

# “Analyzing the relationship between dependent and independent variables in marketing: a comparison of multiple regression with path analysis”

|              |                                                                                                                                                                                                                        |
|--------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| AUTHORS      | Debi P. Mishra<br>Junhong Min                                                                                                                                                                                          |
| ARTICLE INFO | Debi P. Mishra and Junhong Min (2010). Analyzing the relationship between dependent and independent variables in marketing: a comparison of multiple regression with path analysis. <i>Innovative Marketing</i> , 6(3) |
| RELEASED ON  | Thursday, 14 October 2010                                                                                                                                                                                              |
| JOURNAL      | "Innovative Marketing "                                                                                                                                                                                                |
| FOUNDER      | LLC “Consulting Publishing Company “Business Perspectives”                                                                                                                                                             |



NUMBER OF REFERENCES

0



NUMBER OF FIGURES

0



NUMBER OF TABLES

0

© The author(s) 2024. This publication is an open access article.

Debi P. Mishra (USA), Junhong Min (USA)

## Analyzing the relationship between dependent and independent variables in marketing: a comparison of multiple regression with path analysis

### Abstract

Multiple regression models continue to be widely used in marketing. Within the regression framework, researchers have to grapple with and resolve several contentious issues. For example, multicollinearity, non-simultaneous estimation of parameters, inherent measurement error in independent variables, absence of overall goodness of fit indices, and lack of compelling guidelines for adding and deleting model variables are some common estimation problems associated with this method. In the absence of universally acceptable guidelines, researchers often use judgment calls to deal with these issues. Such ad-hoc approaches, in turn, compromise the potential usefulness of multiple regression models. In this paper, we position path analysis as a competing technique that can address in a relatively unambiguous way, many of the above mentioned limitations of multiple regression. We illustrate the superiority of path analysis by reanalyzing data from selected marketing studies that have used multiple regression models. To enable researchers use path analysis more frequently, we provide a technical appendix depicting use of the EQS software for estimating multiple regression models. We discuss several implications of our results and outline avenues for future research.

**Keywords:** multiple regression, path analysis, concept measurement, concept testing, psychometric theory.

### Introduction

Regression based models continue to be widely used in marketing and social sciences (Echambadi and Hess, 2007; Fitzsimons, 2008; Hagerty and Srinivasan, 1991; Irwin and McClelland, 2001; Judd and Kenny, 2010; Zhao, Lynch, and Chen, 2010). In the simplest bivariate case, a dependent variable is predicted from another independent variable. However, with the increasing complexity of marketing phenomena, bivariate models do not adequately capture underlying interrelationships among constructs. In other words, the variance of a dependent variable can be better explained by more than one independent variable. Predicting the dependent variable from more than one independent variable is termed multiple regression.

Despite the widespread popularity of multiple regression models, researchers have to grapple with several contentious issues. First, multicollinearity (high correlations among independent variables) results in biased parameter estimates and researchers do not always agree on the most appropriate approach for addressing this issue (Echambadi and Hess, 2007; Friedman and Wall, 2005; Grewal, Cote, and Baumgartner, 2004).

Second, detection of indirect variable effects in a multiple regression model is not straightforward. For example, an independent variable may affect the dependent variable through a third variable in the model. Typically, a class of intervening variable methods such as mediation analyses, moderating variable techniques, and hierarchical regressions can be used to uncover indirect effects. However, model misspecification remains a major limitation of these

approaches because intervening relationships can be incorporated in the initial model only when the researcher has apriori knowledge of theoretical relationships. In the exploratory stages of research, when theories are not well developed, it is difficult to accurately specify indirect effects in the initial regression model. Hence, it is not surprising that a growing debate is currently raging in marketing about appropriate procedures for using intervening variable methodologies (Baron and Kenny, 1986; Iacobucci, 2008; Iacobucci, Saldhana, and Deng, 2007; Judd and Kenny, 2010; Zhao, Lynch, and Chen, 2010).

Third, multiple regression models do not calculate parameter estimates simultaneously. For example, if a dependent variable ( $Y$ ) is related to an independent variable ( $X_1$ ) through another (mediating) variable ( $X_2$ ), three separate regression models have to be estimated (Baron and Kenny, 1986). Such non-simultaneous estimation might result in biased parameter estimates (Iacobucci, Saldhana, and Deng, 2007; Zhao, Lynch, and Chen, 2010).

Fourth, multiple regression models assume that independent variables are measured without error. In other words, trait variance in an independent variable should be large in relation to measurement error variance. However, a number of studies have shown that the general assumption of low levels of measurement error is highly questionable in practice. For instance, in a study utilizing 70 published studies in marketing, Cote and Buckley (1987) found that traits accounted for less than 50% of construct variance. Similar results were reported in Cote and Buckley (1988). Likewise, Mishra (2000) found that error variance accounted for 64% of total variance in the typical health care measure.

Finally, regression models do not calculate an overall goodness of fit statistic that indicates the degree of congruence between hypothesized structural paths and the underlying phenomenon of interest. To test theory, researchers typically formulate hypotheses and obtain data by administering scale items to respondents. Next, summary associational information is captured via the variance-covariance matrix of variables. Note that covariances are generated because an underlying theory or phenomenon drives responses to be associated with one another (Guttman, 1971). The initial covariance matrix, therefore, reflects theory, and serves as input for subsequent hypothesis testing. After imposing a statistical structure on a set of variables and estimating hypothesized relationships, it is possible to reestimate the original covariance matrix and calculate the discrepancy between the observed and reproduced correlation matrices. Since the original covariance matrix reflects theory, the computed difference is a measure of the degree of fit between the estimated model and underlying theory. Most overall goodness of fit indices e.g., Comparative Fit Index (Bentler, 1990; Ullman and Bentler, 2004) are premised upon the difference between observed and reproduced covariances. Path analytic software typically computes a variety of indices for assessing overall model fit.

In contrast to path analysis, regression models can at best compute a  $R^2$  value, which explains the proportion of variance in the criterion variable that is accounted for by a set of independent variables. While useful, the  $R^2$  statistic provides no indication of the overall fit between data and theory (Iacobucci, Saldhana, and Deng, 2007; Nachtigall, Funke, and Steyer, 2003).

Path analysis appears well suited to address a number of the preceding limitations of multiple regression models. For instance, one can explicitly model multicollinearity, compute parameter estimates simultaneously, detect missing paths, and compute indices for overall goodness of fit. Furthermore, structural equation models allow the researcher to explicitly account for measurement error in variables (Cote and Buckley, 1987; 1988). To the extent that such error can be modeled, the resulting parameter estimates are more robust. Note that in a standard multiple regression model, it is not possible to model measurement error and its impact on parameter estimates.

Since a majority of articles in marketing still use multiple regressions (Zhao, Lynch, and Chen, 2010), it may be worthwhile to employ path analysis and reanalyze data from published studies that have used multiple regressions. Such a comparative analysis is expected to isolate the inherent problems with regressions and provide researchers with a competing technique for estimating the relationship between dependent and independent variable. Furthermore, this

exercise will help us highlight the relative superiority of path analysis over multiple regression models. We expect that the results of this analysis will provide researchers with a clear and compelling alternative to multiple regression. Given that theory development in marketing is heavily dependent on empirical testing, our study answers the call of marketing scholars for better construct measurement approaches (Fitzsimons, 2008; Rindflesich et al., 2008).

In light of these discussions, the central objective of this paper is to position path analysis as a competing approach to multiple regression. In keeping with this goal, this paper is organized as follows. First, multiple regression and path analysis approaches are described and contrasted. Next, the specific criteria employed for evaluating path analytic models are outlined. This is followed by a description of the data collection procedure and a discussion of five studies that employ multiple regression models to test theoretical propositions. The next Section depicts and discusses results of our reanalysis. Finally, we provide guidelines for future research and note limitations of the present study.

## 1. Regression analysis

Regression analysis is concerned with predicting the mean value of a dependent variable  $Y$  from known values of one or more independent variables  $X_i$ . The  $p$  variable model with a dependent variable  $Y$ , and  $p$  independent variables  $X_1, X_2, X_3, \dots, X_p$  can be written as:

$$Y_i = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon_i \dots \dots \dots (1)$$

In equation (1),  $\beta_o$  denotes the intercept,  $\beta_1, \dots, \beta_p$  the partial regression (slope) coefficients, and  $\varepsilon_i$  the residual term. Basically,  $\beta_o$  and  $\beta_i$ 's are estimated to be as close to the corresponding population parameters as possible using the method of ordinary least squares (OLS).

In order to appreciate the problems with multiple regression, let us consider the model depicted in Figure 1, which we subsequently reestimate using path analysis. The empirical model in this study represents hypothesized paths of a compensatory process of organizational commitment (Michaels, Cron, Dubinsky, and Joachimsthaler, 1988).

First, to investigate proposed relationships, four independent multiple regression equations have to be specified. The use of independent regressions may, however, bias parameter estimates. Second, indirect effects cannot be determined explicitly. For instance, since role conflict and role ambiguity are highly correlated (0.63), an additional path between these variables can be expected. Third, four different multiple  $R^2$ 's corresponding to the four regression equations need to be estimated. This, in turn, does not permit the calculation of an overall goodness of fit index for the model. In sum, researchers should

account for typical threats that can undermine the robustness of multiple regressions models.

#### Regression model

$$X_2 = \beta_0 + \beta_1 X_1 + \varepsilon_i,$$

$$X_3 = \beta_0 + \beta_1 X_1 + \varepsilon_i,$$

$$X_4 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon_i,$$

$$X_5 = \beta_0 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon_i,$$

Path model estimated

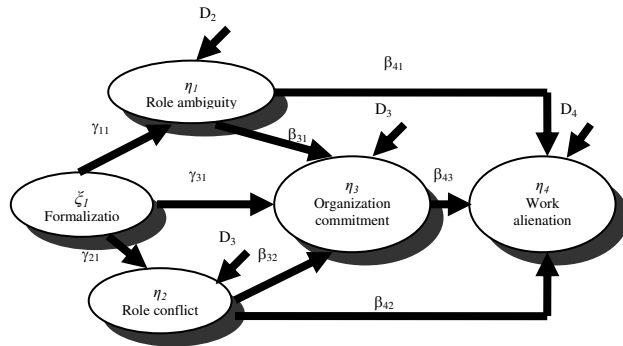


Fig. 1. Path model for reanalysis of Michaels et al. (1988)

In order to tackle the preceding concerns, two approaches within the regression framework have been suggested. First, researchers have tried to address multicollinearity through ridge regression (Mahajan, Jain, and Bergier, 1977; Malthouse, 1999; Zhang and Ibrahim, 2005). Second, in order to identify paths, which are to be added to or deleted from the model, Kang and Senata (1980) recommend the use of residual correlations. Briefly, if the correlation between the residual of a dependent variable and an independent variable is significant, an additional path between the variables may be used. Though, ridge regression and the Kang and Senata (1980) approaches have merit, they are cumbersome to implement and analyze. These techniques neither address the notion of overall fit of the regression model, nor do they tackle the simultaneity concern regarding parameter estimates.

## 2. Path analysis

Path analysis was originally developed by Wright (1921). The starting point for analysis is the summary information about all variables in the model (i.e., the variance-covariance matrix of variables). In the next step, all parameter estimates of the model are estimated simultaneously, and it is possible to compute the discrepancy between the observed correlation matrix and the reproduced correlation matrix. Mathematically, this discrepancy is computed as:  $\Sigma_{(s)} - \Sigma_{(r)}$ , where  $\Sigma_{(s)}$  is the sample correlation matrix and  $\Sigma_{(r)}$  is the reproduced correlation matrix. Multiple criteria are available for computing this discrepancy, i.e., global goodness of fit indices, the distribution of

residuals, and the root mean (or absolute) squared residual (Bentler, 1989, 1990; Ullman and Bentler, 2004). The ability to estimate an overall goodness of fit index represents the main advantage of path analysis over multiple regression. Note that parameter estimates in path analysis are still computed via regression. However, these estimates are computed simultaneously. Path models also allow for the specification of inter-correlations among independent variables, which further aids the computation of unbiased parameter estimates. Moreover, an inspection of the scatter-plot of residuals indicates whether paths have to be added to (or deleted from) the model. In sum, path analysis, which addresses the concerns and limitations of multiple regression is well suited for estimating the relationship between dependent and independent variables.

In the next section, we elaborate upon the path analysis method by outlining the technical criteria for evaluating the output of specific path analysis software programs. Three popular software programs are commonly used to estimate path models i.e., EQS (Bentler, 1989; Bentler and Ullman, 2004; Byrne, 2006), AMOS (Byrne, 2010), and LISREL (Joreskog and Sorbom, 1989). These software programs rely on the same underlying logic and their algorithms produce virtually identical parameter estimates. To a large extent, the use of a particular software is a matter of personal choice. In the present study, we chose EQS for two reasons. First, EQS automatically identifies and imposes bounds on Heywood cases (items with negative error variances). Unidentified Heywood cases, in turn, lead to a number of model estimation problems. Second, the EQS program eschews matrix algebra and is easier to write. An example of an EQS program command file used to reanalyze data in the Michaels et al. (1988) model is depicted in Appendix.

It may be noted that the choice of a particular software is not central to the overall objective of this study since the primary focus is to delineate the superiority of path analysis over multiple regression and not to compare alternative software algorithms.

## 3. Criteria for evaluating path models using EQS

Path analysis output is primarily evaluated by studying goodness of fit between original and reproduced correlations. There are two main approaches to evaluating overall model fit: a) based upon inspecting the residual covariance matrix; and b) assessing overall goodness of fit indices.

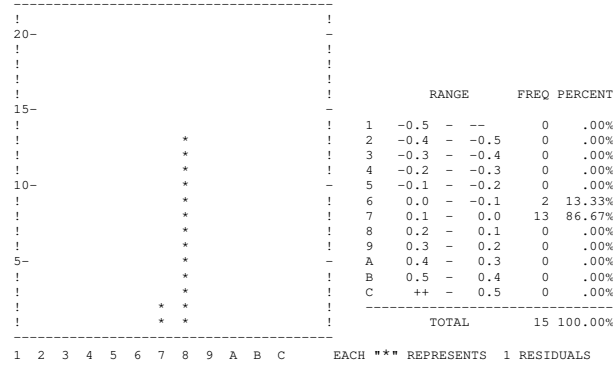


Fig. 2. Distribution of standardized residuals (output from EQS software)

**3.1. Model fit using residual covariance matrix.** A number of specific criteria are used to judge the degree of fit between observed and reproduced covariance matrices. First, the frequency distribution of standardized residuals should be symmetric around zero with small evenly distributed residuals. High residual values suggest correlated errors and poor fit. As an example of good model fit, consider Figure 2 above, which depicts the standardized residual matrix from an EQS output.

Notice that the distribution is symmetrical around zero. Furthermore, 13.33% of residuals lie between 0 and 0.1, while 86.67% of residuals range from 0.1 to 0. Hence, 100% of residuals are close to zero, indicating good degree of fit.

Second, small ( $< 0.05$ ) values of the absolute average of the lower triangular residual matrix and low ( $< 0.05$ ) values of the absolute values of the off-diagonal residual matrix suggest good model fit (Bentler, 1989; Byrne, 2006). Finally, the largest residuals may be inspected to identify variable pairs that distort model fit. For pairs exhibiting high values, structural antecedents or correlates may be modified.

**3.2. Model fit using goodness of fit indices.** To begin with, the  $\chi^2$ -statistic may be inspected to assess degree of fit. The  $\chi^2$ -statistic tests the hypothesis that the model provides a good fit to the data (i.e., the null hypothesis should not be rejected). It may be noted that the  $\chi^2$ -statistic is sensitive to sample sizes greater than 150. Hence, alternative criteria such as the comparative fit index (CFI) should be used to assess model fit. While several fit indices are available, the comparative fit index (CFI) is typically used to assess fit as it takes into account the degree of freedom of the model while avoiding underestimation of fit (Ullman and Bentler, 2004).

Note that significance of parameter estimates is assessed in the normal way, i.e., t-statistic values (computed as the ratio of the parameter estimate to its standard error) in excess of 2 reject the null hypothesis that the parameter estimate is zero.

#### 4. Data collection

To undertake a reanalysis of theoretical models that have used regressions, it is imperative to locate studies that have published the correlation matrix of variables. However, a preliminary inspection of articles appearing in a sample of major marketing and management journals (*Journal of Marketing*, *Journal of Marketing Research*, *Journal of Consumer Research*, *Academy of Management Journal*) revealed that many regression based studies could not be readily reanalyzed because the reported correlation matrix was missing, incomplete, or reported at a level of aggregation that precluded analysis. Hence, a judgment call was made to identify and select a set of studies that: a) were relatively influential as measured by the Social Sciences Citation Index and Google scholar citations; and b) belonged to one of more domains of major managerial decision-making in marketing i.e., product, price, promotion, or place (4P's). Despite emerging debate, there is still general consensus that the 4P's represent an important area of marketing decision-making (Ataman et al., 2010; Narayanan, Desiraju, and Chintagunta, 2004), while researchers are beginning to use Google scholar measures more widely (Harzing and van der Wal, 2008). Based upon our judgmental approach, the following five studies were selected for reanalysis.

**4.1. Description of studies.** The first study (Figure 1) investigates a process model of organizational commitment (Michaels, Cron, Dubinsky, and Joachimsthaler, 1988). Specifically, organizational commitment has three antecedents (role conflict, role ambiguity, and formalization) and one consequence (alienation). With the exception of the formalization-commitment, and the role conflict-alienation paths, all other coefficients are significant. This model was originally tested using multiple regression on a sample of industrial salespeople ( $N = 202$ ).

The second study (Figure 3), deals with the antecedents of inactivity-proneness (IP) among salespeople. It is hypothesized that job satisfaction (JS), perceived image (PI), and performance (P) affect inactivity-

prone to using multiple regression, Wotruba (1990) found significant relations for the JS-IP and the P-IP paths. This study had a sample size of 491.

#### Regression model

$$X_4 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon_i$$

Path model estimated

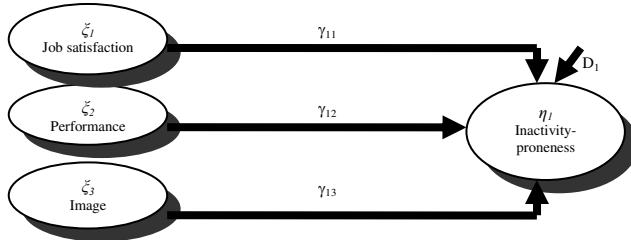


Fig. 3. Path model for reanalysis of Wotruba (1990)

The third study (Figure 4) by Aaker and Keller (1990) investigated the antecedents of extension attitudes. Specifically, quality, transfer, complement, difficulty, and substitute are hypothesized to affect attitudes. Aaker and Keller (1990) found significant relationships only for the transfer-attitude, and the substitute-attitude paths. This regression model was based on a sample of 2140 responses.

#### Regression model

$$X_6 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon_i$$

Path model estimated

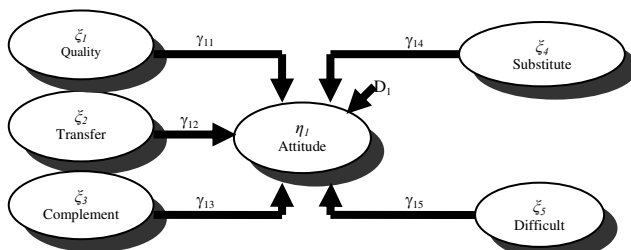


Fig. 4. Path model for reanalysis of Aaker and Keller (1990)

The fourth study (Figure 5) by Heide and Miner (1992) hypothesized that there are seven antecedents of flexibility in the context of a buyer-supplier relationship. Using a regression model, the authors found that only two paths (replace buyer-flexibility, replace supplier-flexibility) were non-significant in a sample of 137 key informants.

Finally, in his regression model, Sethi (2000) studied how cross-functional team characteristics and contextual influences affected product quality (Figure 6). A regression procedure was used to test a series of hypotheses ( $N = 141$ ).

## 5. Results and discussion

As shown in Table 1, regression and path analysis procedures for the Michaels et al. study (1988) produce remarkably congruent estimates for the various paths. Furthermore, the high CFI (0.995), non-significant ( $\chi^2 = 389.9$ ,  $df = 10$ ,  $p = 0.080$ ) and low

value (0.007) of the average off diagonal standardized residual (AOSR) suggest that the model fits the data well. It seems that the path model offers no additional improvement over the multiple regression model. It is possible that in this case data do not violate any of the assumptions of a regression model. Specifically, multicollinearity or measurement error are minimal. This is a special case, where path analysis and multiple regression produce identical results. Ideally, if data were well behaved, regression models and path models would yield the same result.

#### Regression model

$$X_8 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \varepsilon_i$$

Path model estimated

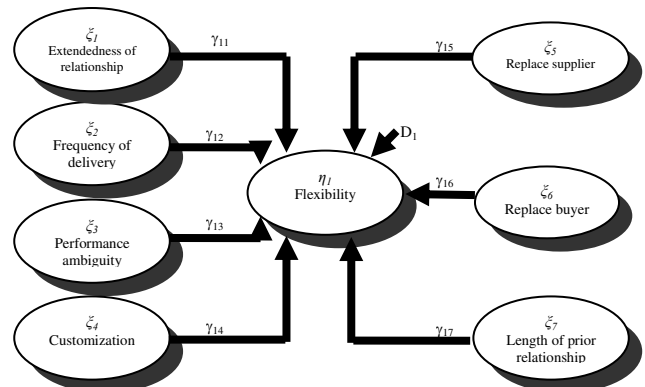


Fig. 5. Path model for reanalysis of Heide and Miner (1992)

#### Regression model

$$X_7 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \varepsilon_i$$

Path model as estimated

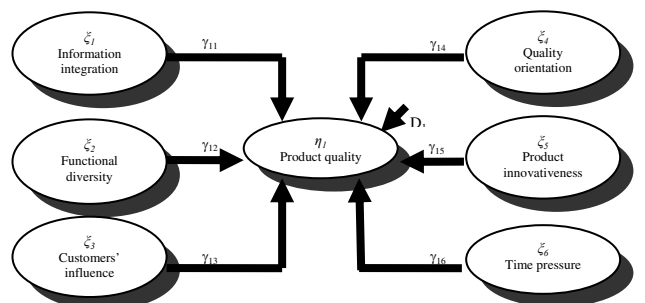


Fig. 6. Path model for reanalysis of Sethi (2000)

Table 1. Reanalysis results for Michaels et al. (1988)

| Path          | Regression coefficient | Path analysis coefficient |
|---------------|------------------------|---------------------------|
| $\gamma_{11}$ | -.56*                  | -.96*                     |
| $\gamma_{21}$ | -.25*                  | -.60*                     |
| $\gamma_{31}$ | .13                    | .32                       |
| $\beta_{31}$  | -.19*                  | -.28*                     |
| $\beta_{32}$  | -.33*                  | -.33*                     |
| $\beta_{41}$  | .21*                   | .12*                      |
| $\beta_{42}$  | .07                    | .02                       |
| $\beta_{43}$  | -.47*                  | -.19*                     |

Notes: AOSR<sup>a</sup> = .007,  $\chi^2 = 389.909$ ,  $df = 10$ ,  $p = .080$ , CFI<sup>b</sup> = .995.

Source: \*  $p \leq .05$ ; <sup>a</sup>average off-diagonal squared residuals; <sup>b</sup>comparative fit index.

A reanalysis of the Wotruba (1990) study via path analysis provides good evidence of model fit (Table 2). The CFI (0.999), AOSR (0.025) and non-significant ( $\chi^2 = 575.61$ ,  $df = 6$ ,  $p = 0.184$ ) suggest that data fits the model well. The path analysis also suggests stronger significant paths between performance and inactivity proneness. Surprisingly, in contrast to the regression model, the relationship between image and inactivity proneness is non-significant. The discrepancy between the two statistical approaches is perhaps caused by multicollinearity among independent variables which could be explicitly modeled in path analysis.

Table 2. Reanalysis results for Wotruba (1990)

| Path          | Regression coefficient | Path analysis coefficient |
|---------------|------------------------|---------------------------|
| $\gamma_{11}$ | -.74 <sup>*</sup>      | -.71 <sup>*</sup>         |
| $\gamma_{12}$ | -.12 <sup>**</sup>     | -.07 <sup>*</sup>         |
| $\gamma_{13}$ | -.44 <sup>*</sup>      | -.04                      |

Notes: AOSR<sup>a</sup> = .025,  $\chi^2 = 575.611$ ,  $df = 6$ ,  $p = .184$ , CFI<sup>b</sup> = .999.  
Source: \*  $p \leq .05$ ; \*\*  $p \leq .01$ ; <sup>a</sup>average off-diagonal squared residuals; <sup>b</sup>comparative fit index.

Results of the Aaker and Keller (1990) reanalysis are depicted in Table 3. In particular, path analysis uncovers three additional significant paths (quality-attitude, substitute-attitude, and complement-attitude). Furthermore, the high CFI value (0.9) and low AOSR (0.048) indicates good model fit. Note that while the  $\chi^2$ -statistic is significant (1433.8,  $df = 15$ ,  $p < 0.000$ ) suggesting poor model fit, the statistic is not very reliable when sample sizes are large.

Table 3. Reanalysis results for Aaker and Keller (1990)

| Path          | Regression analysis coefficient | Path analysis coefficient |
|---------------|---------------------------------|---------------------------|
| $\gamma_{11}$ | -.01                            | .23 <sup>*</sup>          |
| $\gamma_{12}$ | .12 <sup>*</sup>                | .25 <sup>*</sup>          |
| $\gamma_{13}$ | -.02                            | .16 <sup>*</sup>          |
| $\gamma_{14}$ | -.06                            | .09 <sup>*</sup>          |
| $\gamma_{15}$ | .12 <sup>*</sup>                | .13 <sup>*</sup>          |

Notes: AOSR<sup>a</sup> = .048,  $\chi^2 = 1433.803$ ,  $df = 15$ ,  $p < .0001$ , CFI<sup>b</sup> = .901.  
Source: \*  $p \leq .05$ ; <sup>a</sup>average off-diagonal squared residuals; <sup>b</sup>comparative fit index.

Reanalysis of the Heide and Miner (1992) study indicates that in contrast to their regression estimates, the path between flexibility and length of prior relationship becomes insignificant as a result of path analysis. One is led to believe that the multiple regression model produced biased parameter estimates because distributional assumptions may have been violated. Path analysis estimates for the Heide and Miner study are depicted in Table 4.

Table 4. Reanalysis results for Heide and Miner (1992)

| Path          | Regression analysis coefficient | Path analysis coefficient |
|---------------|---------------------------------|---------------------------|
| $\gamma_{11}$ | .56 <sup>*</sup>                | .51 <sup>*</sup>          |
| $\gamma_{12}$ | .01                             | .01                       |
| $\gamma_{13}$ | -.08                            | .02                       |
| $\gamma_{14}$ | .07 <sup>**</sup>               | .09 <sup>*</sup>          |
| $\gamma_{15}$ | -.004                           | -.01                      |
| $\gamma_{16}$ | -.04 <sup>*</sup>               | -.05 <sup>*</sup>         |
| $\gamma_{17}$ | -.24 <sup>*</sup>               | .03                       |

Notes: AOSR<sup>a</sup> = .048,  $\chi^2 = 147.100$ ,  $df = 28$ ,  $p = .170$ , CFI<sup>b</sup> = .962.  
Source: \*  $p \leq .05$ ; \*\*  $p \leq .01$ ; <sup>a</sup>average off-diagonal squared residuals; <sup>b</sup>comparative fit index.

Finally, as shown in Table 5, path analysis reveals that in contrast to the Sethi (2000) results, customers' influence on product quality is no longer statistically significant. Note that the estimated path analytic model is robust and provides a good degree of fit {CFI = .992, AOSR = .063, ( $\chi^2 = 64.896$ ,  $df = 21$ ,  $p = .425$ )}.

Table 5. Reanalysis results for Sethi (2000)

| Path          | Regression analysis coefficient | Path analysis coefficient |
|---------------|---------------------------------|---------------------------|
| $\gamma_{11}$ | .15 <sup>*</sup>                | .15 <sup>*</sup>          |
| $\gamma_{12}$ | .02                             | -.01                      |
| $\gamma_{13}$ | .18 <sup>*</sup>                | .07                       |
| $\gamma_{14}$ | .25 <sup>*</sup>                | .24 <sup>*</sup>          |
| $\gamma_{15}$ | -.48 <sup>*</sup>               | -.18 <sup>*</sup>         |
| $\gamma_{16}$ | -.01                            | -.01                      |

Notes: AOSR<sup>a</sup> = .063,  $\chi^2 = 64.896$ ,  $df = 21$ ,  $p = .425$ , CFI<sup>b</sup> = .992.  
Source: \*  $p \leq .05$ ; <sup>a</sup>average off-diagonal squared residuals; <sup>b</sup>comparative fit index.

## Conclusions

The main objective of this study has been to position path analysis as a competing measurement approach to multiple regression. The results indicate that path analysis may be gainfully employed to investigate theoretical relationships among variables when multicollinearity and non-simultaneous estimation of parameters are present. Specifically, in four of the five studies, the original regression estimates underwent significant changes. Furthermore, additional significant paths were uncovered in two studies, while in two models previously significant paths became non-significant as a result of reanalysis. Notice that to maintain our focus on the technical aspects of path analysis and regressions, we refrain from an in-depth discussion about each specific theoretical construct in the estimated model and how changes in a particular empirical estimate might affect theory development. However, given that theory testing is an important goal of science, researchers should choose the most appropriate scientific method to advance theory.

Note that we do not advocate path analysis to the exclusion of regression. On the other hand, we

behoove upon researches to pay greater attention to path models in future. This point is important because marketing relies very heavily on perceptual data. More often than not, perceptual data collected in a particular theoretical context is expected to be correlated. Given the ability of path analysis to explicitly model correlations and measurement error, marketers should therefore use this method more often.

Though, this study has been exploratory, the results provide researches with a forum for further discussion about competing measurement approaches in marketing. Specifically, the recent debate in marketing about measurement issues concern the limitations of existing regression based approaches (Grapentine, 2000; Iacobucci, Saldhana, and Deng, 2007; Zhao, Lynch, and Chen, 2010). The main criticism leveled at regression is the non-simultaneous estimation of parameters especially while conducting mediation and intervening variable analysis. To address simultaneity concerns, researchers suggest the superiority of path analytic and structural equations modeling because they “estimate everything simultaneously instead of assuming that equations are independent” (Zhao, Lynch, and Chen, 2010, p. 205). In this vein, our discussion of path analytic approaches provides researchers with a statistical approach that allows for simultaneous estimates of parameters.

Finally, the reanalysis approach illustrated in this paper is broadly congruent with the goals of replication and reinquiry research (Wilk, 2001). Specifically, if the goal of science is to cumulate findings via theory

testing, it behooves upon researchers to employ appropriate statistical methods for testing construct interrelationships. By outlining the relative superiority of path analysis, our study stresses the importance of employing appropriate testing approaches as a first step toward cumulating findings.

The results of this study have to be viewed against certain limitations. First, to keep the study tractable, only data from five marketing studies are reanalyzed. Through it is not possible to generalize the findings with certainty, this exercise represents an initial attempt to compare two competing techniques for studying the relationship among dependent and independent variables. Second, we could not completely assess the effects of non-normality on the results because published correlation matrices were analyzed. An ideal comparative approach should make use of raw data. By using raw data, one can detect outlying observations and make better theoretical predictions from the model. Finally, the reported correlation matrices represent averages construct scores as opposed to intercorrelations among measures. To this extent, we could not explicitly control for measurement error.

### Acknowledgements

The study was partially funded by a PDQWL study grant from the State University at Binghamton. We acknowledge the helpful comments provided by seminar participants at the State University of New York and the Melbourne Business School.

### References

1. Aaker, D.A., K.L. Keller. Consumer Evaluation of Brand Extensions // *Journal of Marketing*, 1990. – No. 54 (1). – pp. 27-41.
2. Ataman, M. B., H. J. Van Heerde, C. F. Mela. The Long-Term Effect of Marketing Strategy. // *Journal of Marketing Research*, 2010. – No XLVII (5).
3. Baron, R.M., D.A. Kenny. Moderator-mediator Variables Distinction in Social psychological Research: Conceptual, Strategic, and Strategic Considerations//*Journal of Personality and Social Psychology*, 1986. – No. 51 (6). – pp. 1173-82.
4. Bentler, P.M. EQS: Structural Equations Program Manual – Los Angeles CA: BMDP Statistical Software, (1989).
5. Bentler, P.M. Comparative Fit Indices in Structural Models // *Psychological Bulletin*, 1990. – No. 107. – pp. 238-246.
6. Byrne, B.M. Structural Equations Modeling with EQS. – New York: Lawrence Erlbaum, 2006.
7. Byrne, B.M. Structural Equations Modeling with AMOS. – New York: Routledge, 2010.
8. Cote, J.A., R. Buckley. Estimating Trait, Method, and Error Variance: Generalizing Across 70 Construct Validation Studies // *Journal of Marketing Research*, 1987. – No. 27 (August), – pp. 315-8.
9. Cote, J.A., R. Buckley. Measurement Error and Theory testing in Consumer Research: an Illustration of the Importance of Construct Validation. // *Journal of Consumer Research*, 1988. – No. 14 (March). – pp. 579-582.
10. Echambadi, R., J.D. Hess. Mean-Centering Does Not Alleviate Collinearity Problems in Multiple Regression Models // *Marketing Science*, 2007. – No. 26 (3). – pp. 438-445.
11. Friedman, L., M. Wall. Graphical Views of Suppression and Multicollinearity in Multiple Linear Regression // *American Statistician*, 2005. – No. 59 (2). – pp. 127-136.
12. Fitzsimons, G. J. Death to Dichotomizing // *Journal of Consumer Research*, 2008. – No. 35 (June). – pp. 5-8.
13. Grapentine, T. Path Analysis vs. Structural Equation Modeling // *Marketing Research*, 2000. – No. 12 (3). – pp. 13-20.
14. Grewal, R., J.A. Cote. H. Baumgartner. Multicollinearity and Measurement Error in Structural Equation Models: Implications for Theory Testing // *Marketing Science*, 2004. – No. 23 (4). – pp. 519-529.
15. Guttman, L. Measurement as Structural Theory // *Psychometrika*, 1971 – No. 36 (4). – pp.329-347.

16. Hagerty, M.R., V. Srinivasan. Comparing the Predictive Powers of Alternate Multiple Regression Models // Psychometrika, 1991. – No. 56 (1). – pp. 77-85.
17. Harzing, A.K., R. Wal. Google Scholar as a New Source for Citation Analysis//Ethics in Science and Environmental Politics, 2008. – No. 8. – pp. 1-13.
18. Heide, J.B., A.S. Miner. The shadow of the Future: Effects of Anticipated Interaction and Frequency of Contact on Buyer-Seller Cooperation // Academy of Management Journal, 1992. – No.35 (2). – pp. 265-91.
19. Iacobucci, D. Mediation Analysis. – Thousand Oaks, CA: Sage, 2008.
20. Iacobucci, D., N. Saldhana, X. Deng. A Meditation on Mediation: Evidence that Structural Equations Models Perform Better Than Regressions // Journal of Consumer Psychology, 2007. – No. 17 (2). – pp. 139-53.
21. Irwin, J.R., G.H. McClelland. Misleading Heuristics and Moderated Multiple Regression Models // Journal of Marketing Research, 2001. – No. 38 (February). – pp. 100-109.
22. Joreskog, K.G., D. Sorbom. LISREL 7: A Guide to the Program and Applications. – Chicago, IL: SPSS, Inc, 1989.
23. Judd, C.M., D.A. Kenny. Data Analysis in Social Psychology: Recent and Recurring Issues. In the Handbook of Social Psychology, 5<sup>th</sup> ed., ed. Daniel Gilbert, Susan T. Fiske, and Gardiner Lindzey. – New York: McGraw-Hill, 2010. – pp. 233-65.
24. Kang, K.M., D. Senata. Path Analysis: an Explication. in Development in Statistics, (3), Ch. 4, P.K. Krishnaiah, ed. – New York: Academic Press, Inc., 1980. – pp. 217-246.
25. Malthouse, E.C. Ridge Regression and Direct Marketing Scoring Models // Journal of Interactive Marketing, 1999. – No. 13 (4). – pp. 10-23.
26. Mahajan, V., A.K. Jain, M. Bergier. Parameter Estimation in Marketing Models in the Presence of Multicollinearity: an Application of Ridge Regression // Journal of Marketing Research, 1977. – No. 14 (4). – pp. 586-591.
27. Michaels, R.E., W.L. Cron., A.J. Dubinsky, E.A. Joachimsthaler. Influence of Formalization on the Organizational Commitment and Work Alienation of Salespeople and Industrial Buyers // Journal of Marketing Research, 1988. – No. 25 (4) – pp. 376-83.
28. Mishra, D.P. An Empirical Assessment of Measurement Error in Health-Care Survey Research // Journal of Business Research, 2000. – No. 48. – pp. 193-205.
29. Nachtigall, C., U. Kroehne, F. Funke, R. Steyer. (Why) Should We Use SEM?: Pros and Cons of Structural Equation Modeling // Methods of Psychological Research Online, 2003. – No. 8 (2). – pp. 1-22.
30. Narayanan, S., R. Desiraju, P. Chintagunta. Return on Investment Expenditures for Pharmaceutical Expenditures: the Role of Marketing Mix Interactions // Journal of Marketing, 2004. – No. 68 (October). – pp. 90-105.
31. Rindfleisch, A., A.J. Malter, S. Ganesan., C. Moorman. Cross-Sectional versus Longitudinal Survey Research: Concepts, Findings, and Guidelines // Journal of Marketing Research, 2008. – Vol. XLV (June), – pp. 261-279.
32. Sethi, R. New Product Quality and Product Development Teams // Journal of Marketing, 2000. – No. 64 (2). – pp. 1-14.
33. Ullman, J.B., P.M. Bentler. Structural Equations Modeling in Handbook of Data Analysis, M. Hardy, and A. Bryman (Eds.). – Calif: Sage, 2004.
34. Wilk, R.W. The Impossibility and Necessity of Re-Inquiry: Finding Middle Ground in Social Science // Journal of Consumer Research, 2001. – 28 (September). – pp. 308-312.
35. Wotruba, T. The Relationship of Job Image, Performance, and Job Satisfaction to Inactivity-Proneness of Direct Salespeople // Journal of the Academy of Marketing Science, 1990. – No. 18 (2). – pp. 113-121.
36. Wright, S. Correlation and Causation // Journal of Agricultural Research, 1921. – No. 20 (June), – pp. 557-85.
37. Zhang, J., M. Ibrahim. A Simulation Study on SPSS Ridge Regression and Ordinary Least Squares Regression Procedures for Multicollinearity Data // Journal of Applied Statistics, 2005. – No. 32 (6). – pp. 571-588.
38. Zhao, X., J.G. Lynch Jr., Q. Chen. Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis // Journal of Consumer Research, 2010. – No. 37 (August). – pp. 197-206.

## Appendix

### Sample EQS Program Command Lines

**Title:** Path analysis example based on correlation matrix of Michaels et al. "Influence of formalization on ORGN commitment and work alienation".

**Specifications:** CAS=215; VAR=5; ME=GLS.

**Labels:** V1=FORMALIZ; V2=ROLEAMB; V3=ROLECON; V4=COMMIT; V5=ALIEN.

**Equations:** V2=\*V1+E2; V3=\*V1+E3; V4=\*V1+ \*V2 + \*V3 + E4; V5=\*V2 + \*V4 + \*V3 + E5.

**Variances:** E2 to E5=.

**Matrix:**

1.  
 -.57; 1.  
 -.25; .63; 1.  
 .32; -.47; -.48; 1.  
 -.36; .47; .42; -.61; 1.

**Standard deviations:** 5.3; 8.8; 12.8; 12.9; 5.1.

**Means:** 38; 21.7; 38.9; 82.4; 11.6.

**End.**