# Avoiding spurious moderation effects:
# An information-theoretic approach to moderation analysis

**Ahmad Daryanto**

Lancaster University Management School, Lancaster University, Lancaster LA1 4YX
Lancaster, United Kingdom
Email: a.daryanto@lancaster.ac.uk

1

# Avoiding spurious moderation effects:

# An information-theoretic approach to moderation analysis

**Abstract** Researchers typically use moderated regression models to examine the presence of linear moderation effects in their studies. However, researchers rarely conduct a robustness check following a significant moderation effect to investigate whether the moderation effect is spurious. The misleading moderation can occur when a predictor and a moderator variable correlate and the true nature of the relationships between predictors and a dependent variable are nonlinear. In this paper, we propose and illustrate the use of an information theoretic approach in moderation analysis with the aim of avoiding spurious moderation effects. We demonstrate our suggested procedure using Monte Carlo simulations and real data from published studies.

# 1   Introduction

The use of moderation analysis is ubiquitous in business, management, and marketing research, not least due to the accessibility of publications on the conceptual foundations of the involved statistical procedures (e.g., Aiken et al. 1991) and availability of new, user-friendly statistical computer programs (e.g.,Hayes 2012). Despite the popularity of moderation analysis, however, hardly any papers investigate the possibility that the significant moderation effect reported in those papers is potentially misleading or spurious. The moderation effect can be spurious if there is a nonlinear relationship (i.e., typically captured in empirical research by a quadratic term) between a dependent variable and an independent variable (Lubinski and Humphreys 1990; MacCallum and Mar 1995; Ganzach 1997) leading to a Type I error.  If $X$ and $Z$ are two main variables of interest, a moderation effect is captured by the cross-product of $X$ and $Z$, commonly termed as an interaction term. Testing a moderation effect is typically undertaken by estimating the following moderated multiple regression model:

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ + error \tag{1}$$

The coefficient, $\beta_3$ in the above equation represents a moderation effect. $\beta_3$ can be found to be significant in a sample even when there is no true moderation in the population, termed as spurious moderation.  This situation can occur due to the correlation between $X$ and $Z$. If $X$ and $Z$ are correlated, then $XZ$ and the quadratic terms $X^2$ and $Z^2$ are also correlated, which results in an overlap between the variance explained by the interaction term, $XZ$ and the variance explained by the two quadratic terms, $X^2$ and $Z^2$ (Ganzach, 1997).

There have been mixed recommendations on the statistical procedures that can be used to detect the presence of quadratic effects for avoiding spurious moderation effects. For instance,  Lubinski and Humphreys (1990) suggested that researchers must compare the

effect sizes (i.e., using R$^2$) associated with moderation and quadratic effects and select a model that yield the strongest effects. Others had suggested that researchers should enter quadratic terms in their moderated regression model and tested the inclusion of the curvilinear terms using either a join F-test or T-tests (Cortina 1993; Ganzach 1997). For the later tests, if the F or T-tests accept the null hypothesis of no quadratic effects, then the significant moderation effect is unlikely to be spurious. MacCallum and Mar (1995) found that the result of these tests can be affected by measurement errors in *X* and Z, which might not always lead to a correct model selection. Lance (1988) introduced a residual centering approach for testing moderation effects with the main purpose of reducing the multicollinearity problems between *XZ* and its components, *X* and *Z*. Lance (1988)'s method requires the use of the residuals from a regression of XZ on its component terms, *X* and *Z* in place of the *XZ* in Equation (1) above termed as the orthogonalized version of *XZ*. Common to these methods is that researchers do not need to compare the model in Equation (1) to other plausible models that might fit the data, with the exception that the model in Equation (1) is usually compared with a model with only the main terms in the equation (i.e., $Y = \beta_0 + \beta_1 X + \beta_2 Z + error$) when one uses a hierarchical regression analysis.

Despite these various recommendations, more than two decades ago, MacCallum and Mar called for researchers' attention to "…pursue the development of better methods for … distinguishing between multiplicative and quadratic models" (MacCallum and Mar, 1995, p.418). The present paper responds to this long-forgotten call. In the present research, We introduce and suggest a relatively new method using an information theoretical approach that can help researchers in preventing misleading interaction due to curvilinear effects when detecting the presence of a true moderation effect. This paper focuses on moderated regression models as depicted in Equation (1) with continuous moderators because quadratic terms for dichotomous variables are trivial.

In this paper, as our focus is on preventing or controlling for the presence of curvilinear effects, we propose that researchers can apply our suggested information-approach as a robustness check for a significant moderation effect that is found using Equation (1). For simplicity, throughout this paper, we focus on testing one moderation effect with a model specification represented with Equation (1). We call our suggested modelling endeavor a two-step modelling approach. In the first step of the procedure, researchers formulate their theory-driven moderation effect hypothesis and test the effect using a moderated regression model as depicted in Equation (1). The significance of each coefficient in the model is assessed using the traditional p-value hypothesis testing. If the interaction term in Equation (1) is significant, researchers can proceed to test the quality of their model by conducting the second step, otherwise researchers should revisit their theory about the moderation effect.

In the second step, researchers evaluate the quality of their moderated regression models against other competing models that are plausible given the data. The key feature of the second step is that researchers are working with a set of multiple (plausible) models with the model of Equation (1) being included in the set. The second step requires researchers to use an information-theoretic approach to make an inference regarding the quality of each model in the plausible set. In this step, using statistical criteria commonly used in information-theoretic approaches, researchers can assess whether their moderated regression models are either the best among the plausible models or have considerable support or otherwise, receive no support at all. As the information-theoretic approach dictates (Burnham and Anderson, 2003), the key decision that should be undertaken using the approach is how to select models that are considered to be plausible. Fortunately, for moderated regression analysis, the selection of models that can be included in the set of plausible models can be

formulated by varying the slope coefficient of a linear additive model, namely Equation (1) without the interaction term, i.e., $Y = \beta_0 + \beta_1 X + \beta_2 Z + error$.

We organize our paper as follows. In the next section, using Monte Carlo simulations, We demonstrate the issue of spurious moderation effects due to the presence of nonlinearity between an independent variable and a dependent variable. In the next section, we explain the logic of the information-theoretic approach. Next, we explain how to construct a set of plausible or candidate models required in the information-theoretic approach. In the next section, using two Monte Carlo studies, we illustrate the use of an information-approach in testing moderation effects. Note that in typical Monte Carlo studies, the functional form of the true model used to generate data is known. Using known functional forms does not reflect real research situations. Furthermore, using a pure mathematical model specification without any links to a particular research context lacks research contextualization. Therefore, we provided a demonstration of our suggested procedure using real data from two published studies. Finally, we end our paper with a discussion and conclusion section.

## 2 Demonstrating spurious moderation effects

Using Monte Carlo simulations, we demonstrate the occurrence of spuriousness in a miss-specified moderation model when the true model contains a quadratic effect. In our Monte Carlo simulations, we vary two population parameters: the correlation between $X$ and $Z$ (using 0, .1, .3, .5, and .7) used to generate bivariate normal samples, and the sample size (using $N =$ 75, 300, 500, 1000). We chose these sample sizes to represent a typical sample size range encountered in marketing and business research. We also introduce measurement errors in the $X$ and $Z$ variables according to this formula: observed score=True score + measurement error (MacCallum and Mar 1995), where the array of the true scores is generated from a given correlation above. The measurement errors are sampled from a normal distribution with a mean of zero. The variance of the error for each variable was calculated as error variance= (1-

6

reliability)/reliability. We vary two reliability coefficients in our study using 0.4 and 0.8 representing low and high reliability.

We estimated a moderated regression model, which includes linear and interaction terms when the true population model contains a linear and quadratic term. We should emphasize here that there is a difference between our population model and that of MacCallum and Mar (1995)'s study, which is that in our model We only use a quadratic term from one of the components of the interaction term, so that the interaction term might pick up the left-out variable effect when the correlation between $X$ and $Z$ is not equal to zero (Balli and Sørensen 2013). The use of the quadratic term related to only one component of the interaction term represents a simple and pure form of model misspecification.

In total, there are 20 simulation scenarios for each value of a reliability coefficient. In each of the scenarios, we generated 10,000 replication samples, where values for $Y$'s as a dependent variable are calculated according to a quadratic regression model as $Y = X + X^2 + \varepsilon, \varepsilon \sim N(0,1)$. Next, using the simulated values for $X$, $Z$, and $Y$, we estimated a miss-specified moderated regression model as $Y = X + Z + XZ + \varepsilon, \varepsilon \sim N(0,1)$. In each simulation scenario, we calculated the probability of obtaining a significant interaction effect in the miss-specified model. We plotted the probability of obtaining a significant interaction against different sample sizes for various values of the correlation between $X$ and $Z$. The results of the simulation are presented in Panel A and Panel B of Figure 1. The two graphs of Figure 1 show the effect of model miss-specification on the probability of Type I error measured as the proportion of p-value associated with the interaction term that falls below .05. Panel A and Panel B show results of simulation for reliability coefficients of 0.4 and 0.8, respectively. As can be seen from both graphs, even with a modest sample size of 100 and moderate correlation, the probability of Type I error is already high (i.e., greater than .7).

[INSERT FIGURE 1 HERE]

# 3    An information-theoretic approach

The information-theoretic approach is a relatively new paradigm in inference (Burnham and Anderson 2003), which is based on the following principles. First, there is no true model in sciences and there are multiple models that one can develop to estimate the true reality—in this paper, we use the term plausible and candidate models interchangeably. Hence, in using an information-theoretic approach, one is not examining one model only but considering multiple models that are deemed plausible. This is in contrast with the traditional p-value testing approach, which requires only one model to be assessed. Second, the quality of each candidate model can be assessed, which means that the candidate models can be compared among each other. The concept of model quality in an information-theoretic approach is related to the amount of information contained in the model. When a candidate model is used to estimate the true model, there would be information loss. As a result of the assessment, the models can be ranked from best to worst, with the best model being the one that provides the best approximation to the true model. The assessment of the quality of each candidate model and the selection of the best to worst among the candidate models is typically guided by the Akaike information criterion (AIC) method and Distance and Akaike Weights, which are other measures derived from AIC values. We explain these measures below.

**The uses of AIC, Distance and Akaike Weights**

In an information-theoretic approach, the Akaike information criterion (AIC) is a common measure used to evaluate the quality of each candidate model. The AIC of a model is given in the following form (Akaike 1974):

$$\text{AIC} = -2\log L\left(\hat{\theta}\right) + 2K$$

where $L(\hat{\theta})$ is the maximized likelihood function given the data and the model, and $K$ is the number of free parameters in the model. The candidate model with minimum AIC value is often selected as the best model to fit the data. However, if all the candidate models are poor, the model with minimum AIC value is still the best model among the candidate models. Thus, it is very important that the candidate models are carefully developed.

The AIC value of a model can be rescaled with the following transformation: $D =$ AIC – minimum AIC in the set, where AIC is the AIC value of a model and $D$ is the difference in AIC of the model with respect to the AIC of the best candidate model—a model with the smallest AIC in the set. The $D$ value can be interpreted as the estimates of the distance between a model and the best model (Burnham and Anderson, 2003). As a result of this transformation, the $D$ value of the model with the minimum AIC will be zero. In general, the larger the $D$ value of a model, the less plausible it is that the model is the best approximating model. The advantage of using the $D$ value is that it allows for model comparison such that one can rank the candidate models from best to worst. The simple rules of thumb have been given in the literature that are useful in assessing the relative quality of candidate models: candidate models having $D \leq 2$ are perceived to have substantial support, $4 \leq D \leq 7$ have considerably less support, and while models having $D > 10$ have essentially no support (Burnham and Anderson, 2003).

The coefficient $D$ of a model can also be transformed using the exponential transformation to exp ($-D/2$), which can be interpreted as the likelihood of the model given the data. Further, these transformations can be normalized such that they sum to 1:

$$w_i = \frac{\exp(-D_i/2)}{\sum_{r=1}^{R}(-D_r/2)}$$

where $R$ is the number of candidate models, $D_i$ is the distance of model $i$, and $w_i$ is called the Akaike weight of model $i$. The Akaike weight, $w_i$, can be interpreted as the probability that

model *i* is, in fact, the best approximating model to the data. Note that the Akaike weight is a function of the Ds of all models in a given set, therefore, if one adds or drops a model from the plausible set, then the Akaike weights must be recalculated for all models in the newly created set.

As mentioned above, the key decision that one has to make in using the information-theoretic approach is in selecting what models to include in the candidate set. The development of model specifications that should be included in the candidate set should be considered carefully. Candidate models can be developed from various sources. First, candidate models can be derived by considering modelling situations when hypothesized hypotheses are not supported. For instance, if moderation effect is hypothesized, models with main effects only are candidate models. Second, researching substantive literature on the domain of the scientific enquiry can also lead to different model specifications. Third, models reported in previous research in research domains of interest can also be considered as candidate models. Fourth, candidate models can be developed by modifying the hypothesized model specifications. Fortunately, for moderated regression, the choice of models to be included in the plausible set can be possibly derived from varying the slope coefficient of the linear additive model. In the next section, we explain how to derive a set of candidate models given that the focus of one's research is to test a model specification of Equation (1).

It is important to note that another criterion frequently mentioned in the literature of model selection is the Bayesian Information Criterion (BIC). In this paragraph, we want to highlight why in this paper, we use AIC but not BIC. In this paper, we propose the use of the multi-model inference that is based on the information-theoretic approach for model selection. The information-theoretic approach is based on the K-L (Kuhlback-Leibler) information. AIC is founded on information theory and defined as an expected estimate of K-L information (Burnham and Anderson 2004). In contrast to AIC, BIC is not based on K-L

information, thus, should not be considered as an information-criteria within the information-theoretic paradigm. BIC is a Bayesian model selection approach, which is based on Bayes factor and often considered as a dimension-consistent criterion (Burnham and Anderson, 2004, p.263).

Another difference between AIC and BIC lies on the nature of the modelling philosophy regarding the existence of a true model. AIC assumes that there is no true model; that is a model is just an approximation to reality. Thus, the true model need not be present in the set of the plausible model. In contrast, BIC assumes that there is a true model among the candidate models. We would like to quote Burnham and Anderson (2003) here: "We recommend the class of information-theoretic criteria that are estimates of relative K-L information such as AIC, $AIC_C$[1] for general use in the selection of a parsimonious approximating model for statistical inference for sample sizes that occur in practice" (p.289).

Another AIC-like criterion that is quite popular in practice is the consistent AIC (CAIC). Anderson et al. (1998) compared the performance of AIC vs CAIC using Monte Carlo studies. They found that AIC performs better than CAIC where CAIC has a tendency to select underfitting models. To this end, Anderson et al. (1998) concluded that "CAIC-selected models were too simple, the parameter estimators were often substantially biased, the asymptotic sampling variances were substantially too small, and the achieved coverage was often substantially below the nominal level" *(p. 263).*

## 4   Deriving Plausible Models

The familiar form of the basic moderated regression model is given by Equation (1) presented in the introduction section. The interaction term in Equation (1), *XZ*, captures the

---

[1] $AIC_C$ is AIC with correction for a large number of model parameters (*K*) and small sample sizes.

moderation effect. The origin of this model comes from expressing the slope coefficient corresponding to $X$ as a function of $Z$ in a linear model as in the following additive model (Aiken et al. 1991):

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \varepsilon \tag{2}$$

That is, from this additive model, moderation can be expressed as:

$$\beta_1 = \alpha + \delta Z \tag{3}$$

Substituting Equation (3) into Equation (2), we will obtain a model with a similar form as in Equation (1):

Now, we show that a quadratic term can enter into Equation (1) if $\beta_1$ contains a quadratic term. That is, we can write,

$$\beta_1 = \alpha + \delta Z + \pi Z^2 \tag{4}$$

Substituting Equation (4), into Equation (2), we obtain a model that contains a linear, quadratic, and interaction term in the following form.

$$Y = \beta_0 + \alpha X + \beta_2 Z + \delta XZ + \pi XZ^2 + \varepsilon \tag{5}$$

Similarly, we can express the slope coefficient corresponding to $Z$ as a linear and quadratic function of $X$.

$$\beta_1 = \alpha + \delta X + \pi X^2 \tag{6}$$

We obtain the following equation:

$$Y = \beta_0 + \alpha X + \beta_2 Z + \delta XZ + \pi ZX^2 + \varepsilon \tag{7}$$

The above derivations show that expressing the slope coefficient as a linear and quadratic function of either a predictor or moderator, we can obtain an equation that is more general than Equation (1). That is, the two interaction terms in Equation (5) and Equation (7) are byproducts of combining two equations: Equation (4) with Equation (1) and Equation (6) with Equation (1). Thus, not only Equation (1), but both Equation (5) and Equation (7) should be considered as plausible models with respect to testing moderation effects.

12

Other plausible models can also be derived by expressing $X$ as a linear function of $Z$, which captures a correlation between the two variables.

$$X = \alpha + \gamma Z + \tau \qquad (8a)$$

Where $\tau$ is an error term. Substituting $X$ in Equation (1) with the form in Equation (8a), we obtain a purely quadratic model in $X$ given by the following Equation:

$$Y = \beta_0' + \beta_1'X + \beta_2'X^2 + \varepsilon \qquad (9)$$

Similarly, if we write

$$Z = \alpha + \gamma X + \tau \qquad (8b)$$

We obtain:

$$Y = \beta_0' + \beta_1'Z + \beta_2'Z^2 + \varepsilon \qquad (10)$$

Thus, Equation (9) and Equation (10) are two other candidate models that should be included in the set of plausible models.

As explained in the previous section, past researchers have recommended that quadratic terms should be entered into Equation (1) as a way to control for curvilinear effects (Lubinski and Humphreys 1990; Cortina 1993; Ganzach 1997). Specifically, they suggested that the model that should be tested is a model that includes a linear, quadratic and interaction term as in the following equation.

$$Y = \beta_0 + \beta_1X + \beta_2Z + \beta_3X^2 + \beta_4Z^2 + \beta_5XZ + \varepsilon \qquad (11)$$

This form, however, cannot be directly derived from a linear additive model of Equation (2). The recommendation has been made because capturing nonlinearity effects is typically done by adding the squared terms into the model. As our focus is in building a set of plausible models, Equation (11) can also be considered as a plausible model. The list of candidate models is presented in Table 1. In Table 1, we present the specifications of the eight candidate models derived in this section and the model parameters in the first column and the second columns, respectively. It is important to note that, in the context of model selection,

the model parameters are regression coefficients plus residual variance ($\sigma 2$) (Burnham and Anderson, 2003, p. 13).

In the next section, we illustrate our recommended procedures. We use two cases to describe our procedures. In both cases, we generate data using Monte Carlo simulations. In the first case, we consider a true model that is purely quadratic. In the second case, we consider a true model that contains a moderation effect.


[INSERT TABLE 1 HERE]


## 5    Illustrations

### 5.1    Study 1: True model is $Y = X + X^2 + \varepsilon$.

We illustrate, by use of simulation, how an information-theoretic approach can be used as an additional tool to avoid spurious moderation. We simulate $Y = X + X^2 + \varepsilon, \varepsilon \sim N(0,1)$, where $X$ and $Z$ are sampled from a bivariate normal distribution. Both variables have a zero mean and standard deviation of one and the correlation between the two variables equals 0.2. We generate 300 cases. Estimating Equation (1), the coefficient of the interaction $XZ$ is significant ($p < 0.001$). In fact, the coefficient of the main term $X$ is also significant. Researchers who are unaware of the issue of spuriousness would have concluded that the result provides evidence to support their moderation effect hypothesis. Next, we estimate various model specifications that serve as the candidate models to the true model including the true model itself. These candidate models were discussed in the previous section (see the first column of Table 2). We present the results of our simulation in Table 2.

Table 2 provides information regarding the quality of the eight candidate models, which includes the information about the number of parameters in each model, the residual

sum of squares, AIC values, distance (D) and weight. These values are common statistics reported in an information-theoretic study. Using the distance (D) < 2 criterion (Burnham and Anderson, 2003), models 5, 6, and 7 can be considered as highly plausible, where each model seems to provide better fit to the data compared to the other candidate models listed in the first column of the Table 2. Inspecting Table 2, it is obvious that the moderated regression model, model 4, is among the least plausible models, thus this result should raise a doubt about the hypothesized moderation effect and model 4 should not be selected for further discussion. Thus, given that the true relationship in the true model is curvilinear, using the information- theoretic approach protects researchers from making a Type I error and informs researchers about the quality of their moderated regression model in lieu of other candidate models. If one estimates the traditional moderated regression model alone using a p-value approach, in this scenario, one will be highly likely to conclude that a moderation effect exists when in fact, it does not.  Note that as we can see in Table 2, the best model is model 5 (D=0.000), while the true model is only the fifth best model in the set (D=3.063). As the point of the paper is in comparing the target model (i.e., model 4 in the table) with other competing models in the set, it does not matter which model is being selected as the best model. Our objective is to evaluate to quality of model 4 with respect to the competing models. With the true functional form used to generate the data in this example, obviously, model 4 has no support.


[INSERT TABLE 2 HERE]


5.2    Study 2: True model is $Y = 1 + 2X + 4Z + 7XZ + \varepsilon$

Now, we consider a true model that contains a moderation effect. We simulate $Y = 1 + 2X + 4Z + 7XZ + \varepsilon, \varepsilon \sim N(0,1)$, where $X$ and $Z$ are sampled from bivariate normal

distribution. Both *X* and *Z* have a zero mean and standard deviation of one and the correlation between *X* and *Z* is 0.2. We generate 300 cases. Similar to Case 1 simulation, we fit eight candidate models to the generated data. We present the results of our simulation in Table 3.

[INSERT TABLE 3 HERE]

Table 3 provides information regarding the quality of the eight models, which includes the information about the number of parameters in each model, the residual sum of squares, AIC values, distance and weight. Using the distance (D) < 2 criterion (Burnham and Anderson, 2003), model 4 (D = 0) is selected as the best model that provides a better fit to the data compared to the other candidate models. Testing the significance of the main and interaction terms of model 4, it is no surprise that all terms are significant because model 4 has the same model specification as that of the true model. Thus, in this particular example, we can conclude that the moderation effect found is very unlikely to be spurious.

## 6    Demonstrations using real data

The above section used simulated data to illustrate the use of an information theoretic approach in avoiding spurious moderation effects. In this section, we present how the method works using real data as, unlike the simulation data above, we do not know what the true model is. For this purpose, we use two data sets used in Babcock et al. (2017) and Balakrishnan et al. (2017), respectively. Both data sets are hosted at the Open Science Framework's website and are free to download.  Open Science Framework (OSF) is an open source research project that facilitates an open collaboration among researchers in psychology to replicate previous research in order to enhance our understanding of psychological phenomena and research credibility, which is very important for scientific progress (Collaboration 2015; Camerer et al. 2018).

6.1    Study 3. Replicating Babcock et al.'s (2017) study

We used a data set from Babcock et al. (2017)[2] to re-examine whether the moderation effect that they found was robust against the influence of potential quadratic effects. Babcock et al. (2017) examined the moderation effect of empathy on the relationship between social class on utilitarian moral judgment. Babcock et al. (2017) was a replication research attempting to reproduce findings from Piff et al. (2012) study. Babcock et al. (2017) conducted two replication studies. In both replication studies, they conducted a between-subject experiment manipulating empathy using a fictitious scenario (i.e., empathy condition vs. control condition). Social class was measured with self-reported income and utilitarian moral judgment was measured via an allocation task where at the end of the experiments, participants were asked to indicate how many experimental dollars they would take away (Babcock et al. 2017, p.3). Of the two studies, only in the second replication study did Babcock et al. (2017) successfully obtain similar findings to those reported in Côté et al. (2013). That is, they found that empathy moderates the relationship between social class and utilitarian moral judgments, such that the impact of social class on utilitarian moral judgment was weakened when empathy was elicited; whereas social class is positively related to utilitarian moral judgment when empathy was not elicited (i.e., in a control condition).

For the purpose of our demonstration for the robustness check of significance of the moderation effect using the information-theoretic approach, we used a data set (N=301) from the second study of Babcock et al. (2017). We used three variables provided in the data set: Condition, Income and Utilitarian Judgment. Condition is a dummy variable representing an experimental group where value 0 indicates a control condition and value 1 indicates an empathy condition. Income indicates annual household income measured in dollars and

---

[2] https://osf.io/wq6nr/

Utilitarian moral judgment is a single item measured in dollars. The descriptive statistics of these variables are presented in Table 4. First, we examine the moderation effect of empathy on the relationship between income and utilitarian moral judgment by fitting a moderated regression model. In the moderated regression model, empathy, income and an interaction term created from the multiplication of empathy and income act as the independent variables and utilitarian moral judgment acts as the dependent variable. We present the regression results in Table 5. Table 5 shows that Condition, Income and the interaction term, Condition x Income, are significant ($p < 0.05$), replicating the results obtained by Babcock et al. (2017).

Next, we conducted an information-theoretic approach to check the robustness of the model against nonlinearity. We present our results in Table 6. Our results show that models 2, 4, 6, and 7 have substantial support as the D values of these models are less than 2 (Burnham and Anderson, 2003). Model 4, which is a moderated regression model tested previously using the p-value approach, has the lowest D value. Thus, model 4 represents the best model among all plausible models in the set of models listed in Table 6. Interestingly, despite our results suggesting that models 2, 6, 7 have substantial support, when using the p-value approach, only in model 6 and model 7 are the terms $X$, and $XZ$ significant. We also see that none of the coefficients in model 2 are significant. It seems that the presence of the insignificant terms in model 2 may merely contribute to the reduction of the AIC and D values, but their presence has no substantial interpretations as they are not significant. We conclude that this replication study shows the robustness of the results reported by Babcock et al. (2017) regarding the moderation effect of empathy on the relationship between social class and utilitarian moral judgment. That is, the significant moderation effect is very unlikely to be misleading or spurious due to nonlinearity effects. Furthermore, the information-theoretic approach supplies us with evidence that the moderation model has good quality with respect to other plausible models.

## 6.2    Study 4.  Using data from  Balakrishnan et al. (2017)

We used a data set (N=264) from one of the Balakrishnan et al. (2017) replication studies[3]. Specifically, we used a data set of their first replication study provided in the SPSS format on the open science framework website. In their studies, Balakrishnan et al. (2017) examined the moderation effect of greed on the relationship between social class— measured with socio-economic status (SES) and unethical behaviour (e.g., cheating, lying). The goal of their studies was to directly replicate study 7 in Piff et al. (2012).  Piff et al. (2012) reported that there is a significant positive relationship between SES and unethical behaviour and greed was found to moderate the relationship. That is, Piff et al. (2012) showed that upper-class individuals were more likely to behave unethically than lower class individuals. However, when individuals were prompted to think about the benefits of acting greedily, individuals in both high and low SES indicated a similar propensity to engage in unethical behaviour. Balakrishnan et al. (2017) conducted four replication studies. The four studies failed to reproduce the moderation effect reported in Piff et al. (2012).  In a meta-analysis of the findings of the four studies and the findings of the original study by Piff et al. (2012), Balakrishnan et al. (2017) found that there was indeed no moderating effect of greed on the relationship between SES and unethical behaviour.

To demonstrate the use of the information theoretic approach in preventing spurious moderation effect that we recommend in this paper, the non-significant moderation effect of greed on the relationship between SES and unethical behaviour would not serve our purpose. Furthermore, we want to show how the method works using real data without prior knowledge of the results of previous analysis. Therefore, we want to test a novel idea regarding the moderator effect of the relationship between SES and unethical behaviour. We looked for other variables available in the data set that might be used as a moderator of the

---

[3] Data is downloaded from https://osf.io/nqc8t/.

relationship between SES and unethical behaviour. We found that we could use religiousness as the potential moderator. We hypothesized that religiousness may moderate the relationship between SES and unethical behavior. Following Piff et al. (2012) that shows a negative main effect of SES on unethical behavior, we hypothesize that there is a positive relationship between SES and unethical behaviour. Furthermore, we hypothesize that this positive relationship will be moderated by religiousness, such that the effect will be weaker for individuals who are high in religiosity than for those who are low in religiosity.

We used religiousness and SES variables in Balakrishnan et al. (2017) data set. Both variables were measured with a single item. The item for religiousness was "how religious are you?" with a response answer ranging from 1=not all to 7=deeply. Item for SES was "Think of a ladder with 10 rungs representing where people stand in your local community. Please indicate which rung you would belong to in your community.", with a response ranging from 1=bottom rung to 10=top rung. We used one variable in the data set that belongs to the unethical behaviour scale, which measures an individual's propensity to engage in unethical behaviour with a response answer ranging from 1= very unlikely to 7=very likely[4]. The correlation among these variables is presented in Table 7. As shown in the table, the correlation between SES and Unethical Behavior was significant ($p<0.01$). Other correlations are not significant. We present our moderated regression results in Table 8.

[INSERT TABLE 8 HERE]

[INSERT TABLE 9 HERE]

Table 8 shows that both the main effect of SES and Religiousness on Unethical Behaviour were not significant. Interestingly, the interaction effect between SES and

---

[4] In their replication study, Balakhrisnan et al. (2017) used 12 items to measure unethical behaviour. For the purpose of our demonstration, we only used one item coded in data as UnethicalBeh7.

Religiousness was significant ($p<0.05$). Next, we conducted an information-theoretic approach to check the robustness of the model against nonlinearity.

We present the results of the information-theoretic approach in Table 9. Our results show that only model 7 has substantial support with D-value equal to zero. This model contains all the linear terms, all the quadratic terms, an interaction term, and $XZ^2$ interaction term. Interestingly, testing each regression coefficient of model 7 using the p-value approach, only two terms are (marginally) significant: $XZ$ was marginally significant ($p=0.059$) and $XZ^2$ was significant ($p = 0.020$). Model 4, which is the target model, essentially has no support given that model 4's D-value is greater than 10 (Burnham and Anderson, 2003). Note that, testing each regression coefficient in model 4 with the p-value approach conducted in the previous step, the interaction effect was found to be significant and none of the main effects were significant. As model 4 is only capturing linear relationships, thus given the results presented in Table 9, we should cast a doubt about the results of the moderated regression model using the p-value approach. It is very likely that the moderation effect found using the p-value approach in the previous step is spurious. It seems that the nonlinearity in the religiousness variable, represented with the quadratic term $X^2$, is the source of the spuriousness. As we can see from Table 9, the quadratic term, $X^2$, was significant in models 5, 6, 8 and the term $ZX^2$ was significant in model 7.

# 7 Discussion

In the present research, we focus on avoiding curvilinear effects that can lead to misleading moderation. Specifically, the emphasis of the paper is what to do after researchers find their hypothesized significant effects were significant, i.e., a robustness check. To convey the novel ideas of using an information-theoretic approach in undertaking a moderation analysis, the present research contains five studies: three Monte Carlo simulation

studies using a pre-specified data generation process and two demonstration studies using real data. The first Monte Carlo study provides a demonstration of how a spurious moderation effect can occur. For this demonstration, we choose a nonlinear function that is purely quadratic on its predictor. We noted that Lubinski and Humphreys (1990) were among the first to illustrate this spurious moderation effects phenomena. But, although, we are not the first in demonstrating spurious moderation effects, our aim with the replication is to revive the issue and bring it to the marketing domain in which moderated regression analysis is widely used and one of the marketing researchers' primary tools in testing moderation effects. The second and third Monte Carlo studies demonstrate how an information-theoretic approach can be used as an additional tool to avoid spurious moderation effects. These two Monte Carlo studies also introduce our suggested two-step procedure in testing a moderation effect that combines the use of the traditional p-value testing in the first step and the use of an information approach in the second step. The fourth and the fifth studies used real data from published studies which offers working examples that demonstrate how our approach works in a lifelike situation.

Our primary goal in the present research is to contribute to the improvement of researchers' methodological repertoire in conducting moderation analysis. We recommend that researchers adopt an information theoretic approach in undertaking a moderated regression analysis as part of a robustness check procedure following a significant moderation effect. The procedure can be undertaken following the p-value approach in testing the coefficients of the moderated regression model. In essence, our suggested procedure comprises of two steps: (1) assessing the significance of moderation using the traditional p-value approach and (2) assessing the quality of the moderation model using the information-theoretic approach. The key feature of the information-theoretic approach is that the moderation regression model is compared to a set of plausible or candidate models.

Fortunately, the functional form of the plausible models can be formulated by varying the slope coefficient of a linear additive model. In total, we identified eight plausible models where seven models are developed by varying the slope coefficient and one model is inserted into the set of plausible models because of it has been used in the moderation literature.

It is important to note that there are three models in the plausible set that contains an interaction term as well as the squared terms of the main predictors ($X^2$ or $Z^2$) (see Table 1). Adding non-linear terms to a regression model has been known to be a standard procedure in regression analysis to capture non-linear relationships that might exist in the data (Wooldridge 2015). Thus, this may bring up the question of why an information-theoretic approach is needed at all (i.e., in step 2 of our suggested procedure) as adding quadratic terms into Equation (1) already seems to safeguard researchers from obtaining spurious moderation due to nonlinearities. In fact, we identified a very small number of marketing papers that conducted a robustness check following a significant moderation effect by adding a squared term of a predictor into moderated regression model (e.g., Homburg et al. 2012; DeCarlo and Lam 2016; Alavi et al. 2018). Our reasons are twofold. First, by using an information-theoretic approach, researchers protect themselves not only from making an erroneous conclusion about spurious moderation effects but also giving them information about the quality of their moderated regression models in lieu of other competing models. Finding alternatives models that might be better than researchers' moderation regression models might create further research enquiries. For example, a hypothesized moderation effect might be theoretically justified with the inclusion of nonlinear effects. Second, past research has shown that merely adding quadratic terms into Equation (1) can increase the probability of Type II error (i.e., failing to reject the null moderation hypothesis when in fact, the true model includes moderation) when the true model does not include quadratic terms (Ganzach 1998).

# 8    Conclusion

In this paper, we demonstrated the use of an information-theoretic approach in moderation analysis using simulated data and real data from published studies. As shown by the results of our demonstrations, the use of multi-model inference using the information-theoretic approach has a promise for researchers who want to evaluate the quality of their moderated regression model in lieu of alternative plausible models as well as prevent them from making an erroneous conclusion regarding the presence of their hypothesized moderation effects. The multi-model inference based on a model of Equation (1) has been integrated into a user-friendly SPSS macro and R script to produce findings in this paper (e.g., Table 4) and is available to download from our personal website.

To conclude, the issue of spuriousness in moderation, to some extent, has been neglected in marketing literature as well as in the wider literature of business research. We hope that this paper will increase researchers' awareness of the potential shortcomings of curvilinear effects in moderation.  Specifically, we encourage researchers to consider the use of an information-theoretic approach in their works and include our suggested procedures in their research toolbox.

# 9    References

Aiken L. S., West, S. G., Reno, R. R. (1991). Multiple regression: Testing and interpreting interactions: Sage.

Akaike H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control 19 (6): 716-723.

Alavi S., Habel, J., Guenzi, P., Wieseke, J. (2018). The role of leadership in salespeople's price negotiation behavior. Journal of the Academy of Marketing Science 46 (4): 703-724.

Anderson D., Burnham, K., White, G. (1998). Comparison of AIC and CAIC for model selection and statistical inference from capture-recapture studies. Journal of Applied Statistics 25 (2): 263-282.

Babcock S., Li, Y., Sinclair, V. M., Thomson, C., Campbell, L. (2017). Two replications of an investigation on empathy and utilitarian judgement across socioeconomic status. Scientific Data 4: 160129.

Balakrishnan A., Palma, P. A., Patenaude, J., Campbell, L. (2017). A 4-study replication of the moderating effects of greed on socioeconomic status and unethical behaviour. Scientific data 4: 160120-160120.

Balli H. O., Sørensen, B. E. (2013). Interaction effects in econometrics. Empirical Economics: 1-21.

Burnham K. P., Anderson, D. (2003). Model selection and multi-model inference. USA: Springer.

Burnham K. P., Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. Sociological methods & research 33 (2): 261-304.

Camerer C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., Isaksson, S., Manfredi, D., Rose, J., Wagenmakers, E.-J., Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. Nature Human Behaviour 2 (9): 637-644.

Collaboration O. S. (2015). Estimating the reproducibility of psychological science. Science 349 (6251): aac4716.

Cortina J. M. (1993). Interaction, nonlinearity, and multicollinearity: Implications for multiple regression. Journal of Management 19 (4): 915-922.

Côté S., Piff, P. K., Willer, R. (2013). For whom do the ends justify the means? Social class and utilitarian moral judgment. Journal of Personality and Social Psychology 104 (3): 490.

DeCarlo T. E., Lam, S. K. (2016). Identifying effective hunters and farmers in the salesforce: a dispositional–situational framework. Journal of the Academy of Marketing Science 44 (4): 415-439.

Ganzach Y. (1997). Misleading interaction and curvilinear terms. Psychological Methods 2 (3): 235.

Ganzach Y. (1998). Nonlinearity, Multicollinearity and the Probability of Type II Error in Detecting Interaction. Journal of Management 24 (5): 615-622.

Hayes A. (2012). Process SPSS macro. Computer Software and Manual] Available from http://www afhayes com/public/process pdf.

Homburg C., Klarmann, M., Staritz, S. (2012). Customer uncertainty following downsizing: The effects of extent of downsizing and open communication. Journal of Marketing 76 (3): 112-129.

Lance C. E. (1988). Residual centering, exploratory and confirmatory moderator analysis, and decomposition of effects in path models containing interactions. Applied Psychological Measurement 12 (2): 163-175.

Lubinski D., Humphreys, L. G. (1990). Assessing spurious" moderator effects": Illustrated substantively with the hypothesized (" synergistic") relation between spatial and mathematical ability. Psychological bulletin 107 (3): 385.

MacCallum R. C., Mar, C. M. (1995). Distinguishing between moderator and quadratic effects in multiple regression. Psychological Bulletin 118 (3): 405.

Piff P. K., Stancato, D. M., Côté, S., Mendoza-Denton, R., Keltner, D. (2012). Higher social class predicts increased unethical behavior. Proceedings of the National Academy of Sciences 109 (11): 4086-4091.

Wooldridge J. M. (2015). Introductory econometrics: A modern approach: Nelson Education.

Panel A.  Reliability of *X, Z* and *Y* is equal to *0.4*.
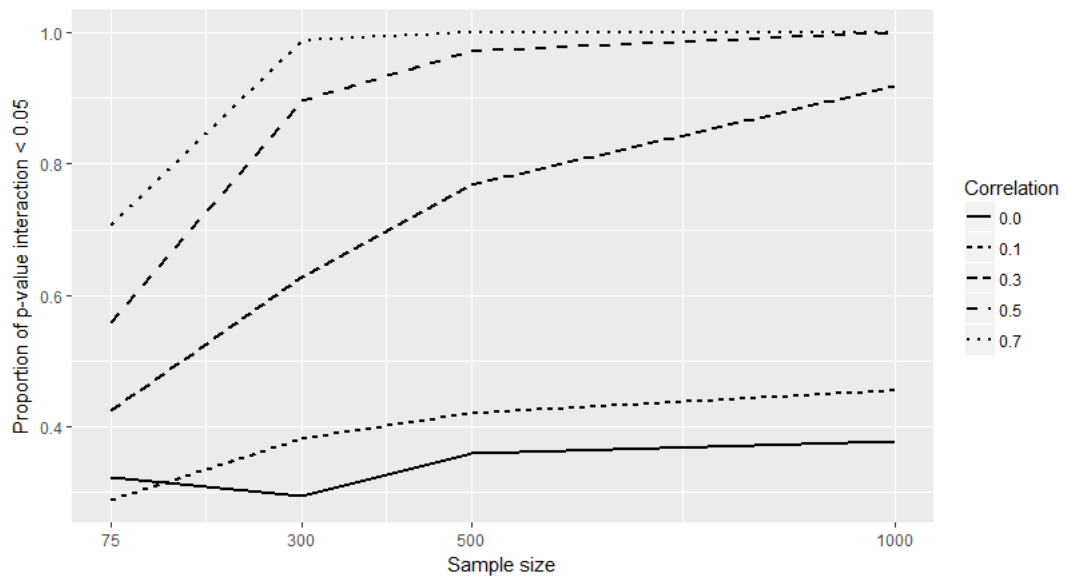


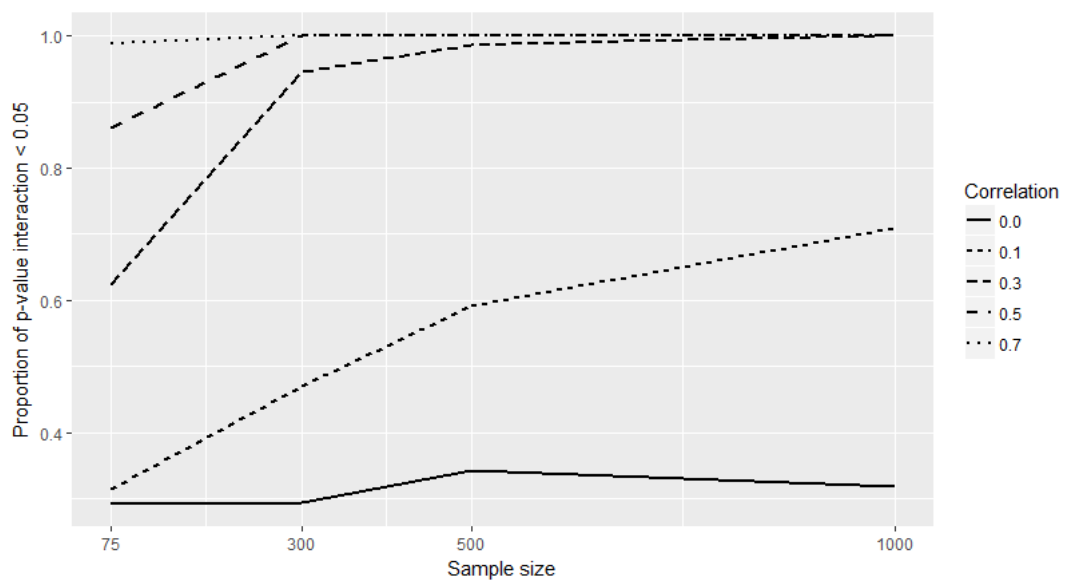Panel B. Reliability of *X, Z,* and *Y* is equal to 0.8



**Fig 1.** The effect of sample size and the correlation between *X* and *Z* on estimating linear, interaction model when the true model is quadratic: $Y = X + X^2 + \varepsilon$, $\varepsilon \sim N(0,1)$, $X \sim N(0,1)$.

Table 1. Candidate models for a moderation effect model.

| No | Model Specifications | $K$ |
|---|---|---|
| 1 | $Y=\beta_0+\beta_1 X+\beta_1 Z$ | 4 |
| 2 | $Y=\beta_0+\beta_1 X+\beta_1 X^2$ | 4 |
| 3 | $Y=\beta_0+\beta_1 Z+\beta_1 Z^2$ | 4 |
| 4 | $Y=\beta_0+\beta_1 X+\beta_1 Z+\beta_1 XZ$ | 5 |
| 5 | $Y=\beta_0+\beta_1 X^2+\beta_2 Z^2+\beta_3 X+\beta_4 Z$ | 6 |
| 6 | $Y=\beta_0+\beta_1 X^2+\beta_2 Z^2+\beta_3 X+\beta_4 Z+\beta_5 XZ$ | 7 |
| 7 | $Y=\beta_0+\beta_1 X^2+\beta_2 Z^2+\beta_3 X+\beta_4 Z+\beta_5 XZ+\beta_6 ZX^2$ | 8 |
| 8 | $Y=\beta_0+\beta_1 X^2+\beta_2 Z^2+\beta_3 X+\beta_4 Z+\beta_5 XZ+\beta_6 XZ^2$ | 8 |

Note: Regression error terms are omitted from each model. $K$=model parameters=regression coefficients plus residual variance ($\sigma^2$) (Burnham and Anderson, 2003, p. 13).

Table 2. Simulation of Case 1

| True model is $Y = X + X^2 + \varepsilon, \varepsilon \sim N(0,1)$. | | | | | |
|---|---|---|---|---|---|
| **Model** | **K** | **RSS** | **AIC** | **D** | **W** |
| 1. $y = x + z$ | 4 | 971.439 | 1212.740 | 418.996 | 0.000 |
| 2. $y = x + x^2$ | 4 | 242.823 | 796.807 | 3.063 | 0.092 |
| 3. $y = z + z^2$ | 4 | 1118.743 | 1255.095 | 461.351 | 0.000 |
| 4. $y = x + z + xz$ | 5 | 913.361 | 1196.246 | 402.502 | 0.000 |
| 5. $y = x^2 + z^2 + x + z$ | 6 | 237.173 | 793.744 | 0.000 | 0.424 |
| 6. $y = x^2 + z^2 + x + z + xz$ | 7 | 236.668 | 795.104 | 1.360 | 0.215 |
| 7. $y = x^2 + z^2 + x + z + xz + zx^2$ | 8 | 235.584 | 795.727 | 1.984 | 0.157 |
| 8. $y = x^2 + z^2 + x + z + xz + xz^2$ | 8 | 236.127 | 796.417 | 2.673 | 0.112 |

Note: True model is $Y = X + X^2 + \varepsilon, \varepsilon \sim N(0,1)$. $X$ and $Z$ are bivariate normal with mean zero and standard deviation of one. The sample size is 300. For simplicity, in the first column, we did not include model parameters including an intercept and an error term in each equation. K=number of estimated parameters, RSS=residual sum of squares, AIC=Akaike information criterion (because of the relatively small sample size, AIC correction was used), D=distance, W=weight.

Table 3. Simulation of Case 2

| True model is: $Y = 1 + 2X + 4Z + 7XZ + \varepsilon, \varepsilon \sim N(0,1)$. | | | | | |
|---|---|---|---|---|---|
| **Model*** | **K** | **RSS** | **AIC** | **D** | **W** |
| 1. $y = x + z$ | 4 | 25857.559 | 2197.214 | 1347.400 | 0.000 |
| 2. $y = x + x^2$ | 4 | 26583.657 | 2205.522 | 1355.708 | 0.000 |
| 3. $y = z + z^2$ | 4 | 21064.680 | 2135.712 | 1285.898 | 0.000 |
| 4. $y = x + z + xz$ | 5 | 287.827 | 849.814 | 0.000 | 0.643 |
| 5. $y = x^2 + z^2 + x + z$ | 6 | 19829.875 | 2121.590 | 1271.776 | 0.000 |
| 6. $y = x^2 + z^2 + x + z + xz$ | 7 | 287.043 | 852.996 | 3.182 | 0.131 |
| 7. $y = x^2 + z^2 + x + z + xz + zx^2$ | 8 | 287.008 | 854.959 | 5.145 | 0.149 |
| 8. $y = x^2 + z^2 + x + z + xz + xz^2$ | 8 | 284.556 | 852.396 | 2.582 | 0.177 |

*Note: $X$ and $Z$ are bivariate normal with mean zero and standard deviation of one. The sample size is 300. For simplicity, in the first column, we did not include model parameters including an intercept and an error term in each equation. K=number of estimated parameters, RSS=residual sum of squares, AIC=Akaike information criterion (Because of the relatively small sample size, AIC correction was used), D=distance, W=weight.

Table 4.

Means, Standard Deviations and Correlations among Experimental condition, Income and Utilitarian Moral Judgment of Balakrishnan et al. (2017) data (N=264)

| Variable | Mean (SD) | 1. | 2. | 3. |
|---|---|---|---|---|
| 1. Condition | 0.465 (0.499) | 1.00 | | |
| 2. Income | 4.419 (2.014) | -.028 | 1.00 | |
| 3. Utilitarian moral judgment | 2.468 (1.724) | .025 | 0.042* | 1.00 |

Note: Condition is a manipulation variable with 0=control, 1=empathy. * p-value $< 0.05$.

Table 5. Results of the regression of the effect of Experimental Condition, Income and their interaction on Utilitarian Moral Judgment of Balakrishnan et al. (2017) data (N=301)

| Variable | Coefficient | Standard error | T-value | P-value | Lower 95% CI | Upper 95% CI |
|---|---|---|---|---|---|---|
| Constant | 1.832 | 0.331 | 5.529 | 0.000 | 1.183 | 2.482 |
| Condition | 1.000 | 0.479 | 2.088 | 0.038 | 0.061 | 1.938 |
| Income | 0.133 | 0.068 | 1.972 | 0.049 | 0.001 | 0.266 |
| Condition x Income | -0.206 | 0.099 | -2.089 | 0.038 | -0.400 | -0.013 |

Note: The regression coefficients are unstandardized estimates. * p-value $< 0.05$.

Table 6. Results of the multi-model inference using data from Babcock et al. (2017)

| Hypothesized model is $Y = X + Z + XZ + \varepsilon, \varepsilon \sim N(0,1)$. | | | | | |
|---|---|---|---|---|---|
| **Model*** | **K** | **RSS** | **AIC** | **D** | **W** |
| [a]1. $y = x + z$ | 4 | 892.762 | 1189.451 | 2.391 | 0.094 |
| 2. $y = x + [x^2]$ | 3 | 894.400 | 1188.003 | 0.943 | 0.193 |
| 3. $y = z + z^2$ | 4 | 893.114 | 1189.570 | 2.509 | 0.088 |
| 4. $y = x + z + xz$ | 5 | 879.833 | 1187.060 | 0.000 | 0.310 |
| 5. $y = [x^2] + z^2 + x + z$ | 5 | 892.518 | 1191.369 | 4.308 | 0.036 |
| 6. $y = [x^2] + z^2 + x + z + xz$ | 6 | 879.744 | 1189.030 | 1.969 | 0.116 |

| | | | | | |
|---|---|---|---|---|---|
| 7. $y = [x^2] + z^2 + \boldsymbol{x} + z + \boldsymbol{xz} + [zx^2]$ | 6 | 879.744 | 1189.030 | 1.969 | 0.116 |
| 8. $y = [x^2] + z^2 + x + z + xz + xz^2$ | 7 | 879.100 | 1190.809 | 3.749 | 0.048 |

*Note: As in Babcock et al. (2017), *X* is a dummy variable denoting an experimental condition (0=control condition, 1=empathy condition). *Z* denotes self-reported income; *Y* denotes utilitarian moral judgment. The sample size is 301. Because of the relatively small **K**, AIC correction was **not** used. [a] Because *X* is a dummy variable codes as 0/1, to avoid perfect multicollinearity, the squared term of *X* does not appear in models 2, 5, 6 ,7 and model 8; the missing squared term is represented with a bracket term $[x^2]$. [b] only indicates variables of models with D < 2 except for model 1. Variables written in bold under the first column are significant (p < 0.05).

Table 7.
Means, Standard Deviations and Correlations among SES, Religiousness
and Unethical behaviour of Balakrishnan et al. (2017) data (N=264)

| Variable | Mean (SD) | 1. | 2. | 3. |
|---|---|---|---|---|
| 1. SES | 5.201 (1.901 | 1.00 | | |
| 2. Religiousness | 3.352 (2.146) | .113 | 1.00 | |
| 3. Unethical behaviour* | 1.424 (1.129) | .308** | 1.98** | 1.00 |

[a]Note: In their replication study, Balakhrisnan et al. (2017) used 12 items to measure unethical behaviour. For the purpose of our demonstration, we only used one item coded in data as UnethicalBeh7. ** p-value < 0.01.

Table 8. Results of the regression of the effect of SES, Religiousness and their interaction on Unethical Behaviour

| Variable | Coefficient | Standard error | T-value | P-value | Lower 95% CI | Upper 95% CI |
|---|---|---|---|---|---|---|
| Constant | 0.800 | 0.328 | 2.441 | 0.015 | 0.158 | 1.442 |
| SES | 0.059 | 0.062 | 0.964 | 0.336 | -0.061 | 0.18 |
| Religiousness | -0.087 | 0.085 | -1.031 | 0.304 | -0.254 | 0.079 |
| SES x Religiousness | 0.034 | 0.015 | 2.202 | 0.029* | 0.004 | 0.064 |

Note: The regression coefficients are unstandardized estimates. * p-value < 0.05.

Table 9. Results of the multi-model inference using data from Balakrishnan et al. (2017)

| Hypothesized model is $Y = X + Z + XZ + \varepsilon, \varepsilon \sim N(0,1)$. | | | | | |
|---|---|---|---|---|---|
| **Model*** | **K** | **RSS** | **AIC** | **D** | **W** |
| 1. $y = x + z$ | 4 | 295.503 | 786.961 | 17.314 | 0.000 |
| 2. $y = x + x^2$ | 4 | 286.634 | 778.915 | 9.268 | 0.007 |
| 3. $y = z + z^2$ | 4 | 323.318 | 810.709 | 41.062 | 0.000 |
| 4. $y = x + z + \boldsymbol{xz}$ | 5 | 290.091 | 784.081 | 14.434 | 0.000 |
| 5. $y = \boldsymbol{x^2} + z^2 + \boldsymbol{x} + z$ | 6 | 276.959 | 773.851 | 4.204 | 0.083 |
| 6. $y = \boldsymbol{x^2} + z^2 + \boldsymbol{x} + z + xz$ | 7 | 274.180 | 773.188 | 3.541 | 0.116 |
| 7. $y = \boldsymbol{x^2} + z^2 + \boldsymbol{x} + z + xz + \boldsymbol{zx^2}$ | 8 | 268.485 | 769.647 | 0.000 | 0.680 |
| 8. $y = \boldsymbol{x^2} + z^2 + \boldsymbol{x} + z + xz + xz^2$ | 8 | 272.140 | 773.216 | 3.569 | 0.114 |

*Note: $X$ is a self-reported measure of socio-economic status, SES_r in the data set; $Z$ is a self-reported measure of religiousness; and $Y$ is a self-reported measure of unethical behaviour, UnethicalBeh7 in the data set. The sample size is 264. Because **K** is relatively small and sample size is relatively large, AIC correction ($AIC_c$) was **not** used. [b] only indicates variables of models with D < 2 except for model 1. Variables written in bold under the first column were found to be significant (either p < 0.001, p < 0.01 or p < 0.05).