

Moderating Effects of Management Control Systems and Innovation on Performance. Simple Methods for Correcting the Effects of Measurement Error for Interaction Effects in Small Samples.

Germà Coenders^{*}, Josep Bisbe[†], Willem E. Saris[‡], Joan Manuel Batista-Foguet[†]

Departament d'Economia
Universitat de Girona

Girona, June 2003

Abstract

In the accounting literature, interaction or moderating effects are usually assessed by means of OLS regression and summated rating scales are constructed to reduce measurement error bias. Structural equation models and two-stage least squares regression could be used to completely eliminate this bias, but large samples are needed. Partial Least Squares are appropriate for small samples but do not correct measurement error bias. In this article, disattenuated regression is discussed as a small sample alternative and is illustrated on data of Bisbe and Otley (in press) that examine the interaction effect of innovation and style of use of budgets on performance. Sizeable differences emerge between OLS and disattenuated regression.

Keywords: measurement error, interaction effects, disattenuation, small samples, moderated regression, reliability, cronbach's alpha.

JEL Classification: C43, C51, C30, M41, O31

^{*} Address: Departament d'Economia, Universitat de Girona, Campus Montilivi, 17071 Girona, Spain. E-mail: germa.coenders@udg.es

[†] ESADE, Universitat Ramon Llull, Barcelona, Spain.

[‡] Department of Methods and Techniques. University of Amsterdam, the Netherlands.

1. Introduction

Interaction effects are relevant to a number of research problems in management and accounting. When an interaction effect is present, the effect of one variable on the dependent variable is different for different values of another explanatory variable. If variables are measured without error, ordinary least squares (OLS) regression can be used to estimate and test interaction effects. If variables are measured with error, OLS lead to biased estimates. In this case, the management literature (e.g. Abernethy and Brownell, 1999; Li and Atuahene-Gima, 2001) has often used summated rating scales (SRS, Likert, 1932, for an introduction see Spector, 1992) by averaging multiple indicators of each variable of interest, and then using these SRS as variables in the OLS regression model. SRS are constructed to reduce the effect of measurement error bias. However this bias is not eliminated completely as the SRS are not perfectly reliable, even if more so than single items.

In many research problems in the management literature, sampling units are firms, and thus large sample sizes are hard to obtain. Many methods for modelling variables measured with error have been suggested in the literature but they either rely on large samples or do not properly correct measurement error bias:

- Structural equation models (Goldberger & Duncan, 1973; for a non-technical introduction see for instance Raykov & Marcoulides, 2000) can be used to completely eliminate measurement error bias. In recent years, different developments have been proposed to examine interaction effects under this approach (Kenny & Judd, 1984; Jaccard & Wan, 1995; Jöreskog & Yang, 1996; Ping, 1995). The merits of structural equation models for tackling this type of problems are universally understood, though their appropriateness for such small samples is still controversial. Moreover, their application to modeling interaction effects requires complex non-linear constraints among parameters, specialised software and a great deal of expertise.
- Two-stage least squares regression (Koopmans & Hood, 1953; see Johnston, 1972 for an introduction, Bollen, 1996 for its application to measurement error correction and Bollen & Paxton 1998 for its use in interaction models) also eliminates the bias completely but does so by using a limited information estimator, and as such is specially vulnerable to small samples (Bollen & Paxton, 1998) and to specification errors (Jöreskog & Sörbom, 1989). In particular, the estimates depend on the arbitrary choice of which indicators are used as regressors and which as instruments.
- Partial least squares (Wold, 1975; see Fornell & Cha, 1994 for a non-technical introduction) constitute also a rather complex procedure that is not so far from OLS on SRS, from which it differs by the fact that the weights of the indicators are not equal but computed from the optimisation of certain criteria. In the same way as OLS on SRS, they have the particularity that they do not eliminate measurement error bias, as they are only consistent under perfect reliability or an infinite number of items per dimension

(Fornell & Cha, 1994; Dijkstra, 1983; O’Loughlin & Coenders, 2002; Wold, 1982). On the contrary, this technique shares with OLS regression the property of providing optimal predictions and has successfully been applied for predictive purposes or whenever the aim of the analysis is exploratory, the theory is weak, or the number of variables is too large for formal modelling. Wold (1982) introduced the term “soft modeling” to refer to these situations. In this paper we concentrate on other purposes of modeling, namely estimation and testing (for which unbiasedness is a key requirement) and our model is small and based on a well-grounded theory.

- OLS regression on factor scores (e.g. Hair, Anderson, Tatham & Black, 1998) is similar to partial least squares in that it computes composite scores from a weighted sum of the indicators and in that it does not correct for measurement error bias. Even if certain criteria for the computation of factor scores optimize their reliability, this reliability would only be equal to one under perfect reliability of the items or an infinite number of them. The weights of the items used to compute the composite scores do differ from partial least squares, but not to a large extent. Anyway, changing the weights of the indicators is reported to have a minor impact on results (McDonald, 1996).

In this article, disattenuated regression on SRS (Lord & Novick, 1968) is discussed as an alternative method that is unbiased while being less demanding than the aforementioned procedures, both in terms of statistical expertise and of sample size. In short, this method estimates the reliability of the SRS, uses this information to compute the variances of the SRS that would have been obtained in absence of measurement error, and substitutes these variances in the covariance matrix from which OLS estimates are computed. Obtaining a good reliability estimate is of course a crucial issue. Two reliability estimates are discussed in the paper, α (Chronbach, 1951) and Ω (Heise & Bohrnstedt, 1970). α can be obtained from the covariances among items while Ω requires estimating a factor analysis model (Spearman, 1904a, for a non-technical introduction see Kim & Mueller, 1978a, 1978b), but Ω is based on less stringent assumptions. The estimation of the reliability of interaction terms cannot be done by just using the α and Ω formulae, but a simple alternative exists.

The procedure will be illustrated on an example of Bisbe and Otley (in press) who estimated and tested the interaction effect from use of management control systems (MCS) and innovation on performance using OLS on SRS. The results obtained under OLS and disattenuated regression will be compared.

The paper is structured as follows. First we summarize the model of Bisbe and Otley (in press) and the data collection mode and measurement instruments used. After that we briefly discuss the alternative methods to take measurement errors into account and present the model. Then we provide the results of the different analyses and sizeable differences emerge depending on whether OLS or disattenuated regression is selected.

2. Use of management control systems, innovation and performance

The management literature has long considered innovation to be one of the major determinants of long-term organizational performance in contemporary environments (e.g. Clark & Fujimoto, 1991). Most empirical studies (e.g. Capon, Farley, Lehmann & Hulbert, 1992) have shown a positive relationship between innovation and performance. Understanding how an organisation can use its control systems to support both innovation and performance has also emerged as an important research question (Shields, 1997)

Simons (1991, 1995) defines two different styles of use of MCS (e.g. budget systems, balanced scorecards, project management systems): a diagnostic and an interactive style of use. When used diagnostically, MCS are employed for setting pre-established standards and monitoring and correcting deviations, and attract the managers' attention only on an exception basis. When used interactively, MCS focus on strategic uncertainties and become a recurring forum and agenda for a continuous and challenging debate in which top managers are involved. Following Simons' framework, it can be expected that, by orientating the contents and the adequacy of the innovation initiatives, an interactive use of MCS will positively influence the success of innovation initiatives (Simons, 1991, 1995). It is plausible that the influence of the interactive use of MCS on the effect of innovation on performance is achieved by providing focus (van de Ven, 1986), by acting as an internal integrative capability (Verona, 1999) and by providing a lever for permanently fine-tuning. The enhancement of the relationship innovation-performance because of the presence of an interactive use of MCS should be particularly strong when innovation is high. Thus, the relationship between innovation and performance would be affected by the extent to which MCS are used interactively (Bisbe & Otley, in press).

In statistical terms, this effect is referred to as a moderated causal relationship in which an interaction effect is present (Hartmann & Moers, 1999; Jaccard, Turrisi & Wan, 1990; Luft & Shields, 2003; Shields & Shields, 1998). Interaction effects may be tested through moderated regression analysis, that is a regression analysis including the multiplicative term of the variables that interact as an additional variable (Hartmann & Moers, 1999). The presence of this interaction effect will make the treatment of measurement error a bit more complicated, as shown below.

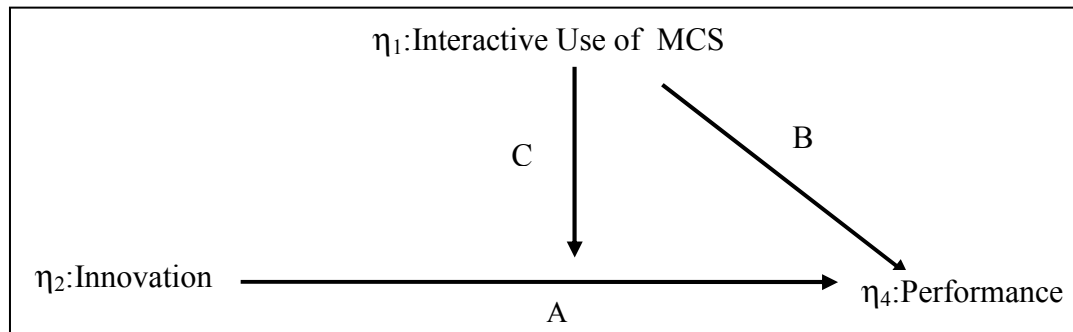


Figure 1: Conceptual framework

Bisbe (2002) and Bisbe and Otley (in press), drawing on Simons' framework (Simons, 1990, 1991, 1995) estimate and test the model in Figure 1 that includes the *interaction effect* between interactive use of MCS and innovation on performance (arrow C). Besides, Figure 1 includes the *main effects* of both innovation and interactive use of MCS on performance (arrows A and B). Hartmann and Moers (1999) and Irwin and McClelland (2001) argue that when an interaction effect is included in a model, all main effects of the variables that interact have to be introduced as well (even in cases where they would be theoretically irrelevant) in order to get meaningful estimates of the interaction effect.

3. Subjects, data collection and instruments

Data were gathered through the administration of a written questionnaire to a sample of chief executive officers of medium-sized, mature manufacturing firms with headquarters located in Catalonia, Spain. (Bisbe, 2002). Mature medium-sized firms were defined as those with an annual turnover of between 18 and 180 million euro, with between 200 and 2000 employees and founded at least ten years before the survey was administered. Exploitation of the Dun & Bradstreet/CIDEM 2000 database (referred to 1998) resulted in 120 firms fulfilling the screening criteria of size, life cycle and headquarters' location.

Once revised after pilot tests, questionnaires were distributed and traced following Dillmann's (2000) guidelines. Out of the 120 distributed questionnaires, 58 were returned, all of which were complete. Thus, the process yielded a 48.33% response rate. This compares well with the response rate of similar studies. Cases where the executives reported not to have been in their current position for at least three years (n=18) were excluded. The resulting useable sample size was n=40. The measurement instruments used for each of the variables in Figure 1 are described next.

3.1 Interactive use of budgets (MCS)

This paper focuses on the interactive use of a specific management control system, namely budgets. In the paper of Bisbe and Otley (in press), the balanced scorecard and project management systems were also considered.

Based on Simons' work (Simons, 1990, 1995, 2000), styles of use of budgets were defined in terms of the patterns of attention posed by top managers. Using Simons' framework, and developing the instruments suggested by Abernethy and Brownell (1999) and Davila (2000), Bisbe (2002) and Bisbe and Otley (in press) developed a 3-item instrument containing:

- item 1) Degree to which information from the control system is discussed face-to-face merely on an exception basis.
- item 2) Extent to which it demands frequent and regular attention from the top manager.
- item 3) Extent to which it demands frequent and regular attention from operating managers at all levels of the organisation.

The chief executive officer was asked to rate the items on a scale ranging from 1 to 7. If budgets were not used at all in a company, the items were scored zero.

3.2. Product innovation

Product innovation is understood from an output perspective and is defined as the development and launching of products that are in some respect unique and distinctive from existing products. The measure of product innovation used in Bisbe (2002) and Bisbe and Otley (in press) was drawn from instruments used by Capon et al. (1992), Thomson and Abernethy (1998) and Scott and Tiesen (1999) and contained 3 items in a 1 to 7 scale:

- item 4) Rate of introduction of new products.
- item 5) Tendency of firms to pioneer.
- item 6) Part of the product portfolio corresponding to recently launched products.

3.3. Performance

In accordance with previous research (i.e. Gupta & Govindarajan, 1984; Venkatraman & Ramanujam, 1986; Kaplan & Norton, 1996; Chenhall & Langfield-Smith, 1998, Otley, 1999), the construct performance was defined as the degree of goal attainment along several dimensions, including both financial and non-financial aspects. An adaptation of the self-rating instrument developed for evaluation of strategic business unit effectiveness (Govindarajan, 1984; Gupta & Govindarajan, 1984, Govindarajan & Gupta, 1985; Govindarajan, 1988; Chong & Chong, 1997; Chenhall & Langfield-Smith, 1998) has been used. The adaptation of the instrument proposed by Bisbe and Otley (in press) captures eight questions related to both financial (sales growth rate, revenue growth rate, return on

investment, profit/sales ratio) and customer perspectives (customer satisfaction, customer retention, customer acquisition and increase in market share). As widely accepted (see cites above), a single item (item 7) was constructed from the assessments of the firm's performance on those eight aspects compared with the industry average. When constructing item 7, the eight questions were weighted according to their perceived importance by the chief executive officers themselves, who were asked to take the firm strategy into account when assessing this importance.

4. Data analysis

4.1. Traditional method for dealing with measurement error

SRS are often used in the management literature when an unobservable concept, assumed to be unidimensional, is measured by multiple indicators. A SRS is computed as either the sum or the average of these indicators (in this paper we assume, without loss of generality that they are averaged; for summated items Equations 1 and 2 will change somewhat). This has a threefold purpose:

- Properly defining a composite construct by combining observable variables.
- Increasing measurement reliability (i.e. precision) by averaging out random errors of measurement from single indicators. This also results in higher discrimination as the composite index global range is larger.
- Increasing parsimony as only equations relating the composites (of which there are fewer than variables) are needed.

Under this approach, the analysis is very simple because one can use an OLS regression in which the SRS are used as variables. This simplicity has fostered a widespread use in the accounting and management literature (Abernethy & Brownell, 1999; Abernethy & Stoelwinder, 1995; Lal & Hassel, 1998; Li & Atuahene-Gima, 2001).

The drawback of OLS on SRS is that measurement error correction is not complete. It has long been known that a sum or an average of several measures is more reliable than just one measure (the first proof, in the field of astronomy, can be found in Simpson, 1755). However, this average is only perfectly reliable when the number of items approaches infinity or the reliability of each item approaches one.

As a result, the OLS estimates of regression coefficients will be biased (usually negatively, which is known as attenuation) due to measurement error. Biased estimates limit the use of the regression equations to purely predictive purposes; no inferences about population parameters or relationships among variables can be made. This is especially

serious for testing interaction effects, which tend to be of lower magnitude than main effects, and may easily go undetected if attenuated.

4.2. An alternative method for dealing with measurement error

Measurement error bias can be avoided by using disattenuated regression (e.g. Lord & Novick, 1968, although the idea can be traced back to Spearman, 1904b). A first step to estimate a disattenuated regression is to estimate the reliability of the SRS. Reliability is defined as 1 minus the percentage of variance of the SRS that corresponds to random measurement error. The product of the total variance of the SRS and reliability yields the so-called true variance. A disattenuated regression proceeds as a OLS regression in which true variances are substituted for total variances. Any OLS regression software that accepts covariance matrices as way of data input can thus perform a disattenuated regression. Any structural equation modeling software will also do the job.

Reliability of a SRS can be computed as Chronbach's α (Chronbach, 1951) under the assumption that items are at least tau-equivalent (e.g. Bollen, 1989, p.215-216). This assumption implies that all items are an unweighted sum of the true score plus a random error term. These random error terms are assumed not to contain any systematic component (the items thus measure the true score and only one true score), and to be mutually uncorrelated. A strong test of this assumption is hard to perform as requires special data collection designs and data analysis models and is beyond the scope of this article. However, an easily observed consequence of tau-equivalence is that all covariances among all pairs of items are equal (the opposite, of course, does not hold, i.e. covariances may be equal and yet items may not be tau-equivalent). Chronbach's α is a very popular measure and its computation is performed by most commercial software packages as:

$$\alpha = \frac{p}{p-1} \left(1 - \frac{\sum_{j=1}^p \text{var}(item_j)}{p^2 \times \text{var}(SRS)} \right) \quad (1)$$

where p is the number of items, $\text{var}(item_j)$ the variance of of the j th item and $\text{var}(SRS)$ the variance of the SRS constructed as the average of the items.

If the tau-equivalence assumption is fulfilled, the disattenuated regression estimates obtained in this way are consistent. Otherwise, α is only a lower bound for reliability (see Cortina, 1993, for a review of the properties of α), which the literature considers as being conservative and thus not too harmful. However, too low reliability estimates imply that the method will perform too strong a correction for attenuation, and thus the regression coefficient estimates will tend to be inflated, which is by no means conservative.

Unfortunately, practitioners usually do not perform any test of the tau-equivalence assumption before applying α .

A myriad of alternative estimates of reliability that are based on more relaxed assumptions is available. Among them, one of the simplest is Heise and Bohrnstedt's Ω . In order to estimate Ω , a unidimensional factor analysis model must be fitted to the items in each dimension. The estimates of the model will include the so-called communalities or percentages of true score variance in each item. Reliability is estimated as:

$$\Omega = 1 - \frac{\sum_{j=1}^p [\text{var}(\text{item}_j) \times (1 - h_j)]}{p^2 \times \text{var}(SRS)} \quad (2)$$

where h_j is the communality of the j th item. A factor analysis model requires p (number of items) to be equal or larger than three.

This measure assumes that items are congeneric. This assumption implies that all items are a *weighted* sum of the true score plus a random error term, which makes it possible that the units of measurement of the different items be different or that the contribution of the true score to the different items be different. As before, these random error terms are assumed not to contain any systematic component and to be mutually uncorrelated. As before, a strong test of this assumption is hard to perform. However, an easily observed consequence of congeneric measurement when the number of items is equal or larger than four is that the unidimensional factor analysis model fits the inter-item correlations well (the opposite does not hold, i.e. the one-factor model may perfectly fit the correlations and yet items may not be congeneric). If the model is estimated by maximum likelihood, most commercial software packages will produce a χ^2 test of the fit of the model to the correlations. Otherwise, the residual correlations may be examined one by one to check that they are all small.

In this paper we use the described simple approaches to estimate reliability of a SRS in order to perform a disattenuated regression, and we compare their results to OLS regression using Bisbe and Otley's (in press) data. The largest differences are expected between using OLS and disattenuated regression. Some differences between using α and Ω are also expected if the items turn out not to be tau-equivalent, in which case the latter should be preferred.

As stated in the introduction, other possible approaches must be discarded for our particular application; structural equation models and two-stage least squares because our sample size is small, and partial least squares and OLS regression on factor scores because, like OLS on SRS, they do not correct for attenuation, which is needed for a non predictive application like ours.

4.3. The model

The model is composed of two parts, one relating the variables, called structural part, and one making reliability explicit, called measurement part. The measurement part takes care of disattenuation.

4.3.1. Structural part

The relationships in Figure 1 are formalised in Equation 3. For simplicity we do not use a different notation for endogenous and exogenous variables:

$$\eta_4 = \beta_{41}\eta_1 + \beta_{42}\eta_2 + \beta_{43}\eta_3 + \zeta_4 \quad (3)$$

where:

- η_1 is the interactive use of budgets corrected for measurement error, centred with zero mean.
- η_2 is innovation corrected for measurement error, centred with zero mean.
- $\eta_3 = \eta_1\eta_2$ is the interaction term corrected for measurement error, centred with zero mean after computing the product. As in moderated regression analysis (e.g. Hartmann & Moers, 1999; Irwin & McLelland, 2001; Jaccard et al. 1990) the interaction is constructed as the product of both variables that interact.
- η_4 is performance corrected for measurement error, centred with zero mean.
- ζ_4 is a disturbance term, with zero expectation and uncorrelated with η_1 , and η_2 .

Mean centring of the interacting variables is always advisable in order to prevent collinearity (Li, Harmer, Duncan, Duncan, Acock & Boles 1998; Irwin and McClelland, 2001).

The interaction implies that the effect of the interactive use of budget on performance depends on the innovation level or that the effect of innovation on performance depends on the level of interactive budget use. The interpretation of a main effect like β_{42} is that occurring when the value of the other variable is zero (if η_1 and η_2 are mean centred, for the mean value of the other variable). Hartmann and Moers (1999) and Irwin and McLelland (2001) advise some caution in the interpretation of a main effect on its own, as it depends on this arbitrary centring decision. Standardisation of the η variables cannot be done as it would prevent η_3 from being equal to the product of η_1 and η_2 . This implies that standardised parameter estimates have no interpretation (Jackard et al. 1990).

4.3.2. Measurement part

For each variable the same measurement equation is assumed:

$$y_j = \eta_j + e_j \quad \text{where } j=1, \dots, 4 \quad (4)$$

where:

- y_1 and y_2 are not items but SRS of the interactive use of budget and innovation.
- $y_3 = y_1 y_2$ is the product indicator used for the interaction term.
- $y_4 = \text{item7}$, the single weighted item measuring performance.
- e_j are measurement errors with zero expectation. Besides, we assume that $e_1, e_2, e_4, \eta_1, \zeta_2$ and ζ_4 are mutually uncorrelated.

The properties of the e_3 term are more complicated as it involves the product of two variables. Jöreskog and Yang (1996) show that these properties can be derived from the measurement properties of the two variables being multiplied. In particular, if both true (η) and observed (y) scores are centred around their means, the error and true components of the interaction can be expressed as:

$$y_3 = y_1 y_2 = (\eta_1 + e_1) \times (\eta_2 + e_2) = \eta_3 + e_3 \quad (5)$$

where $e_3 = \eta_1 e_2 + \eta_2 e_1 + e_1 e_2$ and $\eta_3 = \eta_1 \eta_2$. Mean centring is a crucial issue because in this case from the results in Jöreskog and Yang (1996) it can be derived that the following covariances involving e_3 are zero:

$$\text{cov}(e_3, e_1) = \text{cov}(e_3, \eta_1) = \text{cov}(e_3, \eta_2) = \text{cov}(e_3, e_2) = 0 \quad (6)$$

4.3.3. Computation of reliability and true variance

The reliability of y_1 and y_2 is estimated as α and Ω or assumed to be equal to 1 under the OLS approach. Table 1 first shows the necessary information, that is, the variances and covariances of the two sets of items, and the communality estimates in two factor analysis models, as provided by a standard package such as SPSS 10.1. For instance, for the interactive use of budgets we have:

$$\alpha_1 = \frac{3}{2} \left(1 - \frac{3.259 + 2.743 + 2.615}{9 \times 2.125} \right) \quad (7)$$

$$\Omega_1 = 1 - \frac{3.259 \times (1 - 0.755) + 2.743 \times (1 - 0.778) + 2.615 \times (1 - 0.366)}{9 \times 2.125} \quad (8)$$

The items measuring interactive use of budgets have markedly different covariances (the ratio of the smallest over the largest covariance is 1.6) and thus the application of Chronbach's α is dubious. This results in α and Ω being somewhat different for this dimension. All α 's and Ω 's are substantially different from 1, much too different for OLS to be appropriate.

To estimate the reliability of the single indicator y_4 a direct question about overall performance was also included in the questionnaire. This question might play the role of an external criterion. If both measurements (this single question and the item7 discussed in Section 3.3) are valid for performance (measure performance and only performance), their correlation is the geometric mean of the reliabilities of both. The high value of this correlation at 0.73 makes validity a reasonable assumption. We take this figure as the reliability of the dependent variable for the disattenuated regression approach. Of course the OLS approach assumes it to be equal to 1.

Statistics for items in y_1 (budget use)				Statistics for items in y_2 (Innovation)			
Variances and covariances				Variances and covariances			
	item1	item2	item3		item4	item5	item6
item1	3.259			item4	2.984		
item2	2.291	2.743		item5	1.994	3.208	
item3	1.535	1.429	2.615	item6	1.577	1.682	1.887
$\text{var}(y_1)=\text{var}(\text{SRS}_1)$	2.125			$\text{var}(y_2)=\text{var}(\text{SRS}_2)$	2.065		
α_1	0.824			α_2	0.848		
	h_1	h_2	h_3		h_1	h_2	h_3
communalities	0.755	0.778	0.366	communalities	0.626	0.663	0.705
Ω_1	0.840			Ω_2	0.852		

Table 1: Computation of α and Ω

Once the reliability of y_j is obtained, its product by the total variance of y_j is the true score variance of y_j or the variance of y_j corrected for measurement error, that is, the estimate of the variance of η_j . For instance, for the interactive use of budgets and using α we would have:

$$\text{var}(\eta_1)=\alpha_1 \times \text{var}(y_1)=0.824 \times 2.125=1.751 \quad (9)$$

Table 2 shows the covariance matrix of y_1 to y_4 and the variances corrected for measurement error (as covariances do not change).

raw variances and covariances among dimensions				
	y_1	y_2	y_3	y_4
y_1	2.125			
y_2	0.464	2.065		
y_3	-0.286	-0.242	4.283	
y_4	0.078	-0.604	-0.593	1.003
true variances (corrected for measurement error)				
	y_1	y_2	y_3	y_4
α	1.751	1.751	3.282	0.732
Ω	1.784	1.759	3.354	0.732

Table 2: Raw and corrected variances and covariances among dimensions

The presence of interaction terms makes measurement error correction somewhat more complicated. One might think that the computation of the reliability of y_3 could be done by selecting the 9 products of items item1×item4, item1×item5,..., item3×item6 as tau-equivalent or congeneric indicators of y_3 and then estimating α or Ω in the usual way. However, these pairs of items overlap and thus can lead to correlated measurement errors. As an alternative, 3 non overlapping pairs could be used (e.g. item1×item4, item2×item5 and item3×item6) but then y_3 would fail to be equal to y_1y_2 , η_3 would fail to equal $\eta_1\eta_2$ and thus could not be interpreted as a proper interaction term anymore.

By using some properties of variances of products of bivariate normal variables (Anderson, T.W., 1984, p. 49, equation 26) we can obtain an appropriate true score variance for $\eta_1\eta_2$:

$$\text{var}(\eta_3) = \text{var}(\eta_1\eta_2) = \text{var}(\eta_1)\text{var}(\eta_2) + \text{cov}^2(\eta_1, \eta_2) \quad (10)$$

where $\text{var}(\eta_1)$ and $\text{var}(\eta_2)$ can be estimated from the combination of $\text{var}(y_1)$, $\text{var}(y_2)$ and the reliability estimates, while $\text{cov}(y_1, y_2)$ is by itself an unbiased estimate of $\text{cov}(\eta_1, \eta_2)$. For instance, under the Ω approach, $\text{var}(\eta_3)$ in Table 2 is computed as:

$$\text{var}(\eta_3) = 1.784 \times 1.759 + 0.464^2 = 3.354 \quad (11)$$

Kolmogorov-Smirnov tests of normality of y_1 and y_2 did not allow us to reject normality even at the non-conservative $\alpha=10\%$, which makes this approach to be appropriate for our data.

```

DISATTENUATED. REGR. OMEGA.
OBSERVED VARIABLES
Y1 Y2 Y3 Y4
COVARIANCES
  1.784
  0.464  1.759
-0.286 -0.242  3.354
  0.078 -0.604 -0.593  0.732
SAMPLE SIZE: 40
RELATIONSHIPS
Y4 = Y1 Y2 Y3
END OF PROBLEM

```

Figure 2: Sample LISREL8 input file

The covariance matrix in Table 2 with the true variances substituted is the data input required for a regression program to perform a disattenuated regression. Figure 2 shows a sample LISREL8 (Jöreskog & Sörbom, 1993) syntax file used to estimate Equation 3 using disattenuated regression with the Ω reliability estimate. Any other software program that can estimate a regression equation from a covariance matrix supplied by the user can be used.

5. Results

The estimates for the structural part of the model are displayed in Table 3. The percentage of explained variance for performance is very different depending on whether correction for attenuation is employed or not. The R^2 obtained without correcting for attenuation is in the range of those encountered in the extant organisational literature (Abernethy & Brownell, 1999; Abernethy & Stoelwinder, 1995; Lal & Hassel, 1998; Li & Atuahene-Gima, 2001). Both disattenuated regression variants provide a similar and substantially higher R^2 . The regression coefficients are also lower for OLS and higher and roughly similar for both methods correcting for attenuation, though slightly higher for α as a low bound for reliability and thus a high bound for correction is employed. For this particular data set, the significant relationships at the 5% level are the same for all approaches, though this does not need to be the case under all circumstances. In fact, β_{43} is significant by a much narrower margin for OLS than for any of the disattenuated regression approaches. Though in this case α and Ω has yielded similar estimates, this does not need to be the case under all circumstances.

parameter	OLS			Disattenuated using α			Disattenuated using Ω		
	estimate	t	approx. p-value	estimate	t	approx. p-value	estimate	t	approx. p-value
β_{41}	0.088	0.9	0.374	0.119	1.6	0.118	0.117	1.5	0.142
β_{42}	0.330	3.4	0.002	0.404	5.3	0.000	0.401	5.2	0.000
β_{43}	0.151	2.3	0.027	0.200	3.7	0.001	0.196	3.6	0.001
R^2	0.295			0.508			0.502		

Table 3: Estimates under 3 approaches

If we take the approach that is correct under the mildest assumptions (disattenuated regression using Ω), the relevance of the interaction effect is revealed by its t-value equal to 3.6 which clearly leads to rejecting the hypothesis $\beta_{43}=0$. The sign of the coefficient is positive, and thus the hypothesis imbedded in arrow C of Figure 1 is supported: the effect of innovation on performance will be higher for higher interactive use of budgets.

Testing the significance of the interaction effect implies that the main effects β_{42} and β_{41} must also be in the model regardless of whether they are significant or not (Hartmann & Moers, 1999; Irwin & McClelland, 2001). In fact, as a main effect refers to a value of zero of the other main effect, its statistical significance depends the origin of the variables (e.g. on whether they are mean centred or not) and is thus rather arbitrary.

6. Discussion

In the analysis using disattenuated regression with the Ω reliability estimate, we found support for a positive interaction effect of interactive use of budgets and innovation on performance. Estimates and R^2 greatly differ from Bisbe and Otley's (whose method was OLS on SRS), which argues for the fact that large biases can be encountered when correction for attenuation is omitted. Correction for attenuation has traditionally been carried out with Chronbach's α . The use of Heise and Bohrnstedt's Ω is not substantially more complex and relies on milder assumptions and should thus be preferred on a general basis, even if for our particular data set, results obtained for α and Ω were quite similar. In the particular case of interaction effects one must of course check the condition that SRS be normally distributed, which, being weighted sums of variables, will relatively often be the case. Disattenuated regression can also be applied for regression models with only main effects (e.g., without interactions). The procedure is essentially the same but becomes much simpler as Equations 5, 6, 10 and 11 are not needed and normality is not required.

The standard errors (and thus the t-values) of disattenuated regression tend to underestimate the uncertainty due to measurement error, thus leading to the possibility of true null hypotheses being rejected. Correct standard errors can only be obtained by methods that are more demanding in terms of sample size like structural equation models

and two-stage least squares. We attempted to use two-stage least squares on the small Bisbe and Otley's sample and repeated the estimation by changing the selection of regressor and instrumental variables, and we obtained sets of estimates that showed even higher differences than those observed between disattenuated and OLS regression.

Partial least squares are claimed to produce correct standard errors with resampling methods like the jackknife or the bootstrap. However, even if partial least squares themselves do not require large sample sizes, resampling methods like the ones mentioned do (e.g. Efron, 1979; LePage & Billard, 1992; Efron & Tibshirani, 1993; Shao & Tu, 1995). Besides, since partial least squares do not correct for measurement error bias, one may wonder what the use is for a correct standard error around a biased point estimate. We attempted to use partial least squares on our data and results regarding point estimates were very similar to those of OLS. This finding is in accordance with that of McDonald (1996) who shows that merely changing the weights of the items used to compose the scale has a minor impact on the results. Nevertheless, there has been a growing interest in the partial least squares method (even if the technique dates back to the 1970's, half of its applications reported in the Social Sciences Citation Index in march 2003 were published in 1999 or later). This is probably due to a mystification of the "soft modeling" term (McDonald, 1996). In fact, many of these applications of partial least squares are for non-predictive purposes, for which the presence of bias is a fundamental drawback. Our particular application could not be furthest from a soft-modelling situation: our aim is parameter estimation and testing, the number of variables is small, and theory is well grounded.

In the particular case of interaction effects, partial least squares have still one further disadvantage. Under this technique, the interaction term is built as a weighted sum of product indicators (e.g. $item1 \times item4$, $item2 \times item5$,...) that fails to be equal to $y_1 y_2$, which compromises its interpretation as a proper interaction term.

OLS regression on factor scores was also tried on Bisbe and Otley's data set. Results were very similar to those obtained under partial least squares, and thus were also not far from those of OLS regression on SRS. In fact, partial least squares scores are very close to being principal components, which are considered by some as a special case of factor scores. Principal components analysis, even if a statistical technique on its own right, was primitively used as an estimation method for factor analysis models.

Of course, disattenuated regression on factor scores is also a feasible alternative. Formulae to compute the reliability of a factor score are different from those used to compute the reliability of a SRS, but are still relatively simple. When we tried this approach on Bisbe and Otley's data we obtained similar results to disattenuated regression on SRS. The main disadvantage of factor scores is that the weights used to compute the scores are sample dependent, which may make comparative research problematic.

All these considerations enable us to make some recommendations for practitioners. If the aim of the analysis is parameter estimation or theory testing, then unbiasedness is a key requirement. In this situation, if sample size is large enough, structural equation models

and two stage least squares will be the approaches of choice. The relative interchangeability of both approaches opens application up to both researchers with an econometric and a psychometric or sociometric background. For small sample sizes, disattenuated regression should be used. Reliability should best be computed by using Ω , as it relies on milder assumptions than the by now more widely used α .

If the aim of the analysis is prediction, then the appropriate techniques would be OLS regression on SRS, OLS regression on factor scores and partial least squares, no matter what the sample size is. The resulting equations will yield optimal predictions of the dependent composite scores conditional on the explanatory composite scores, but parameter estimates will not necessarily reflect any population characteristic or relationship.

References

- Abernethy, M.A., & Brownell, P. (1999), The role of budgets in organizations facing strategic change: an exploratory study, *Accounting, organizations and Society* 24, 189-204.
- Abernethy, M.A., & Stoelwinder, J.U. (1995), The role of professional control in the management of complex organizations, *Accounting, Organizations and Society* 20, 1-17.
- Anderson, T.W. (1984), *An introduction to multivariate statistical analysis* (2nd ed.), Wiley, New York.
- Bisbe, J. (2002), *The interactive use of management control systems on product-innovating firms*, PhD Thesis, ESADE, University Ramon Llull.
- Bisbe, J., & Otley, D. T. (in press), The effects of the interactive use of management control systems on product innovation, *Accounting, organizations and society*.
- Bollen, K. A. (1989), *Structural equations with latent variables*, Wiley, New York.
- Bollen, K. A. (1996), An alternative two stage least squares (2SLS) estimator for latent variable equations, *Psychometrika* 35, 89-110.
- Bollen, K.A., & Paxton, P. (1998), Interactions of latent variables in structural equation models, *Structural Equation Modeling* 5, 267-293.
- Capon, N., Farley, J.U., Lehmann, D.R., & Hulbert, J.M. (1992), Profiles of product innovators among large US manufacturers, *Management Science* 38, 157-169.
- Chenhall, R.H., & Langfield-Smith, K. (1998), The relationship between strategic priorities, management techniques and management accounting: an empirical investigation using a systems approach, *Accounting, Organizations and Society* 23, 243-264.
- Chong, V.K., & Chong, K.M. (1997), Strategic choices, environmental uncertainty and SBU performance: a note on the intervening role of management accounting systems, *Accounting and Business Research* 27, 268-276.
- Chronbach, L.J. (1951), Coefficient alpha and the internal structure of tests, *Psychometrika* 16, 297-334.

- Clark, K.B., & Fujimoto, T. (1991), *Product development performance*, Harvard Business School Press, Boston.
- Cortina, J.M. (1993), What is coefficient alpha? An examination of theory and applications, *Journal of Applied Psychology* 78, 98-104.
- Davila, A. (2000), An empirical study on the drivers of management control systems design in new product development, *Accounting, Organizations and Society* 25, 383-410.
- Dijkstra, T. (1983), Some comments on maximum likelihood and partial least squares methods, *Journal of Econometrics* 22, 67-90.
- Dillman, D.A. (2000), *Mail and Internet Surveys*, Wiley, New York.
- Efron, B. (1979), Bootstrap methods: another look at the jackknife, *Annals of Statistics*, 7, 1-26.
- Efron, B., & Tibshirani, R. J. (1993), *An Introduction to the bootstrap*, Chapman & Hall, New York.
- Fornell, C., & Cha, J. (1994), Partial least squares, in Bagozzi, R.P. (ed.), *Advanced methods in marketing research*, Blackwell, Cambridge (pp. 52-78).
- Goldberger, A. S., & Duncan, O. D. (1973), *Structural equation models in the social sciences*, Academic Press, New York.
- Govindarajan, V. (1984), Appropriateness of accounting data in performance evaluation: an empirical examination of environmental uncertainty as an intervening variable, *Accounting, Organizations and Society* 9, 33-48.
- Govindarajan, V. (1988), A contingency approach to strategy implementation at the business-unit level: integrating administrative mechanisms with strategy, *Academy of Management Journal* 31, 828-853.
- Govindarajan, V., & Gupta, A.K. (1985), Linking control systems to business unit strategy: impact on performance, *Accounting, Organizations and Society* 10, 51-66.
- Gupta, A.K., & Govindarajan, V. (1984), Business unit strategy, managerial characteristics and business unit effectiveness at strategy implementation, *Academy of Management Journal* 27, 24-41.
- Hair, J.F. Anderson, R.E. Tatham, R.L., & Black, W.C. (1998), *Multivariate data analysis*, Prentice Hall, Englewood Cliffs, NJ.
- Hartmann, F.G., & Moers, F. (1999), Testing contingency hypothesis in budgetary research: an evaluation of the use of moderated regression analysis, *Accounting, Organizations and Society* 24, 291-315.
- Heise, D.R., & Bohrnstedt, G.W. (1970), Validity, invalidity and reliability, in Borgatta, E.F., & Bohrnstedt, G.W. (eds), *Sociological Methodology 1970*, Jossey-Bass, San Francisco (pp. 104-129).
- Irwin, J.R., & McClelland, G. (2001), Misleading heuristics and moderated multiple regression models, *Journal of Marketing Research* 38, 100-109.
- Jaccard, J., Turrisi, R., & Wan, C.K. (1990), *Interaction effects in multiple regression*. Sage, Newbury Park.

- Jaccard, J., & Wan, C.K. (1995), Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: multiple indicator and structural equation approaches, *Psychological bulletin* 116, 348-357.
- Johnston, J. (1972), *Econometric methods* (2nd ed), McGraw Hill, New York.
- Jöreskog, K. G., & Sörbom, D. (1989), *LISREL 7. A guide to the program and applications* (2nd. ed), SPSS Inc., Chicago.
- Jöreskog, K. G., & Sörbom, D. (1993), *LISREL8: structural equation modeling with the SIMPLIS command language*, Scientific Software International, Chicago.
- Jöreskog, K. G., & Yang, F. (1996), Nonlinear structural equation models: the Kenny-Judd model with interaction effects, in Marcoulides, G.A., & Schumacker, R.E. (eds), *Advanced structural equation modeling*, Lawrence Erlbaum, Mahwah (pp. 57-88).
- Kaplan, R.S., & Norton, D.P. (1996), *The balanced scorecard*, Harvard Business School Press, Boston.
- Kenny, D.A., & Judd C.M. (1984), Estimating the non-linear and interactive effects of latent variables, *Psychological Bulletin* 96, 201-210.
- Kim J.O., & Mueller, C.W. (1978a), *Introduction to factor analysis. What it is and how to do it*, Sage, Thousand Oaks.
- Kim J.O., & Mueller, C.W. (1978b), *Factor analysis. Statistical methods and practical issues*, Sage, Thousand Oaks.
- Koopmans T.C., & Hood W.C. (1953), The estimation of simultaneous linear economic relationships, in Hood W.C. & Koopmans, T.C. (eds), *Studies in Econometric Method*, Yale University Press, New Haven (pp. 112-199).
- Lal, M., & Hassel, L. (1998), The joint impact of environmental uncertainty and tolerance of ambiguity of top managers' perceptions of the usefulness of non-conventional management accounting information, *Scandinavian Journal of Management* 14 , 259-271.
- LePage, R., & Billard, L. (1992), *Exploring the limits of bootstrap*, Wiley, New York.
- Li, H., & Atuahene-Gima, K. (2001), Product innovation strategy and the performance of new technology ventures in China, *Academy of Management Journal* 44, 1123-1134.
- Li, F., Harmer, P., Duncan, T.E. Duncan, S.C., Acock, A., & Boles, S. (1998), Approaches to testing interactions effects using structural equation modeling methodology, *Multivariate Behavioral Research* 33, 1-39.
- Likert, R. (1932), A technique for the measurement of attitudes, *Archives of Psychology* 22, 140.
- Lord, F. M., & Novick, M. R. (1968), *Statistical theories of mental test scores*, Addison-Wesley, Reading.
- Luft, J., & Shields, M.D. (2003), Mapping management accounting: graphics and guidelines for theory-consistent empirical research, *Accounting, Organizations and Society* 28,, 169-250.
- McDonald, R.P. (1996), Path analysis with composite variables, *Multivariate Behavioral Research* 31, 239-270.

- O'Loughlin, C., & Coenders, G. (2002), Application of the European customer satisfaction index to postal services. Structural equation models versus partial least squares. *Working Papers of the Department of Economics, University of Girona 4*, 1-28. <http://www.udg.edu/fcee/economia/english/document.htm>
- Otley, D.T. (1999), Performance management: a framework for management control systems research, *Management Accounting Research 10*, 363-382.
- Ping, R.A. (1995), A parsimonious estimating technique for interaction and quadratic latent variables, *Journal of Marketing Research 32*, 336-347
- Raykov, T., & Marcoulides, G.A. (2000), *A first course in structural equation modeling*, Lawrence Erlbaum, Mahwah.
- Scott, T.W., & Tiessen, P. (1999), Performance measurement and managerial teams, *Accounting, Organizations and Society 24*, 263-285.
- Shao, J., & Tu, D. (1995), *The Jackknife and bootstrap*, Springer, New York.
- Shields, M. (1997), Research in management accounting by North Americans in the 1990s', *Journal of Management Accounting Research 9*, 3-61.
- Shields, J.F., & Shields, M.D. (1998), Antecedents of participative budgeting, *Accounting, Organizations and Society 23*, 49-76.
- Simons, R. (1990), The role of management control systems in creating competitive advantage: new perspectives, *Accounting, Organizations and Society 15*, 127-143.
- Simons, R. (1991), Strategic orientation and top management attention to control systems, *Strategic Management Journal 12*, 49-62.
- Simons, R. (1995), *Levers of control*, Harvard Business School Press, Boston.
- Simons, R. (2000), *Performance measurement and control systems for implementing strategies*, Prentice Hall, Upper Saddle River.
- Simpson, T. (1755), A letter to the right Honorable George Earl of Marclesfield, president of the Royal Society on the advantage of taking the mean of a number of observations in practical astronomy, *Philosophical Transactions of the Royal Society 49*, 82-93.
- Spearman, C. (1904a), General intelligence, objectively determined and measured, *American Journal of Psychology 15*, 201-293.
- Spearman, C. (1904b), The proof and measurement of association between two things, *American Journal of Psychology 15*, 72-101.
- Spector, P. E. (1992), *Summated rating scale construction. An introduction*, Sage, Newbury Park
- Thomson, G., & Abernethy, M.A. (1998), Product innovation and management control system design: the concept of fit, *Paper presented at the 4th International Management Control Systems Research Conference, Reading (UK), July 1998*.
- van de Ven, A. (1986), Central problems in the management of innovation, *Management Science 32*, 590-607.

Venkatraman, N., & Ramanujam, V. (1986), Measurement of business performance in strategy research: a comparison of approaches, *Academy of Management Review* 4, 801-814.

Verona, G. (1999), A resource-based view of product development, *Academy of Management Review* 24, 132-142

Wold, H. (1975), Path models with latent variables: the NIPALS approach, in Blalock, H.M. et al. (eds.), *Quantitative sociology. International perspectives on mathematical and statistical modeling*, Academic Press, New York.

Wold, H. (1982), Soft modeling: the basic design and some extensions, in Jöreskog, K. G., & Wold, H. (eds), *Systems under indirect observation. Causality, structure, prediction*, Vol. 2, North Holland, Amsterdam (pp. 1-54).