2007

Partial aggregation for complex structural equation modelling (SEM) and small sample sizes: an illustration using a multi-stakeholder model of cooperative interorganisational relationships (IORs) in product innovation'

Tania von der Heidt
Southern Cross University

Don R. Scott
Southern Cross University

Publication details

von der Heidt, T & Scott, DR 2007, 'Partial aggregation for complex structural equation modelling (SEM) and small sample sizes: an illustration using a multi-stakeholder model of cooperative interorganisational relationships (IORs) in product innovation' paper presented to the 21st ANZAM 2007 Conference, Sydney, 4-7 December.

Partial aggregation for complex structural equation modelling (SEM) and small sample sizes: An illustration using a multi-stakeholder model of cooperative interorganisational relationships (IORs) in product innovation

Tania von der Heidt

School of Commerce and Management, Lismore, Australia

Email: tania.vonderheidt@scu.edu.au

Don Scott

School of Commerce and Management, Lismore, Australia

Email: don.scott@scu.edu.au

Partial aggregation for complex structural equation modelling (SEM) and small sample sizes: An illustration using a multi-stakeholder model of cooperative interorganisational relationships (IORs) in product innovation

ABSTRACT

Where the research task involves collecting quantitative strategic data from senior management, low response rates and corresponding small sample sizes tend to be the norm. When structural equation modelling (SEM) is proposed, the challenge of undertaking sound research can be further compounded by the need to test a complex research models. A simple statistical technique which can help in this case is data reduction through partial aggregation of measurement models within the structural model.

This paper provides a brief review of literature on aggregation and composites – the foundation for aggregation. Then it illustrates the use of partial aggregation in SEM for a six-factor model of cooperative product innovation, including three distinct factors characterising interorganisational relationships (IORs) in relation to four different stakeholder groups. Sixteen multi-item measurement models were developed and empirically tested using survey data from CEOs of 120 Australian manufacturing companies. Partial aggregation was achieved by replacing each measurement models with a single summated scale or composite. The outcomes of this approach are a set of well-fitting SEM models at an overall multi-stakeholder level and a stakeholder-specific level.

Preferred Stream: Research Methods

Keywords: Structural equation modelling, sample, survey

1. INTRODUCTION

The present-day strategic business researcher using SEM is faced with a dual problem: On the one hand, strategic-level research models are expected to deliver new or more in-depth insights, meaning that they are often more complex. On the other hand, sample sizes are declining, as it is becoming increasingly difficult to obtain the requisite primary data from key informant in business, particularly at the CEO/General Manager-level. The literature shows that the rate of response refusal is growing (Groves and Couper 1998) and that attitude toward research becomes more negative as the frequency of being surveyed increases (Helgeson, Voss et al. 2002). Fortunately, complex structural models can be assessed at different 'levels of abstraction' (Bagozzi and Heatherington 1994) or aggregation. Adopting the appropriate level of aggregation allows researchers to perform meaningful tests of model fit, despite small achieved sample size.

2. APPROACHES TO AGGREGATION

Three (Bentler and Wu 1995) or four (Bagozzi and Heatherington 1994) approaches to assessing complex structural models are recommended in the literature. Bagozzi and Heatherton call these 'levels of abstraction'. According to their conceptualisation, there are two extremes in model assessment called aggregation (most abstract) and total disaggregation (least abstract). In between these two extremes of model assessment are partial aggregation and partial disaggregation.

1. *Total disaggregation* uses each item as a separate indicator of the relevant construct. This provides the most detailed level of analysis for model testing (Bagozzi and Heatherington 1994) with psychometric properties able to be reported for each individual item. However, using this approach can also lead to high levels of error, especially when there is a large number of items (Bagozzi and Heatherington 1994). This could result in a Type II error, where a null hypothesis is accepted when the alternative should be.

2. *Total aggregation* develops a single composite variable made up of the sum of all items measuring a construct (Bagozzi and Heatherington 1994). Such an approach constitutes an aggregation of both dimensions and items. The main advantages of the total aggregation model is simplicity, the ability to capture the essence of the underlying meaning of a concept and the smoothing of random error (Bagozzi and Heatherington 1994; Baumgartner and Homberg 1996). However, it should only be used in situations where measures share sufficient common variance and where the unique properties of subdimensions is not required (Bagozzi and Heatherington 1994).

3. The *partial aggregation* approach involves the aggregation of the indicators of each dimension of the overall construct, whereby each separate underlying factor is retained (Bagozzi and Heatherington 1994). In such a case, a composite variable is created from the items of each separate dimension of the construct and become single indicators of a single factor model. SEM confirmatory factor analysis (CFA) can then be performed to test an overall model. Failure to reject this model would suggest that each of the composite variables measure a single construct (Bagozzi and Heatherington 1994). This approach to model assessment provides greater substantive content for each variable within a smaller matrix, less distraction from accumulated errors and, thereby, greater reliability (Loehlin 1992; Bentler and Wu 1995). Baumgartner and Homberg (1996) recommend that these composites be created from scales for which unidimensionality and reliability are established. Partial aggregation is frequently used to assess complex models, such as by Morgan and Hunt (1994) to assess their commitment-trust theory of relationship marketing.

4. The final approach, *partial disaggregation*, involves the creation of two or more composite variables for each construct (Bentler and Wu 1995; Dabholkar, Thorpe et al. 1996). The composites may be created from identified subdimensions of an indicator construct of the overall latent construct (Bagozzi and Heatherington 1994) or items may be allocated and aggregated randomly as "it is expected that any combination of a construct's variable indicators should yield the same model fit" (Dabholkar, Thorpe et al. 1996, p. 10). Partial disaggregation provides particular benefits of being able to assess a complex higher-order model, whilst reducing the level of random error, more stable estimates from reducing the number of parameters to be estimated and improving approximation of normality distributions (Bagozzi and Heatherington 1994; Baumgartner and Homberg 1996; Dabholkar, Thorpe et al. 1996).

3. COMPOSITES: THE FOUNDATION FOR AGGREGATION

3.1 The comparative advantage of composite scales

Measures at all levels of specificity have been widely used in both theoretical and practical research (Ironson, Smith et al. 1989). Specific (facet) scales are used to differentiate different aspects of a phenomenon, whereas general (composite and global) scales ask the respondent to combine his or her reactions to various aspects of the phenomenon in a single integrated response (Ironson, Smith et al. 1989). The three different types of scale are briefly described below and evaluated below:

- **Facet** (specific) scales are intended to cover, separately, the principal areas within a more general domain (Ironson, Smith et al. 1989). The scale measuring each individual facet should be internally homogeneous and discriminably different from the others (Ironson, Smith et al. 1989).

- **Global** scales (one form of general scale) employ single items that elicit overall impression and summary evaluations. Global scales are sometimes referred to as ‘clinical’ combinations (Einhorn 1972) because they require respondents to combine their evaluations cognitively into a global judgement (Ironson, Smith et al. 1989).
- **Composite** scales (the second form of general scale) assume the whole is equal to the sum of its principal parts (Ironson, Smith et al. 1989). They sometimes require explicit summing of the facets and are sometimes referred to as mechanical composites (Einhorn 1972). Both for predicting a criterion and for representing the overall evaluations of raters, unit-weighted additive linear models have been found to be adequate (Einhorn 1972).

In the case of research involving constructs (or latent variables), a convincing case can be made for **composite** scales (Kumar, Stern et al. 1992):

- Composites can represent the multiple aspects of a concept in a single measure. In this way it addresses the researcher’s dilemma to accommodate the richer descriptions of concepts by using multiple variables, while also maintaining parsimony in the number of variables in the multivariate model (Hair, Black et al. 2006). The predicative ability of mechanical combinations or composites is superior to that of ‘clinical’ combinations (global evaluations) (Einhorn 1972).
- In comparison with global questions, specific questions underlying a composite scale help informants cope with complexity by structuring their task, thereby reducing measurement error and enhancing the probability of obtaining convergence between informant reports (Hair et al. 2006).
- When relationships with other constructs (as opposed to variables) are the focus of inquiry, the *general* (i.e. global and composite) measures of performance provide a better fit than the individual facet scales (Kumar, Stern and Achrol 1992).
- Compared with the proxy, e.g. the single highest-loading variable, the composite represents multiple facets of a concept. Compared with factor scores, which reflect the factor loadings (low and high) of all variables on the factor, the composite is easier to replicate across studies and easier to interpret. Overall, the composite is viewed as a compromise between the use of a single proxy variable and factor score options for data reduction (Hair, Black et al. 2006).

Consequently, the application of composites in academic, as well as applied and managerial research has increased (Hair, Black et al. 2006).

3.2 Composite weighting schemes

Combining component variables into a composite involves deciding (1) how multiple criterion measures are to be weighted and combined into a composite criterion measure, (2) how symptoms and signs are to be weighted and combined into a clinical judgement (Einhorn and Hogarth 1975). These authors examined two weighting schemes - linear multiple regression and unit- (or average-)

weighting: **Linear regression models** yield weights that are optimal in terms of minimising squared error, but consume degrees of freedom in the estimation of those weights. Both for predicting a criterion and for representing the overall evaluations of raters, **unit-weighted additive linear models** (linear composites) have been found to be a viable alternative to standard regression methods because unit weights (1) are not estimated from the data and therefore to not 'consume' degrees of freedom, (2) are estimated without error (have not standard error), (3) cannot reverse the 'true' relative weights of the variables and (4) incorporate prior knowledge into the analysis. Provided that one can state the sign of the zero-order correlation between the independent and dependent variables, one can confidently use a unit-weighting scheme. Also, the unit- or average-weighted scheme has the appeal of a democratic procedure which would be particularly suited to pooling the judgments of experts (Einhorn and Hogarth 1975).

4. AN ILLUSTRATION OF PARTIAL AGGREGATION WITHIN A STRUCTURAL MODEL OF MULTI-STAKEHOLDER COOPERATIVE PRODUCT INNOVATION IORS

4.1 Summary of research model development

Academic interest in analysing the relationship between innovative performance of small and medium-sized enterprises (SMEs) and their recourse to external resources, especially through interorganisational relationships (IORs) has been growing over the last ten to fifteen years. To date the emerging literature has given limited attention to the systematic empirical assessment of the relationship *and* innovation inputs and outputs of cooperative product innovation. Furthermore, most of the literature focuses on cooperative innovation IORs with customer stakeholders, ignoring the potential role played by other external stakeholder groups, such as suppliers, industry partners and research/advisory organisations.

Building upon prior models of cooperative IORs, including Robicheaux and Coleman's (1994) strategy-structure-performance-based conceptual model of marketing channel relationship structure, a six-factor multi-stakeholder model of cooperative IORs in product innovation for Australian manufacturers was developed. The research model synthesised and incorporated concepts and measures drawn from the IOR and product innovation literatures.

The central, IOR structure-based factor or construct (measurement model) developed was *Stakeholder Involvement in Product Innovation* (SIPI). Two factors were used to predict SIPI: *Stakeholder Orientation* (SO) and *Product Innovation Orientation* (PIO). A further two factors represented outcomes of SIPI: *Relationship Quality* (RQ) and *Product Innovation Performance* (PIP). A sixth factor - *Overall Firm Performance* (OFP) - was also measured to assess the broader implications of the model. Two objective, single-item variables were also included in the model – *relative Product innovation spending* (associated with PIO) and *Sales growth* (associated with OFP). The three IOR-oriented constructs SO, SIPI and RQ were specified for *each* of the four external stakeholder groups most likely to be involved in a manufacturer's product innovation – customer, supplier, industry partner and research/advisor. The six-factor model contained sixteen multi-item measurement models totaling 153 observable variables (indicators) and involved 21 hypothesised associations. The hypotheses and a preview of their results (standard coefficients) are presented in **Table 1**.

Table 1 Results of hypothesised associations (std. coefficients)

Hypo-thesis	Association	Stakeholder type				Overall model
		Customer	Supplier	Industry partner	Researcher	
H1	SO-SIPI	.53 ***	.76 ***	.77 ***	.79 ***	.98 ***
H2	PIO-SIPI	.15 *	n.s.	n.s.	n.s.	-.12*
H3	Relative PI spending- SIPI	n.s.	n.s.	n.s.	n.s.	n.s.
H4	PIO-SO	.35 ***	.45 ***	.29 ***	.29 ***	.48 ***
H5	Relative PI spending- SO	n.s.	n.s.	n.s.	.18 *	n.s.
H6	SIPI-RQ	n.s.	.26 **	.42 ***	.35 ***	n.s.
H7	SIPI-PIP technical	n.s.	n.s.	.12 *	n.s.	n.s.
H8	SIPI-PIP market	n.s.	n.s.	n.s.	n.s.	n.s.
H9	RQ-PIP technical	n.s.	n.s.	n.s.	n.s.	n.s.
H10	RQ-PIP market	n.s.	n.s.	n.s.	n.s.	n.s.
H11	RQ-OFP	n.s.	n.s.	n.s.	n.s.	n.s.
H12	RQ-Sales growth	.17 *	n.s.	n.s.	.27 ***	.27 ***
H13	SO-RQ	.33 ***	.33 ***	.50 ***	.53 ***	.63 **
H14	Relative PI spending-RQ	n.a.	.14 *	n.a.	n.a.	.15 *
H15	SIPI-OFP	n.a.	-.20 *	n.a.	n.a.	n.s.
H16	Relative PI spending-PIP technical					.21***
H17	PIP technical-PIP market					.36***
H18	PIO-PIP technical					.67***
H19	PIO-PIP market					.40***
H20	PIP technical-OFP					-.37***
H21	PIP market-OFP					.71***

*** p <.001; ** p<.01; * p < .05; n.s. = not significant; n.a. = not applicable

As strategic data was required, the research method for primary data collection was a survey of CEOs/General Managers of Australian machinery and equipment manufacturers, predominately small and medium-sized manufacturing enterprises (SMMEs). The sample was obtained from four machinery and equipment industry associations' membership databases. Standard questionnaire design actions and response management strategies were implemented to address the typical problem of non-response error associated with mail surveys. However, in this highly competitive and, generally, declining manufacturing sector, survey fatigue and lack of time were two major obstacles to achieving adequate survey response rates. Difficulties were also encountered identifying qualified SMMEs (i.e. those with a sufficiently significant manufacturing base). Data obtained from 120 key informants was used in the research. The survey response rate averaged 12%, ranging from 3.5% of respondents sourced from the Telstra database to 28% to 44% of respondents affiliated with three machinery and equipment-based manufacturing industry associations.

4.2 Sample size issues

Sample size provides the basis for the estimation of sample error and impacts on the ability of the model to be correctly estimated (Hair, Black et al. 2006). As with any statistical method, the critical question is how large a sample is needed? Bentler and Chou (1987) suggest that in SEM the sample size requirements vary for measurement and structural models. In an ideal case, the following Bentler and Chou (1987) rules of thumb need to be satisfied in order to test measurement *and* structural models:

Measurement models

A ratio of ten responses per free parameters is required to obtain trustworthy estimates (Bentler and Chou 1987). Others suggest a rule of thumb of ten subjects per item in scale development is prudent (Flynn and Pearcy 2001). However, if data is found to violate multivariate normality assumptions, the number of respondents per estimated parameter increases to 15 (Bentler and Chou 1987; Hair, Black et al. 2006). In this research, each of the constructs to be measured had four to seven indicators, i.e. eight to fourteen parameters. Applying Bentler and Chou's 10:1 rule of thumb, a sample size of 80 to 140 was required. Applying Flynn and Pearcy's (2001) rule of thumb, a sample size of 40 to 70 would suffice. A split sample (n = 60) used to cross-validate the model met Flynn and Pearcy's criterion, but fell just below Bentler and Chou's rule.

Structural models

A ratio of five responses per free parameter is required to obtain trustworthy estimates (Bentler and Chou 1987). With a total (maximum) of 153 observables or indicators, i.e. maximum of 306 free parameters, the effective sample size required to test the trustworthiness of the model would be 1530. However, a sample size exceeding 400 to 500 becomes 'too sensitive', as almost any difference is detected, making all goodness-of-fit measures indicate poor fit (Hair, Anderson et al. 1995). Furthermore, given the survey limitations, this sample size was far from achievable. For a meaningful model assessment, some form of data reduction was required.

4.3 Converting measurement models to composites

In specifying the structural model, the **partial aggregation** approach was considered most appropriate for two reasons: (1) It retained the separate (four stakeholder) dimensions of the three multi-stakeholder or IOR constructs (SO, SIPI and RQ); (2) It provided an assessment of the multi-stakeholder model of cooperative product innovation that was less distracted by accumulated error, due to the large number of items used in the model.

Each of the sixteen constructs of interest was operationalised using a multi-item scale. Unidimensionality was assessed at the facet (first-order construct) level (e.g. for *each* of the four stakeholder groups – customer, supplier, industry partner and research/advisor), rather than for the entire scale, as per Kumar, Stern et al. (1992). Following scale purification, validation (evidence of unidimensionality, convergent and discriminant validity) and cross-validation of the construct or measurement model, all of the variables loading highly on each of the measurement models were combined using the **simple average unit weighting** (for reasons spelled out in Section 3.2), and the average score of variables was used as the replacement variable. In other words, new variables for use by replacing *each* measurement model with a single summated scale or composite.

4.4 Specifying structural models with aggregated measurement models (composites)

Two types of structural models of stakeholder cooperation in product innovation using composites were developed: (1) One overall (multi-stakeholder) model; (2) Four specific stakeholder models, i.e. one each for customer, supplier, industry partner and researcher/advisor.

Overall (multi-stakeholder) structural model

The overall structural model to be tested contained a total of 16 composites:

- twelve composites, representing four composites (customer, supplier, industry partner and research/advisor) for *each* of the three second-order, stakeholder-based constructs (SO, SIPI and RQ) and
- four[1] composites relating to the remaining three first-order factors (PIO, PIP and OFP).

In addition, two single-item measures (product innovation spending and sales growth) were included in the structural model. Given a sample size of 120, the resulting 18-variable model provides seven cases per measured variable, compared with 15 cases per predictor in standard ordinary least squares multiple regression (University of Texas 2002) and Bentler and Chou's (1987) five cases per parameter.

Specific stakeholder structural models

To test whether there were significant differences *between* stakeholder types in the extent to which they were involved in a manufacturing firm's product innovation, four specific stakeholder models were developed. Each of these four models contained nine measured variables (composites) with 13 cases per measured variable. This is more consistent with the recommended guidelines of the University of Texas (2002). All four stakeholder models were specified the same way.

[1] Originally PIP was conceived as one-dimensional, but subsequent testing revealed two dimensions.

4.5 Fit of structural models

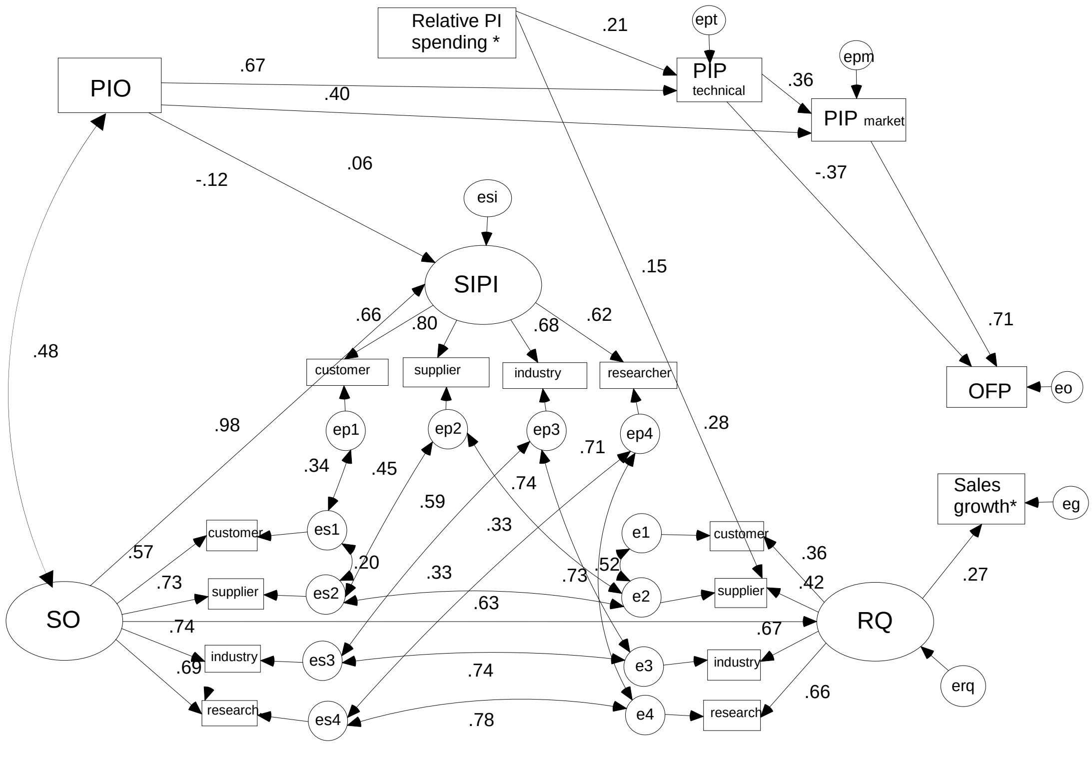
The overall (multi-stakeholder) structural model, as well as four single-stakeholder models, of key relationship- and innovation-oriented antecedents and consequences of cooperative product innovation were tested using SEM AMOS 6.0 software. Meaningful modifications of the hypothesised model were undertaken to improve model fit.[2] As is shown below, the four modified single-stakeholder models and the overall (multi-stakeholder) model were found to provide a satisfactory fit. Statistically significant standard coefficients for each of the latent constructs provided evidence of the importance of each element as an input or outcome of cooperative innovation.

Modified overall (multi-stakeholder) model

A mostly acceptable fit to the data was found, with critical fit indices falling within acceptable limits: Normed Chi-Square $X^2/df = 1.370$; CFI = 0.967. The residual fit indices provided further evidence of satisfactory fit (RMSEA = 0.056; SRMR = 0.060). The probability of significant difference between the observed (empirical) **S** and estimated (implied) Σ covariance matrices was calculated ($p = .006$). The final model for the overall (multi-stakeholder) model showing significant standardised coefficients is presented in **Figure 1**. It contains 12 paths significant at $P \leq .05$ and one path with a lower significance level - PIO to SIPI ($p = .079$).

[2] Modifications comprised the addition of four post-hoc theorised direct paths and inclusion of error correlations between error terms for stakeholder-specific composites relating to SO, SIPI and RQ. In spite of proven discriminant validity having been shown to exist, these variables had the same measurement scale and represented a series of questions on different aspects of a *related topic*, i.e. the nature of the interorganisational relationship (IOR). In addition, when classifying their attitudes toward their relationship with the top firm in each stakeholder group, the respondent had answered the questions relating to each stakeholder group in the same way. These two reasons provided theoretical support for the use of correlated errors (Dunne, Everitt et al. 1993; Kline 2003).

Figure 1Final overall (multi-stakeholder) model of IORs for cooperative product innovation



* single item

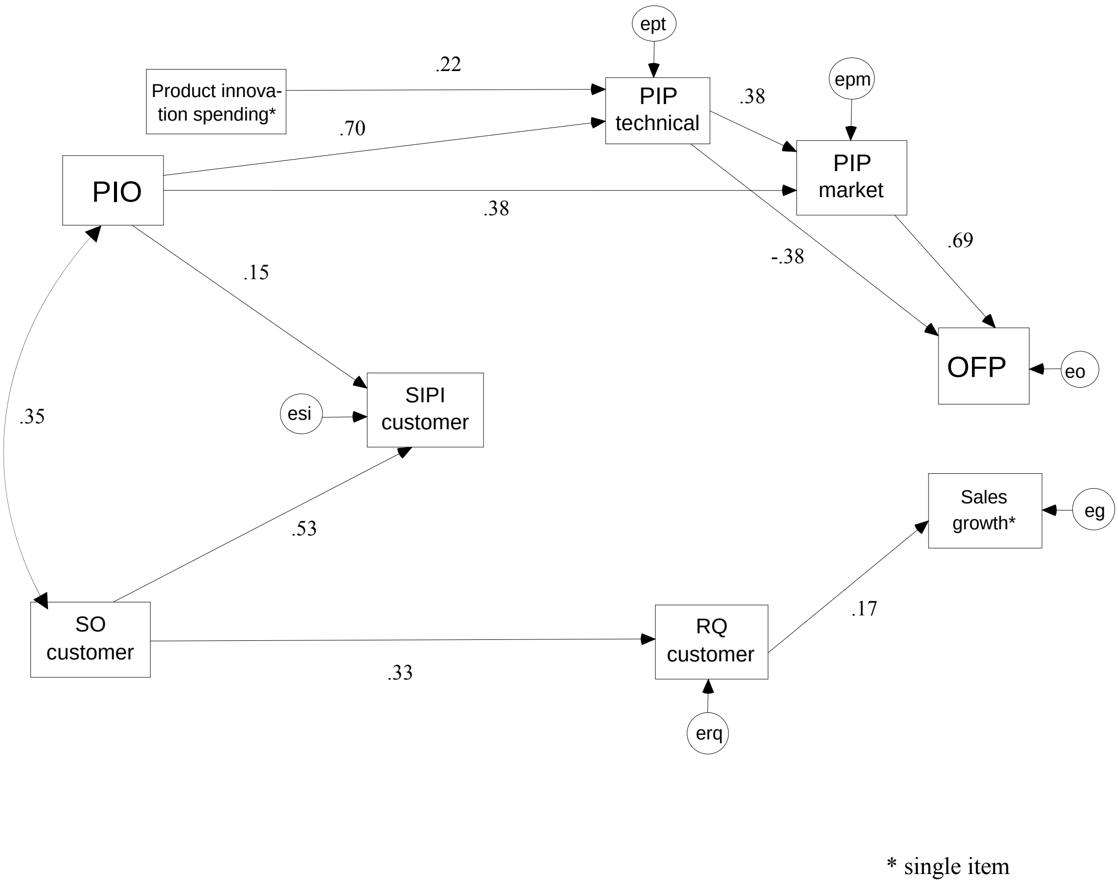
Modified single-stakeholder models

As reported in **Table 2**, a good fit to the data was found for all fit indices for all four single-stakeholder models. To illustrate, the customer model of cooperative product innovation IORs is shown in **Figure 2**.

Table 2 Fit indices for single-stakeholder models

Fit index	Customer	Supplier	Industry partner	Researcher/ advisor	Values indicating good fit
X^2 (df)	17.758 (18)	10.787 (6)	6.931 (18)	12.328 (17)	
P	.472	.822	.991	.780	> .05
X^2/df	.987	.672	.385	.725	≥ 3
SRMR	.046	.030	.035	.027	≤ .08
RMSEA	0	0	0	0	≤ .08
GFI	.969	.981	.987	.978	≥ .09
CFI	1.0	1.0	1.0	1.0	≥ .09

Figure 2 Final customer model of IOR for cooperative product innovation



5. DISCUSSION AND CONCLUSIONS

While data reduction through partial aggregation is not new, its application in a business research context is substantial and still on the rise. This is largely explained by an increase in the complexity of research models, in particular at the strategic-level, and the decline in achieved sample sizes with which to test these models.

The successful application of partial aggregation was illustrated in this paper in relation to a complex, six-factor multi-stakeholder model of cooperative product innovation IORs. It was shown how such an SEM model can be specified at a partially aggregated level, in order to retain the separate (e.g. four stakeholder) dimensions of multiple factors of interest (e.g. the three IOR factors SO, SIPI and RQ), and provide an assessment that is less distracted by accumulated error, due to the large number of items used in the model. Partial aggregation is achieved by replacing each of the measurement models developed with a single summated scale or composite, reducing the number of observables in both an overall (e.g. multi-stakeholder) model and in specific (e.g. single stakeholder) models. As a result, the ratio of responses per free parameter required to obtain trustworthy estimates can exceed Bentler and

Chou’s (1987) guideline, and a meaningful and satisfactory assessment of the model is achievable for both overall and specific model types.

Hence, partial aggregation is recommended as a statistical technique in situations where the researcher aims to test a complex model with limited achieved sample size at two levels: (1) the level of a *combined* or overall model of second-order (multi-dimensional) factors (e.g. a firm’s cooperative product innovation IORs with four different stakeholders) and (2) the level of an *individual* model of first-order (e.g. stakeholder specific) factors. In these challenging situations, partial aggregation can provide researchers with a meaningful test of model fit, which may otherwise not be possible. This is especially useful in the under-researched and hard-to-research area of business strategy pertaining to interorganisation relationships (IORs) between focal firms and multiple stakeholder groups.

REFERENCES

Bagozzi, R. P. and T. F. Heatherington (1994). "A general approach to representing mulitfaceted personality constructs: Application to self-esteem." Structural equation modeling **1**(1): 35-67.

Baumgartner, H. and C. Homberg (1996). "Applications of structural equation modeling in marketing and consumer research: A review." International Journal of Research in Marketing **13**: 139-161.

Bentler, P. M. and C.-P. Chou (1987). "Practical issues in structural modeling." Sociological Methods & Research **16**(1): 78-117.

Bentler, P. M. and E. J. C. Wu (1995). EQS for Windows User's Guide. Encino, CA, Mulitvariate Software Inc.

Dabholkar, P. A., D. I. Thorpe, et al. (1996). "A measure of service quality for retail stores: Scale development and validation." Journal of Academy of Marketing Science **24**(1): 3-16.

Dunn, G., B. Everitt, et al. (1993). Modelling Covariances and Latent Variables using EQS. London, Chapman and Hall.

Einhorn, H. J. (1972). "Expert measurement and mechanical combination." Organizational Behavior and Human Performance **7**: 86-106.

Einhorn, H. J. and R. M. Hogarth (1975). "Unit weighting schemes for decision making." Organizational Behavior and Human Performance **13**: 171-192.

Flynn, L. R. and D. Pearcy (2001). "Four subtle sins in scale development: Some suggestions for strengthening the current paradigm." International Journa of Market Research **43**(4): 409-423.

Garson, G. D. (2006). "Factor analysis." Retrieved 16/8/06, from http://www2.chass.ncsu.edu/garson/pa765/factor.htm.

Hair, J., F. Jr., W. C. Black, et al. (2006). Multivariate Data Analysis. Upper Saddle River, NJ, Pearson Prentice Hall.

Hair, J. F., Jr, R. E. Anderson, et al. (1995). Multivariate Data Analysis. Englewood Cliffs, NJ, Prentice Hall.

Helgeson, J. G., K. E. Voss, et al. (2002). "Determinants of mail-survey design factors and respondent factors." Psychology & Marketing **19**(3): 303-328.

Ironson, G. H., P. C. Smith, et al. (1989). "Construction of a job in general scale: A comparison of global, composite and specific measures." Journal of Applied Psychology **74**(2): 193-200.

Kline, R. B. (2003). How to fool yourself (and others) with SEM. Montreal, Concordia University.

Kumar, N., L. W. Stern, et al. (1992). "Assessing reseller performance from the perspective of the supplier." Journal of Marketing Research **29**(May): 238-253.

Loehlin, J., C. (1992). Latent Variable Analysis. Hillsdale, NJ, Lawrence Erlbaum Associates Inc.

Morgan, R. M. and S. D. Hunt (1994). "The commitment - trust theory of relationship marketing." Journal of Marketing(Jul).

Robicheaux, R. A. and J. Coleman (1994). "The structure of marketing relationships." Journal of the Academy of Marketing Science **22**(Winter): 38-51.

University of Texas. (2002). "Structural equation modeling using AMOS: An Introduction." Retrieved 16/8/06, 2006.