boudreau, M.-C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the Association for Information Systems*, *4*(1), 7.

Givoly, D., & Hayn, C. (2000). The changing time-series properties of earnings, cash flows and accruals: Has financial reporting become more conservative? *Journal of Accounting & Economics*, *29*(3), 287–320. Retrieved from http://search.proquest.com/docview/213053787?accountid=14555

Goodhue, D., Lewis, W., & Thompson, R. (2012). Does PLS have advantages for small sample size or non-normal data? *MIS Quarterly*, *36*(3), 981–1001.

Gow, I. D., Larcker, D. F., & Reiss, P. C. (2016). Causal inference in accounting research. *Journal of Accounting Research*, *54*(2), 477–523.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *The Journal of Marketing Theory and Practice*, *19*(2), 139–152.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). Partial least squares structural equation modeling: Rigorous applications, better results and higher acceptance. *Long Range Planning*, *46*(1–2), 1–12.

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*(5–6), 320–340.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2017). *Advanced issues in partial least squares structural equation modeling* (2nd ed.). Thousand Oaks, CA: Sage.

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

Hampton, C. (2015). Estimating and reporting structural equation models with behavioral accounting data. *Behavioral Research in Accounting*, *27*(2), 1–34.

Henri, J.-F. (2007). A quantitative assessment of the reporting of structural equation modeling information. The case of management accounting research. *Journal of Accounting Literature*, *26*, 76–115. Retrieved from http://search.proquest.com/docview/216304849?accountid=8003%5Cnhttp://pqdd.sinica.edu.tw/twdaoapp/servlet/advanced?query=

Henseler, J., Ringle, C. M., & Sarstedt, M. (2012). Using partial least squares path modeling in advertising research: Basic concepts and recent issues. In S. Okazaki (Ed.), *Handbook of research on international advertising* (pp. 252–276). Cheltenham, UK: Edward Elgar.

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., . . . Calantone, R. (2014). Common beliefs and reality about PLS: comments on Ronkko and Evermann (2013). *Organizational Research Methods*, *17*(2), 182–209.

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., & Sinkovics, R. R. (2009). The use of partial least squares structural equation modeling in marketing research. *Advances in International Marketing*, *20*, 277–320.

Henseler, J., & Sarstedt, M. (2013). Goodness-of-fit indices for partial least squares path modeling. *Computational Statistics*, *28*(2), 565–580.

Hermanns, S. (2006). *Financial information and earnings quality: A literature review*. (April 2006). Available at SSRN: https://ssrn.com/abstract=897722

Hinson, L. A., & Utke, S. (2018). *Structural equation modeling in archival capital markets research*. (July 2, 2018). Available at SSRN: https://ssrn.com/abstract=2614085

IASB. (2010). *The conceptual framework for financial reporting*. International Accounting Standards Board, London.

Jackson, A. B. (2018). Discretionary accruals: Earnings management . . . or Not? *Abacus*, *54*(2), 136–153.

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, *30*(2), 199–218.

Jones, K. L., Krishnan, G. V., & Melendrez, K. D. (2008). Do models of discretionary accruals detect actual cases of fraudulent and restated earnings? An empirical analysis. *Contemporary Accounting Research*, *25*(2), 6. Retrieved from http://search.proquest.com/docview/194220750?accountid=14555

Laksmana, I., & Yang, Y. (2009). Corporate citizenship and earnings attributes. *Advances in Accounting*, *25*(1), 40–48.

Larcker, D. F., & Rusticus, T. O. (2010). On the use of instrumental variables in accounting research. *Journal of Accounting and Economics*, *49*(3), 186–205.

Lee, L., Petter, S., Fayard, D., & Robinson, S. (2011). On the use of partial least squares path modeling in accounting research. *International Journal of Accounting Information Systems*, *12* (4), 305–328.

Leuz, C., Nanda, D., & Wysocki, P. D. (2003). Earnings management and investor protection: An international comparison. *Journal of Financial Economics*, *69*(3), 505–527.

Leuz, C., & Wysocki, P. D. (2016). The economics of disclosure and financial reporting regulation: Evidence and suggestions for future research. *Journal of Accounting Research*, *54*(2), 525–622.

Libby, R., Bloomfield, R., & Nelson, M. W. (2002). Experimental research in financial accounting. *Accounting, Organizations and Society*, *27*(8), 775–810.

Licerán-Gutiérrez, A., & Cano-Rodríguez, M. (2019). A review on the multidimensional analysis of earnings qualilty. *Revista de Contabilidad, 22(1),* 41-60.

MacKenzie, S. B., Podsakoff, P. M., & Jarvis, C. B. (2005). The problem of measurement model misspecification in behavioral and organizational research and some recommended solutions. *Journal of Applied Psychology*, *90*(4), 710–730.

Mackenzie, S. B., Podsakoff, P. M., & Podakoff, N. P. (2011). Construct measurement and validation procedures in mis and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, *35*(2), 293–334.

Mateos-Aparicio, G. (2011). Partial least squares (PLS) methods: Origins, evolution, and application to social sciences. *Communications in Statistics - Theory and Methods*, *40*(13), 2305–2317.

McNichols, M. F., & Stubben, S. R. (2018). Research design issues in studies using discretionary accruals. *Abacus*, *54*(2), 227–246.

Nitzl, C. (2016). The use of partial least squares structural equation modelling (PLS-SEM) in management accounting research: Directions for future theory development. *Journal of Accounting Literature*, 37, 19–35.

Numally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: Mc Graw Hill.

Penman, S. H., & Zhang, X.-J. (2002). Accounting conservatism, the quality of earnings, and stock returns. *The Accounting Review*, 77(2), 237–264. Retrieved from http://search.proquest. com/docview/218538202?accountid=14555

Perotti, P., & Wagenhofer, A. (2014). Earnings quality measures and excess returns. *Journal of Business Finance & Accounting*, 41(5–6), 545–571.

Reinartz, W., Haenlein, M., & Henseler, J. (2009). An empirical comparison of the efficacy of covariance-based and variance-based SEM. *International Journal of Research in Marketing*, 26 (4), 332–344.

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. (2014). Rethinking partial least squares path modeling: Breaking chains and forging ahead. *Long Range Planning*, 47(3), 161–167.

Rigdon, E. E. (2016). Choosing PLS path modeling as analytical method in European management research: A realist perspective. *European Management Journal*, 34, 1–8.

Ringle, C. M., Sarstedt, M., & Schlittgen, R. (2014). Genetic algorithm segmentation in partial least squares structural equation modeling. *OR Spectrum*, 36(1), 251–276.

Ringle, C. M., Sarstedt, M., & Straub, D. W. (2012). A critical look at the use of PLS-SEM in MIS Quarterly. *MIS Quarterly*, 36(1), 3–14.

Rodgers, W., & Guiral, A. (2011). Potential model misspecification bias: Formative indicators enhancing theory for accounting researchers. *The International Journal of Accounting*, 46(1), 25–50.

Roldán, J. L., & Sánchez-Franco, M. J. (2012). Variance-based structural equation modeling: Guideliness for using partial least squares in information systems research. In M. Mora, O. Gelman, A. Steenkamp, & M. S. Raisinghani (Eds.), *Research methodologies, innovations and philosophies in software systems engineering and information systems* (pp. 193–211). Hershey, PA: IGI Global.

Ronkko, M., & Evermann, J. (2013). A critical examination of common beliefs about partial least squares path modeling. *Organizational Research Methods*, 16(3), 425–448.

Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies! *Journal of Business Research*, 69(10), 3998–4010.

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2014). PLS-SEM: Looking back and moving forward. *Long Range Planning*, 47(3), 132–137.

Schipper, K., & Vincent, L. (2003). Earnings quality. *Accounting Horizons*, 17, 97–110.

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

Ullman, J. B. (2006). Structural equation modeling : Reviewing the basics and moving forward. *Journal of Personality Assessment*, 87(1), 35–50.

VanTendeloo, B., & Vanstraelen, A. (2008). Earnings management and audit quality in Europe: Evidence from the private client segment market. *European Accounting Review*, 17(3), 447–469.

Wold, H. (1980). *Model construction and evaluation when theoretical knowledge is scarce* (pp. 47–74). Bureau of Economic Research. Cambridge, United Kingdom.

Wold, H. (1985, September). Systems analysis by partial least squares. *Measuring the unmeasurable*, Llumina Press. NYC, United States.