國立交通大學

統計學研究所

碩 士 論 文

廣義Neyman-Rubin的因果模型在評估迴歸上交互作用的應用

Generalized Neyman-Rubin's causal model for Regression Interaction Assessment

研 究 生：莊揚凱

指導教授：陳鄰安　教授

中 華 民 國 一 百 零 二 年 六 月

廣義 Neyman-Rubin 的因果模型在評估迴歸上交互作用的應用

Generalized Neyman-Rubin's causal model for Regression Interaction Assessment

研　究　生：莊揚凱　　Student：Yang-Kai Chuang

指導教授：陳鄰安　　Advisor：Dr. Lin-An Chen

國　立　交　通　大　學

統計學研究所

碩　士　論　文

A Thesis
Submitted to Institute of Statistics
College of Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in

Statistics

June 2013

Hsinchu, Taiwan, Republic of China

中華民國一百零二年六月

廣義 Neyman-Rubin 的因果模型在

評估迴歸上交互作用的應用

研 究 生：莊揚凱　　　　　　　　指導教授：陳鄰安 博士

國立交通大學統計學研究所

摘要

儘管在經濟，社會和健康科學上，藉由插入一個分析模型的相乘項來檢定交互作用影響的表現是非常常見的，但交互作用是否存在決定於模型型態為一直被批評的地方(Greenland (2009) and Mauderly and Samet (2009))。有些文章努力解決這個爭議性問題但卻導致於更複雜且不清楚的交互作用定義。這讓評估統計交互作用更加困難(Greenland (1980))。我們提出一個有系統的定義介紹，方法和定理將相互關係(結合)參數融入在廣義 Neyman-Rubin 的因果模型。這項創舉帶來許多的優點：

(a) 此方法允許我們定義和測量關於未知統計上交互作用的統計推論的相互關係影響。

(b) 對於統計上交互作用的統計推論全都可從分佈參數的估計理論來建構。

(c) 此因果模型測量一個明確且模型獨立能避免插入爭論的相互關係影響。

(d) 廣義 Neyman-Rubin 的因果分析理論擴展到對於 probit 迴歸的統計交互作用評估。

關鍵字：因果推論；內部相互關係；迴歸分析；統計交互作用。

Generalized Neyman-Rubin's causal model for Regression Interaction Assessment

Student：Yang-Kai Chuang Advisor：Dr. Lin-An Chen

Institute of Statistics
National Chiao Tung University

Abstract

Although the insertion of product terms into analytical to test for presence of interaction effect is very common in economic, social and health sciences, it has long been criticized for that existence of interaction is model dependent (Greenland (2009) and Mauderly and Samet (2009)). The efforts for resolving this criticism leads to multiple but ambiguous definitions of statistical interaction resulting in assessing various but unknown versions of effect (Greenland (2009)). We report that a systematic introduction of definitions, methods and theorems to fit the intercorrelation (association) parameter into a generalized Neyman-Rubin's causal model brings interesting advantages: (a) This approach allows us to define and measure a clean effect of intercorrelation for statistical inferences of unknown statistical interaction. (b) Statistical inferences for statistical interaction all can be constructed from the estimation theory of the distributional parameters. (c) This causal model measures an unambiguous but also model independent effect of intercorrelation that avoids the controversy of insertion. (d) The theory of the generalized Neyman-Rubin's causality is extended to statistical interaction assessment for probit regression.

Key words: Causal inference; intercorrelation; regression analysis; statistical interaction.

誌謝

在碩士這兩年中，我最感謝的是我的指導教授陳鄰安老師，這一年來跟老師一起做研究，老師總是詳細地教導我，當我論文遇到問題時，老師也很耐心地引導我找出疑點並進行改善。我覺得跟老師做研究，學到的不僅僅是學術上的知識，更讓我學習到如何在遇到問題時，找出解決問題的方法與態度，想必這在往後的工作以及待人處世上有深深的影響。也要謝謝我的口試委員許文郁老師、蕭金福老師及洪慧念老師，老師們給予我的建議，使我的論文更加完整。

再來我要感謝交大統研所碩士班的所有同學，沒有你們，我的碩士生活不會如此精采。有好多我們一起做過的事，現在還是歷歷在目，不論是班遊、統研盃及研究室發生的點點滴滴，都讓我的碩士的生活增添了許多歡樂的回憶，謝謝你們這一群好同學好朋友。

最後我要感謝的是我的家人，從小到大不辭辛勞的栽培我，也都尊重我在我求學過程中所做的選擇，有家人的陪伴讓我順利完成學生生活。

莊揚凱　謹誌於

國立交通大學統計學研究所

中華民國一百零二年六月

目錄

表目錄

Generalized Neyman-Rubin's causal model for Regression Interaction Assessment

Abstract

Although the insertion of product terms into analytical model to test for presence of interaction effect is very common in economic, social and health sciences, it has long been criticized for that existence of interaction is model dependent (Greenland (2009) and Mauderly and Samet (2009)). The efforts for resolving this criticism leads to multiple but ambiguous definitions of statistical interaction resulting in assessing various but unknown versions of effect (Greenland (2009)). We report that a systematic introduction of definitions, methods and theorems to fit the intercorrelation (association) parameter into a generalized Neyman-Rubin's causal model brings interesting advantages: (a) This approach allows us to define and measure a clean effect of intercorrelation for statistical inferences of unknown statistical interaction. (b) Statistical inferences for statistical interaction all can be constructed from the estimation theory of the distributional parameters. (c) This causal model measures an unambiguous but also model independent effect of intercorrelation that avoids the controversy of insertion. (d) The theory of the generalized Neyman-Rubin's causality analysis is extended to statistical interaction assessment for probit regression.

Key words: Causal inference; intercorrelation; regression analysis; statistical interaction.

1. Introduction

The notions of "interaction" is common in researches of business, economics, education, sociology, health science and many others but it has long been in literature with controversies surrounding concept of interaction (Greenland (1993), Rothman, Greenland and Walker (1980) and Suhnel (1992)).

The classical regression interaction assessment often inserts product terms in causal (regression) model as

$$y = g(\beta_0, \beta_1 x_1, \beta_2 x_2, \beta_{12} x_1 x_2) + \epsilon \qquad (1.1)$$

to test hypothesis $H_0 : \beta_{12} = 0$ for the presence of interaction. It has made the controversy that the presence or absence of interaction is entirely dependent on the form of the regression model been choosed; for the same data, interaction may appear to be present when using one regression model but absent when another regression model is applied (Mauderly and Samet (2009) and Rothman, Greenland and Walker (1980)). Some efforts for resolving this controversy are done, for examples, by Mullahy (1999) and Ai and Norton (2003) for cross-derivative method measuring so-called the second order interaction and VanderWeele (2009) and VanderWeele and Robin (2008) for sufficient cause interaction in biologic approach. Although the these approaches can measure effect other than product terms but they and some others measure different versions of interaction that can not avoid the concern of Greenland (1993) and Greenland (2009) that makes the users confused for how much and in what direction we can learn from a data set.

A correct resolution of this controversy requires a mechanism that can accurately measure the causal effect of consensus intercorrelation (association) cause in random world (Rothman, Greenland and Walker (1980) and Ai and Norton (2003)). The approach of Chen et al. (2013) extends the concept of biologic isobole (Loewe (1928, 1953)) to define an unknown statistical isobole for inferences of statistical interaction. While this is interesting that allows us to detect if the statistical interaction is present or not, it can not measures the size of statistical interaction and, isobole is not popularly applied in social and economic sciences. Fitting the statistical interaction into a causal model can completely solve this controversy.

The Neyman-Rubin framework (Neyman (1990) and Rubin (1974)) of causality analysis discovering causal relationships between outcome variable and causal variables has become increasingly popular in applied research (Holland (1986), Rubin (2006) and Sekhon (2008)). This interesting approach, also called the difference in difference method, is a popular tool for evaluating the effects of policy interventions in economics and biology (Abadie (2005) for a review). This article attempts to formulate its generalization so that regression interaction analysis can be done with this common framework with expectation of making this advance of interaction assessment more accessible to the general research community. We derive a Neyman-Rubin's causal model from the joint distribution of outcome and cause variables forcing the regression function to have distributional parameters involved that leads to several important advantages: (a) The novel parametrization of imposing distributional parameters in regression model builds a bridge between the Neyman-Rubin's causal model and the parameter of intercorrelation between causes (explanatory variables) allowing us to measure clean causal effect of intercorrelation. (b) This approach is not model dependent that avoids the controversy of insertion of product terms in model. (c) This success in assessment of statistical interaction by Neyman-Rubin's causality analysis may be applied to statistical interaction for other models where we interpret this for probit regression model.

2. Parametrized Regression for Effect Assessment

2.1. A Normal Regression Model

We first clarify two concepts of our interest when we have random variables y, x_1 and x_2 and our interest is the effect of some causes on variable y. Statistically causal study consider a comparison between the effect (outcome) of variable y on taking a treatment $x_1 = x_{1b}$ relative to the effect of taking another treatment $x_1 = x_{1a}$ holding all other factors unchanged. On the other hand, statistically interaction assessment consider a comparison between the effect of variable y when there is no intercorrelation between variables x_1 and x_2 and the effect of variable y when the intercorrelation is present. Two closely concepts have not been unified in analysis. We propose to assess statistical interaction via a generalized causal model.

A proper perspective in a theoretical analysis of causality requires method to assess effect of the cause to understand why the values of a quantity to be explained is affected with causes (Holland (1986)) while looking for the cause of an effect such as approach of sufficient cause approach of Rothman (1976) and Vanderweele and Robins (2008) is not of this kind. For references of general causal analysis and modeling, see Vanderweele and Rubin (2008) and Sekhon (2008) and for review, see Holland (1986), Heckman (2008) and Rothman and Greenland (1998).

Suppose that variables y, x_1 and x_2 have a joint distribution. We propose to study causal effect from a regression model that is parametrized from this joint distribution. Here we consider joint normal distribution as

$$\begin{pmatrix} y \\ x_1 \\ x_2 \end{pmatrix} \sim N_3(\begin{pmatrix} \mu_y \\ \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_y^2 & \sigma_{y1} & \sigma_{y2} \\ \sigma_{1y} & \sigma_1^2 & \sigma_{12} \\ \sigma_{2y} & \sigma_{21} & \sigma_2^2 \end{pmatrix}). \tag{2.1}$$

for interpretation. The conditional expectation of y given (x_1, x_2) under normality assumption is

$$\mu_{norm}(x_1, x_2) = \mu_y + (\sigma_{y1}, \sigma_{y2}) \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}^{-1} \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix}.$$

Setting a fixed vector of distributional parameetrs as $\theta' = (\mu_y, \mu_1, \mu_2, \sigma_y^2, \sigma_1^2, \sigma_2^2, \sigma_{y1}, \sigma_{y2}, \sigma_{12})$, a parametrized regression model is stated in the following theorem.

Theorem 2.1. The regression model under the normal distribution of (2.1) is

$$y(x_1, x_2, \theta) = \beta_0(\theta) + \beta_1(\theta)x_1 + \beta_2(\theta)x_2 + \epsilon \tag{2.2}$$

where $\beta(\theta) = (\beta_0(\theta), \beta_1(\theta), \beta_2(\theta))'$ with

$$\beta_0(\theta) = \mu_y - \frac{(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})\mu_1}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2} + \frac{(\sigma_{y1}\sigma_{12} - \sigma_{y2}\sigma_1^2)\mu_2}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}$$

$$\beta_1(\theta) = \frac{\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12}}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}$$

$$\beta_2(\theta) = \frac{\sigma_{y2}\sigma_1^2 - \sigma_{y1}\sigma_{12}}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}$$

where error variable ϵ has the normal distribution $N(0, \sigma^2_{y|x_1,x_2}(\theta))$ with

$$\sigma^2_{y|x_1,x_2}(\theta) = \sigma_y^2 - (\sigma_{y1}, \sigma_{y2}) \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{y1} \\ \sigma_{y2} \end{pmatrix}.$$

2.2. Can Classical Interaction Detection Methods Deal with Normal Data?

Parametrized normal regression model provides important messages for verification of existent interaction detection methods. Three methods are considered here. Applied econometrics researchers commonly infer about the presence or absence of statistical interaction via testing a hypothesis for existence of something about interaction. The most popular one is assuming the following regression model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon$$

to test hypothesis $H_0 : \beta_{12} = 0$ for existence of product term interaction. It is argued that power of developed statistical tests for the presence of interaction is remarkably low (Geenland (2009) and Mauderly and Samet (2009)) which is not surprised from our derivation (2.2) that the true regression model does not include any product term even the trivariate data is drawn from a normal distribution. An effort in econometrics for avoiding this model dependence disadvantage is verifying the presence or absence of second-order interaction $\frac{\partial^2 E(y|x_1,x_2)}{\partial x_1 \partial x_2}$. This approach is not applicable too for normal data since its true regression model virtually does not have this interaction. Hence, these two popularly methods can measure the product term interaction and second order interaction that can not measure a pure effect of intercorrelation.

Besides the above, departure from additivity is another popularly used test for biologic synergism effect. Following Greenland (1993), the effects of two explanatory variables are defined as mean differences in the absence of the other variable as

$$E_{x_1} = \tau_{cm}(x_1, 0, \theta) - \tau_{cm}(0, 0, \theta) = \beta_1(\theta) x_1$$
$$E_{x_2} = \tau_{cm}(0, x_2, \theta) - \tau_{cm}(0, 0, \theta) = \beta_2(\theta) x_2$$

and the combined (total) effect as the sum of separate effects is

$$E_{x_1,x_2} = \tau_{cm}(x_1, x_2, \theta) - \tau_{cm}(0, 0, \theta) = \beta_1(\theta) x_1 + \beta_2(\theta) x_2.$$

The test by Prentice and Kalbfleisch (1988) defining synergistic effect if the combined effect E_{x_1,x_2} is greater than the sum of separate effects $E_{x_1} + E_{x_2}$

shows no synergism since $E_{x_1,x_2} = E_{x_1} + E_{x_2}$. This criterion of interaction as departure from additivity limits its application since it can not detect the effect of intercorrelation even the data is observed from a normal distribution.

2.3. Quantities to be Explained

To forestall confusion that the reader of the literature on causality encounters unclearly terminologies, we provide step by step the definitions. Benifited from parametrization, a framework of causal model can assess effects on the internal variable y of various causes of interest not restricted to external variables x_1 and x_2. That is, the potential causes in this causal model includes elements in the following set:

$$\text{Potential causes: } x_1, x_2, \theta. \tag{2.3}$$

where the unobservable external (error) variable is not considered in this paper. We need to specify outcomes of interest.

The classical Neyman-Rubin's causality analysis defines the comparison of effects at treatments $x_1 = x_{1b}$ and $x_1 = x_{1a}$ as the difference of two potential outcomes as

$$y(x_{1b}, x_2, \theta) - y(x_{1a}, x_2, \theta). \tag{2.4}$$

This causal inference is a missing data problem because we cannot observe two outcome variables in (2.4) at the same time (Holland (1986)). Outcome quantities to be explained other than the response variable exist in literature, for examples, a utility function $R(y)$ as subjective evaluation of outcome in economic approach (Heckman (2008)) and biology approach (Greenland (1993)), conditional mean by Holland (1986) and variation effect $\sigma^2_{y|x_1,x_2}(x_1, x_2, \theta)$ in social science (Russo (2011)). A framework broadens the range of quantity to be explained is available.

Definition 2.2. Any quantity $\tau(x_1, x_2, \theta)$ that characterizes the regression model (2.2) is called an outcome quantity.

Example 1. Some interests of outcome quantity are:

(1) Outcome variable: $y(x_1, x_2, \theta)$ measuring the outcome of possible experiment

(2) Conditional mean: $\tau_{cm}(x_1, x_2, \theta) = \beta_0(\theta) + \beta_1(\theta)x_1 + \beta_2(\theta)x_2 = (1, x_1, x_2)\beta(\theta)$ measuring the central tendency of the regression model (Holland (1986))

(3) Conditional variance: $\tau_{cv}(\theta) = \sigma_y^2 - (\sigma_{y1}, \sigma_{y2}) \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{y1} \\ \sigma_{y2} \end{pmatrix}$ measuring the conditional variation of the regression model (applied in social science (Russo (2011)))

(4) Regression parameters: $\beta(\theta)$

(5) Regression quantile: $\beta(\gamma, \theta) = (\beta_0(\theta) + z_\gamma \sqrt{\tau_{cv}(\theta)}, \beta_1(\theta), \beta_2(\theta))'$ satisfying $\gamma = P(Y \leq (1, x_1, x_2)\beta(\gamma, \theta)|x_1, x_2)$ for $(x_1, x_2)' \in R^2$ defined by Koenker and Bassett (1978)

(6) Conditional quantile: $\tau_{cq}(x_1, x_2, \gamma, \theta) = (1, x_1, x_2)\beta(\gamma, \theta)$

(7) Reference charts: $\tau_{rc} = \{(1, x_1, x_2)\beta(\gamma, \theta) : \gamma \in (0, 1), (x_1, x_2)' \in R^2\}$

(8) Conditional signal-to-noise ratio at (x_1, x_2):

$$\tau_{sn}(\theta) = \frac{\tau_{cm}(x_1, x_2, \theta)}{\sqrt{\tau_{cv}(\theta)}}.$$

One can measure the effect of causes on the central tendency of outcome's distribution by conditional mean, but it does not provide a complete picture of this distribution. When one is interested in the distributional extreme behavior, the conditional quantile of (6) above is desired for investigation.

2.4. Motivation of Causality Analysis for Statistical Interaction

Causal comparisons for interaction assessment entail contrasts between outcomes in states of presence or absence of the intercorrelation between variables x_1 and x_2 that can be answered from an extension of the classical causality model of Rubin (1974) and Holland and Rubin (1980). The generalized Neyman-Rubin's causal model for interaction assessment then considers the difference

$$\tau(x_1, x_2, \theta|\sigma_{12}) - \tau(x_1, x_2, \theta|\sigma_{12} = 0)$$

holding all other factors including variables x_1 and x_2 and parameters $\theta - \{\sigma_{12}\}$ unchanged. To see if intercorrelation parameter σ_{12} causes effect on the defined outcome quantities, we let $\theta' = (1, 1, 1, 2, 2, 2, 0.7, 0.7, \sigma_{12})$ and compute some true values of outcome quantities under $\sigma = 0.5$ and $\sigma_{12} = 0$ for verification. The results are displayed in Table 1.

Table 1. Effects of intercorrelation on some outcome quantities

Effect quantity	$\sigma_{12} = 0$	$\sigma_{12} = 0.5$
$\tau_{cv}(\theta)$	1.51	1.61
$\beta(\theta)$	$(0.3, 0.35, 0.35)$	$(0.44, 0.28, 0.28)$
$\beta(\gamma, \theta), \gamma = 0.1$	-1.27	-1.18
$\gamma = 0.2$	-0.73	-0.62
$\gamma = 0.3$	-0.34	-0.22
$\gamma = 0.4$	-0.01	0.11
$\gamma = 0.5$	0.30	0.44

$*$ The second and third elements for $\beta(\gamma, \theta)$ and $\beta(\theta)$ are identical in either case of $\sigma_{12} = 0$ and $\sigma_{12} = 0.5$.

The fact that $\tau_{cv}(\theta|\sigma_{12} = 0.5) = 1.61$ that is not equal to $\tau_{cv}(\theta|\sigma_{12} = 0) = 1.51$ reveals that effect of intercorrelation does exist on the mean of outcome variable. This statistical interaction theoretically can not be detected by the test for synergism of Prentice and Kalbfleisch (1988). Its effect on regression parameters and regression quantiles give the same conclusion.

The comparison results shown in Table 1 also indicate that the approach of Neyman-Rubin causal model is appropriate for assessment of regression interaction.

3. The Neyman-Rubin's Causal Model for Interaction Analysis

The classical versions giving regression interactions falling short of formalism necessary for rigorous logical analysis. The usage of Neyman-Rubin's causal model of two treatment levels matches to measure clean effect of intercorrelation for defining statistical interaction.

Definition 3.1. (a) We define the following difference

$$\tau^{+}(x_1, x_2, \theta) = \tau(x_1, x_2, \theta|\sigma_{12}) - \tau(x_1, x_2, \theta|\sigma_{12} = 0), (x_1, x_2)' \in R^2,$$

as the Neyman-Rubin's causal effect of intercorrelation for outcome quantity $\tau(x_1, x_2, \theta)$ where θ is vector true parameters.

(b) We say that Rubin's statistical interaction for outcome quantity $\tau(x_1, x_2, \theta)$ exists if there are (x_1, x_2) such that its causal effect of intercorrelation $\tau^+(x_1, x_2, \theta)$ is not zero (vector).

Here $\tau(x_1, x_2, \theta|\sigma_{12} = 0)$ represents the no-interaction response surface for outcome quantity $\tau(x_1, x_2, \theta)$. Unlike many causal models in statistics are incomplete guides to interpreting data or for suggesting answers to particular policy questions (Heckman (2008)), this causal model clearly specify the mechanism determing how hypothetical interventions are implemented. We explore the interaction assessment in detail while the others can be done analogously.

We denote $\sigma_0 = \sigma_1^2\sigma_2^2 - \sigma_{12}^2$,

$$\beta_0^+(\theta) = \frac{\sigma_{12}(-\sigma_{y1}\sigma_{12} + \sigma_{y2}\sigma_1^2)}{\sigma_1^2\sigma_0}\mu_1 + \frac{\sigma_{12}(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})}{\sigma_2^2\sigma_0}\mu_2,$$

$$\beta_1^+(\theta) = -\frac{\sigma_{12}(-\sigma_{y1}\sigma_{12} + \sigma_{y2}\sigma_1^2)}{\sigma_1^2\sigma_0},\ \beta_2^+(\theta) = -\frac{\sigma_{12}(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})}{\sigma_2^2\sigma_0}.$$

This help in formulating the excess effects of several interaction quantities.

Theorem 3.2. The Neyman-Rubin's causal effect of intercorrelation for some outcome quantities are:

(1) Conditional mean at (x_1, x_2): $\tau_{cm}^+(x_1, x_2, \theta) = (1, x_1, x_2)\beta^+(\theta)$.

(2) Regression parameters: $\beta^+(\theta) = (\beta_0^+(\theta), \beta_1^+(\theta), \beta_2^+(\theta))'$.

(3) Outcome variable: $y^+(x_1, x_2, \theta) = \tau_{cm}^+(x_1, x_2, \theta)$

(4) Conditional variance: $\tau_{cv}^+(\theta) = \frac{\sigma_{y1}^2}{\sigma_1^2} + \frac{\sigma_{y2}^2}{\sigma_2^2} - (\sigma_{y1}, \sigma_{y2})\begin{pmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{pmatrix}^{-1}\begin{pmatrix}\sigma_{y1}\\ \sigma_{y2}\end{pmatrix}$.

(5) Regression quantile: $\beta^+(\gamma, \theta) = (\beta_0^+(\theta) + z_\gamma(\sqrt{\tau_{cv}(\theta)} - \sqrt{\tau_{cv}(\theta|\sigma_{12} = 0)}), \beta_1^+(\theta), \beta_2^+(\theta))'$

(6) Conditional quantile: $\tau_{cq}^+(x_1, x_2, \theta) = (1, x_1, x_2)\beta^+(\gamma, \theta)$

(7) Reference charts: $\tau_{rc}^+(\theta) = \{(1, x_1, x_2)\beta^+(\gamma, \theta) : \gamma \in (0, 1), (x_1, x_2)' \in R^2\}$.

Proof. It is seen that $y^+(x_1, x_2, \theta) = y(x_1, x_2, \theta|\sigma_{12}) - y(x_1, x_2, \theta|\sigma_{12} = 0)$. Then regression model (2.2) indicates that $y^+(x_1, x_2, \theta) = (1, x_1, x_2)(\beta(\theta|\sigma_{12}) - \beta(\theta|\sigma_{12} = 0)) = \tau_{cm}^+(x_1, x_2, \theta)$. The others are straight forward. □

The parametrization for a Neyman-Rubin's causal model is novel that generates several advantages:

(a) It makes no controversy of model dependence occurred in classical statistical interaction assessment (Mantel et al. (1977)).

(b) The approach of causality analysis for interaction assessment resolves the concern by Greenland (1993) to measure clear effect of intercorrelation that makes the users unconfused for how much and in what direction we can learn from a data set.

(c) It creates a framework for effect of the cause to be shown with proper perspective in theoretical analysis of causality (Holland (1986)).

(d) The Neyman-Rubin's causal effect of intercorrelation model then is

$$y^+(x_1, x_2, \theta) = y(x_{1b}, x_2, \theta|\sigma_{12}) - y(x_{1a}, x_2, \theta|\sigma_{12} = 0) \tag{3.1}$$

where, like the classical Neyman-Rubin model, we can only measure one observation, here is $y(x_{1b}, x_2, \theta|\sigma_{12}) = y(x_{1b}, x_2, \theta)$. Interestingly the parametrization leads the causality analysis requiring only estimators of distributional parameters θ. This makes this causality analysis much more easier than the classical Neyman-Rubin's causality analysis.

For statistical inferences of unknown Neyman-Rubin's causal effect, we assume that we have a random sample $(y_i, x_{1i}, x_{2i})'$, $i = 1..., n$ from distributional model (2.1), we denote sample means $\begin{pmatrix} \bar{y} \\ \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} = \frac{1}{n}\sum_{i=1}^n \begin{pmatrix} y_i \\ x_{1i} \\ x_{2i} \end{pmatrix}$, sample covariance matrix $\begin{pmatrix} s_y^2 & s_{y1} & s_{y2} \\ s_{1y} & s_1^2 & s_{12} \\ s_{2y} & s_{21} & s_2^2 \end{pmatrix} = \frac{1}{n-1}\sum_{i=1}^n \begin{pmatrix} y_i - \bar{y} \\ x_{1i} - \bar{x}_1 \\ x_{2i} - \bar{x}_2 \end{pmatrix} \begin{pmatrix} y_i - \bar{y} \\ x_{1i} - \bar{x}_1 \\ x_{2i} - \bar{x}_2 \end{pmatrix}'$.

The mle of θ is $\hat{\theta}_{mle} = (\bar{y}, \bar{x}_1, \bar{x}_2, s_y^2, s_1^2, s_2^2, s_{y1}, s_{y2}, s_{12})'$. We define statistic $\hat{\tau}^+(x_1, x_2, \theta) = \tau^+(x_1, x_2, \hat{\theta}_{mle})$ as the mle of Neyman-Rubin's causal effect of intercorrelation $\tau^+(x_1, x_2, \theta)$. The asymptitic theory of this mle is a direct result of the asymptotic theory of the mle of θ.

Theorem 3.3. The random quantity $n^{1/2}(\hat{\tau}^+(x_1, x_2, \theta) - \tau^+(x_1, x_2, \theta))$ converges to $N(0, \Sigma_{\tau(\theta)})$ with $\Sigma_{\tau(\theta)}(x_1, x_2) = \frac{\partial \tau^+(x_1,x_2,\theta)}{\partial \theta'} V_\theta \frac{\partial \tau^+(x_1,x_2,\theta)}{\partial \theta}$ where $\frac{\partial \tau^+(\theta)}{\partial \theta}$ is the partial derivative of $\tau^+(\theta)$ with respect to θ and V_θ is the Cramer-Rao lower bound of the regression parameters θ.

We further denote

$$\hat\beta_0^+ = \frac{s_{12}(-s_{y1}s_{12}+s_{y2}s_1^2)}{s_1^2 s_0}\bar x_1 + \frac{s_{12}(s_{y1}s_2^2 - s_{y2}s_{12})}{s_2^2 s_0}\bar x_2$$

$$\hat\beta_1^+ = -\frac{s_{12}(-s_{y1}s_{12}+s_{y2}s_1^2)}{s_1^2 s_0},\ \hat\beta_2^+ = -\frac{s_{12}(s_{y1}s_2^2 - s_{y2}s_{12})}{s_2^2 s_0}.$$

We then have mles of the Neyman-Rubin's causal effects for some outcome quantities (parameters):

(1) Conditional mean at (x_1, x_2): $\hat\tau_{cm}^+(x_1,x_2) = (1, x_1, x_2)\hat\beta^+(\theta)$.

(2) Regression parameters: $\hat\beta^+(\theta) = (\hat\beta_0^+(\theta), \hat\beta_1^+(\theta), \hat\beta_2^+(\theta))'$.

(3) Outcome variable: $\hat y^+(\theta) = \hat\tau_{cm}^+(x_1, x_2, \theta)$

(4) Conditional variance: $\hat\tau_{cv}^+(\theta) = \frac{s_{y1}^2}{s_1^2} + \frac{s_{y2}^2}{s_2^2} - (s_{y1}, s_{y2})\begin{pmatrix} s_1^2 & s_{12} \\ s_{12} & s_2^2 \end{pmatrix}^{-1}\begin{pmatrix} s_{y1} \\ s_{y2}\end{pmatrix}$.

(5) Regression quantile: $\hat\beta^+(\gamma,\theta) = (\hat\beta_0^+(\theta) + z_\gamma(\sqrt{\hat\tau_{cv}(\theta|\sigma_{12})} - \sqrt{\hat\tau_{cv}(\theta|\sigma_{12}=0)}), \hat\beta_1^+(\theta), \hat\beta_2^+(\theta))'$.

(6) Conditional quantile at (x_1, x_2): $\hat\tau_{cq}^+(x_1, x_2) = (1, x_1, x_2)\hat\beta^+(\gamma, \theta)$

(7) Reference charts: $\hat\tau_{rc}^+ = \{(1, x_1, x_2)\hat\beta^+(\gamma,\theta) : \gamma \in (0,1), (x_1, x_2)' \in R^2\}$.

Here the Neyman-Rubin's causal effects for outcome variable and conditional mean are identical and their estimators also have the same asymptotic distribution. Advanced statistical inferences for Neyman-Rubin's causal effect $\tau^+(x_1, x_2, \theta)$ can be developed when partial derivative $\frac{\partial \tau^+(x_1,x_2,\theta)}{\partial\theta}$ is derived. The partial derivatives of causal effect for regression parameters and it for conditional variance help the formulation of asymptotic distribution of mle's of some causal effect estimators.

For designing tests and evaluations of them, we list the derived matrices of partial derivatives in Appendix:

(a) Partial derivative of $\beta^+(\theta, \sigma_{12})$ w.r.t. θ': $\frac{\partial\beta^+(\theta)}{\partial\theta'} = \begin{pmatrix} \frac{\partial\beta_0^+(\theta)}{\partial\theta'} \\ \frac{\partial\beta_1^+(\theta)}{\partial\theta'} \\ \frac{\partial\beta_2^+(\theta)}{\partial\theta'} \end{pmatrix}$, a 3×9 matrix (see (a1) of Appendix A).

(b) Partial derivative of $\tau_{cv}^+(\theta)$ w.r.t. θ': $\frac{\partial\tau_{cv}^+(\theta)}{\partial\theta'}$ (see (b1) of Appendix B).

(c) Partial derivative of Conditional quantile w.r.t. θ': $\frac{\partial\beta^+(\gamma,\theta)}{\partial\theta'}$ (see (c1) of Appendix C)

For partial derivative of conditional mean w.r.t. θ', it is $\frac{\partial\tau_{cm}^+(x_1,x_2,\theta)}{\partial\theta'} = (1, x_1, x_2)\frac{\partial\beta^+(\theta)}{\partial\theta'}$

4. Statistical Inferences for Neyman-Rubin's Causal Effect of Intercorrelation

Assessment of statistical interaction on any outcome quantity $\tau(x_1, x_2, \theta)$ can be done by statistical inferences for its Neyman-Rubin's causal effect $\tau^+(x_1, x_2, \theta)$. We consider a simulation with significance level $\alpha = 0.05$ to verify the power of the mle based test. Suppose that we have test statistic T for testing hypothesis $H_0 : \tau^+(x_1, x_2, \theta) = 0$ vs $H_1 : \tau^+(x_1, x_2, \theta) \neq 0$ and T^j represents its value at jth replication. We then define the power function as

$$p = \frac{1}{m} \sum_{j=1}^{m} I(T^j \geq t_\alpha)$$

where t_α is the simulated constant so that the probabilities at various designs to be close to α. In our studies, we let $m = 10,000$ and choose $\alpha = 0.05$.

First we consider conditional mean by testing hypothesis $H_0 : \tau^+_{cm}(x_1, x_2, \theta) = 0$ vs $H_1 : \tau^+_{cm}(x_1 x_2, \theta) \neq 0$. We let $\frac{\hat{\partial}\tau^+_{cm}(x_1,x_2,\theta)}{\partial\theta'}$ and $\hat{V}_\theta = V_{\hat{\theta}_{mle}}$ be mle's of $\frac{\partial\tau^+_{cm}(x_1,x_2,\theta)}{\partial\theta'}$ and V_θ. A test for this hypothesis is defined below:

$$\text{Test 1: rejecting } H_0 \text{ if } \frac{n^{1/2}|\hat{\tau}^+_{cm}(x_1, x_2)|}{\sqrt{\frac{\hat{\partial}\tau^+_{cm}(x_1,x_2,\theta))}{\partial\theta'} \hat{V}_\theta \frac{\hat{\partial}\tau^+_{cm}(x_1,x_2,\theta))}{\partial\theta}}} \geq t_\alpha.$$

The simulated powers when $(x_1, x_2) = (2, 2)$ are displayed in Table 2.

Table 2. Power performance for interaction detection by Test 1

n	$n = 30$	$n = 50$	$n = 100$
$\sigma_{12} = 0$	0.048	0.05	0.052
$\sigma_{12} = -0.2$	0.081	0.101	0.165
$\sigma_{12} = -0.5$	0.299	0.46	0.733
$\sigma_{12} = -0.8$	0.701	0.882	0.988
$\sigma_{12} = 0.2$	0.078	0.101	0.161
$\sigma_{12} = 0.5$	0.196	0.361	0.683
$\sigma_{12} = 0.8$	0.391	0.669	0.967

This test involves asymptotic variance estimate of interaction that is influence by all parameters. We may consider the partial influence of it influenced by only covariance σ_{12}.

We consider the test constructing the test statistic based on the sacle of variance change due to parameter σ_{12} that defines the following test:

$$\text{Test 2: rejecting } H_0 \text{ if } \frac{n^{1/2}|\hat{\tau}^+_{cm}(x_1,x_2)|}{\sqrt{\frac{\partial \tau^+_{cm}(x_1,x_2,\theta)}{\partial \sigma'_{12}} \hat{V}_\theta \frac{\partial \tau^+_{cm}(x_1,x_2,\theta)}{\partial \sigma_{12}}}} \geq t_\alpha$$

where $\frac{\partial \tau^+_{cm}(x_1,x_2,\theta)}{\partial \sigma_{12}} = (1, x_1, x_2)\frac{\partial \beta^+(\theta|\sigma_{12})}{\partial \sigma_{12}}$ with $\frac{\partial \beta^+(\theta|\sigma_{12})}{\partial \sigma_{12}}$ displaying in Appendix (a2). Hopefully this tset is more sensitive in detection a change in covariane.

Table 3. Power performance for interaction detection by Test 2

n	$n = 30$	$n = 50$	$n = 100$
$\sigma_{12} = 0$	0.049	0.052	0.050
$\sigma_{12} = -0.2$	0.080	0.110	0.169
$\sigma_{12} = -0.5$	0.233	0.429	0.719
$\sigma_{12} = -0.8$	0.428	0.815	0.988
$\sigma_{12} = 0.2$	0.077	0.108	0.167
$\sigma_{12} = 0.5$	0.234	0.420	0.710
$\sigma_{12} = 0.8$	0.434	0.817	0.987

It shows that the test considering partial derivative with respect to σ_{12} only does improve the power a bit.

We may also interest in assessment of statistical interaction for the average conditional mean on a region A for variables (x_1, x_2) as

$$\tau^+_{cm}(A|\theta) = \int_A \tau^+_{cm}(x_1,x_2,\theta) f_{12}(x_1,x_2,\theta)dx_1dx_2.$$

One interesting unknown quantity to be verified is the averaging Neyman-Rubin's causal effect $\tau^+_{cm,ave} = \tau^+_{cm}(R^2|\theta)$ which can be shown to be the Neyman-Rubin's causal effect for conditional mean at mean vector (μ_1, μ_2) as $\tau^+_{cm,ave} = \tau^+_{cm}(\mu_1, \mu_2, \theta)$. We would not go further to investigate it in simulation but we will study it in data analysis.

We now consider a test for hypothesis of Neyman-Rubin's causal effect for conditional variance. For testing hypothesis $H_0 : \tau^+_{cv}(\theta) = 0$ vs $H_1 : \tau^+_{cv}(\theta) \neq 0$ for existence of statistical interaction for conditional variance, we define $\hat{\tau}^{+}_{cv}(\theta) = \tau^{+}_{cv}(\hat{\theta}_{mle})$ and the test is defined as

$$\text{rejecting } H_0 \text{ if } \frac{\sqrt{n}|\hat{\tau}^{+}_{cv}(\theta)|}{\sqrt{\frac{\partial \hat{\tau}^{+}_{cv}(\theta)}{\partial \theta'}\hat{V}_\theta \frac{\partial \hat{\tau}^{+}_{cv}(\theta)}{\partial \theta}}} \geq t_\alpha$$

Table 4. Power performance for statistical interction for conditional variance

n	$n = 30$	$n = 50$	$n = 100$
$\sigma_{12} = 0$	0.051	0.056	0.049
$\sigma_{12} = -0.2$	0.116	0.165	0.247
$\sigma_{12} = -0.5$	0.378	0.592	0.844
$\sigma_{12} = -0.8$	0.749	0.932	0.997

We next consider the test for detection of statistical interaction on the regression quantile $\beta(\gamma, \theta)$. We test hypothesis $H_0 : \beta^{+}(\gamma, \theta) = 0$ vs $H_1 : \beta^{+}(\gamma, \theta) \neq 0$ by setting the following test:

$$\text{rejecting } H_0 \text{ if } n\hat{\beta}^{+}(\gamma, \theta)'\left(\frac{\partial \hat{\beta}^{+}(\gamma, \theta)}{\partial \theta'}\hat{V}_\theta \frac{\partial \hat{\beta}^{+}(\gamma, \theta)}{\partial \theta}\right)^{-1}\hat{\beta}^{+}(\gamma, \theta) \geq t_\alpha$$

Table 5. Power performance for statistical interction for regression quantile

n	$n = 30$	$n = 50$	$n = 100$
$\gamma = 0.7$			
$\sigma_{12} = 0$	0.055	0.048	0.048
$\sigma_{12} = -0.2$	0.070	0.093	0.160
$\sigma_{12} = -0.5$	0.267	0.440	0.718
$\sigma_{12} = -0.8$	0.655	0.855	0.989
$\gamma = 0.8$			
$\sigma_{12} = 0$	0.053	0.049	0.054
$\sigma_{12} = -0.2$	0.069	0.088	0.145
$\sigma_{12} = -0.5$	0.253	0.417	0.724
$\sigma_{12} = -0.8$	0.662	0.861	0.992
$\gamma = 0.9$			
$\sigma_{12} = -0.2$	0.056	0.073	0.143
$\sigma_{12} = -0.5$	0.207	0.396	0.704
$\sigma_{12} = -0.8$	0.615	0.853	0.988

The power performance of assessing statistical interactions for conditional variance and regression quantile by testing for hypotheses for their corresponding Neyman-Rubin's causal effects shows these tests are satisfactory.

We consider a real data analysis for further interpretation. The sales data of size 25 from 25 territories for 1994 is available in Dielman (1996) that includes the sales (response variable) y (in US$1000), the amount the company spent on advertising (explanatory variables x_1) and the total amount of bonuses paid (x_2). The estimated regression model by least squares method computed in Dielman (1996) is

$$\hat{y} = -516.4 + 2.47x_1 + 1.86x_2$$

with $R^2 = 85.5\%$. Management department of the Meddicop Company concerned if the cause of bonuses (x_2) paid in 1994 to salesmen is related to the response of sales. We further evaluate if correlation plays a role of affecting the sale outcomes (y) by computing the estimated sales $\hat{\tau}_{cm}(x_1, x_2, \theta) = (1, x_1, x_2)\beta(\hat{\theta}_{mle}|\sigma_{12})$ for (x_1, x_2) in territories 1, 2, 3 and 4 (T_1, T_2, T_3 and T_4) based on mle's of $\hat{\theta}$ except that covariance σ_{12} is replaced by some specified values including its s_{12}. The estimated mean sales ($\hat{\tau}_{cm}(x_1, x_2, \theta)$) are listed in Table 6.

Table 6. Predicted mean sales for first four territories

σ_{12}	$(x_1, x_2) = (-1.89, -1.59)$	$(-1.43, -1.40)$	$(-1.35, -0.17)$	$(-0.892, 0.522)$
$\sigma_{12} = 0$	673.935	792.929	969.53	1153.362
$\sigma_{12} = -0.2$	529.926	674.309	913.38	1146.779
$\sigma_{12} = -0.5$	94.089	318.129	729.952	1109.164
$\sigma_{12} = -0.8$	-1655.61	-1107.133	-28.724	928.929
$\sigma_{12} = 0.2$	768.354	871.874	1000.722	1150.307
$\sigma_{12} = 0.5$	857.058	950.334	1009.448	1120.455
$\sigma_{12} = 0.8$	887.623	1000.22	902.930	966.567
$s_{12} = 0.419$	837.937	932.435	1012.292	1133.086

The predicted mean sales with $\sigma_{12} = 0$ present the outcome results of no-intercorrelation. Then the predicted mean sales for $\sigma_{12} \neq 0$, in terms of (x_1, x_2), are different from it for $\sigma_{12} = 0$ and regression parameters $\beta(\theta)$ for $\sigma_{12} \neq 0$ are also different vectors from it of $\sigma_{12} = 0$. These results show that intercorrelation does making influence on conditional mean and regression parameters. We then are appropriate to consider them as potential outcome quantities.

Table 7. p-values of observation points

Obs #	Z-test	t-test	Obs #	Z-test	t-test
1	0.019*	0.028*	14	0.028*	0.038*
2	0.021*	0.030*	15	0.028*	0.038*
3	0.109	0.122	16	0.818	0.820
4	0.402	0.410	17	0.895	0.896
5	0.381	0.390	18	0.215	0.227
6	0.020*	0.028*	19	0.024*	0.034*
7	0.032*	0.043*	20	0.080	0.093
8	0.108	0.121	21	0.430	0.438
9	0.030*	0.040*	22	0.053	0.065
10	0.359	0.368	23	0.075	0.088
11	0.036*	0.046*	24	0.055	0.067
12	0.750	0.753	25	0.429	0.436
13	0.024*	0.034*			

5. The Neyman-Rubin's Causal effect of Intercorrelation for Binary Variable

Ai and Norton (2003) considered the second-order interaction verification for probit and logic regression models. It is interesting to see if we can assess effect of intercorrelation for binary response variable.

It is very common that the categorical dependent variable is observed from categorization of a continuous explanatory variable while Prock et al. (2004) reported that 84% of epidemiological papers from leading journals made categorization of continuous variables. This categorization is widespread from epidemiology to other areas such as psychology (MacCallum, et al. (2002)) and marketing (Irwin and McClelland (2003)). With categorization, we are allowed to apply parametrization to assess statistical interaction for categorical dependent variables.

Again, we let Y, X_1 and X_2 be continuous random variables with a joint distribution. One categorization is to set a binary variable $I(Y \leq \lambda)$ with outcome quantity as the regression function of the conditional mean defined as $p(x_1, x_2) = E(I(Y \leq \lambda)|X_1 = x_1, X_2 = x_2)$, a probability as a function of (x_1, x_2). For assessment of statistical interaction, most applied scientists consider the model-dependent outcome quantity in the framework of logistic regression as

$$p_{log}(x_1, x_2) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2)}} \tag{5.1}$$

or of probit regression as

$$p_{prob}(x_1, x_2) = \Phi(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2) \tag{5.2}$$

where $\Phi(.)$ is the distribution function of the standard normal distribution and consider a test for hypothesis $H_0 : \beta_{12} = 0$.

Now, we also consider that these random variables follow the normal distribution of (2.1). We then easily obtain the following theorem.

Theorem 5.1. The outcome quantity of conditional mean for binary variable $I(Y \leq \lambda)$ under normality assumption (2.1) is

$$p_{cat}(x_1, x_2) = \Phi(\frac{\lambda - (1, x_1, x_2)\beta(\theta)}{\sqrt{\tau_{cv}(\theta)}}), \tag{5.3}$$

called the probit outcome quantity, where regression coefficients $\beta_0(\theta), \beta_1(\theta)$ and $\beta_2(\theta)$ and error conditional variance $\sigma^2_{y|x_1,x_2}$ are denoted in Theorem 2.1.

The true outcome quantity under normality assumption in (5.3) indicates that either probit one with product term of (5.2) and logistic one are all inappropriate. Following Theorem 3.3, we have the following theorem.

Theorem 5.2. The Neyman-Rubin's causal effect of intercorrelation for probit outcome quantity $p_{cat}(x_1, x_2)$ is

$$p^{+}_{cat}(x_1, x_2) = \Phi(\frac{\lambda - (1, x_1, x_2)\beta(\theta|\sigma_{12})}{\sqrt{\tau_{cv}(\theta|\sigma_{12})}}) - \Phi(\frac{\lambda - (1, x_1, x_2)\beta(\theta|\sigma_{12} = 0)}{\sqrt{\tau_{cv}(\theta|\sigma_{12} = 0)}}).$$

The mle of the Neyman-Rubin's causal effect of intercorrelation is

$$\hat{p}^{+}_{cat}(x_1, x_2) = \Phi(\frac{\lambda - (1, x_1, x_2)\hat{\beta}(\theta|\sigma_{12})}{\sqrt{\hat{\tau}_{cv}(\theta|\sigma_{12})}}) - \Phi(\frac{\lambda - (1, x_1, x_2)\hat{\beta}(\theta|\sigma_{12} = 0)}{\sqrt{\hat{\tau}_{cv}(\theta|\sigma_{12} = 0)}}).$$

that leads to the following theorem.

Theorem 5.3. The random quantity $n^{1/2}(\hat{p}^+_{cat}(x_1,x_2) - p^+_{cat}(x_1,x_2))$ converges to $N(0, \frac{\partial p^+_{cat}(x_1,x_2)}{\partial\theta'} V_\theta \frac{\partial p^+_{cat}(x_1,x_2)}{\partial\theta})$ with

$$\frac{\partial p^+_{cat}(x_1,x_2)}{\partial\theta'} = -\phi(\frac{\lambda-(1,x_1,x_2)\beta(\theta|\sigma_{12})}{\sqrt{\tau_{cv}(\theta|\sigma_{12})}})$$
$$\frac{\sqrt{\tau_{cv}(\theta|\sigma_{12})}(1,x_1,x_2)\frac{\partial\beta(\theta|\sigma_{12})}{\partial\theta'} + (\lambda-(1,x_1,x_2)\beta(\theta|\sigma_{12}))\frac{1}{2}\frac{\partial\tau_{cv}(\theta|\sigma_{12})}{\partial\theta'}}{\tau_{cv}(\theta|\sigma_{12})}$$
$$+\phi(\frac{\lambda-(1,x_1,x_2)\beta(\theta|\sigma_{12}=0)}{\sqrt{\tau_{cv}(\theta|\sigma_{12}=0)}})$$
$$\frac{\sqrt{\tau_{cv}(\theta|\sigma_{12}=0)}(1,x_1,x_2)\frac{\partial\beta(\theta|\sigma_{12}=0)}{\partial\theta'} + (\lambda-(1,x_1,x_2)\beta(\theta|\sigma_{12}=0))\frac{1}{2}\frac{\partial\tau_{cv}(\theta|\sigma_{12}=0)}{\partial\theta'}}{\tau_{cv}(\theta|\sigma_{12}=0)}$$

where $\frac{\partial\tau_{cv}(\theta|\sigma_{12})}{\partial\theta'}$ and $\frac{\partial\tau_{cv}(\theta|\sigma_{12}=0)}{\partial\theta'}$ are displayed in Appendix (b2) and (b3) and $\frac{\partial\beta(\theta|\sigma_{12})}{\partial\theta'}$ and $\frac{\partial\beta(\theta|\sigma_{12}=0)}{\partial\theta'}$ are displayed in Appendix (a3) and (a4).

Different underlying distribution or binary variable results different outcome quantity and their corresponding Neyman-Rubin's causal effect of intercorrelation. We would not go further on it.

6. Concluding Remarks

Whether or not effects of explanatory variables are intercorrelated are frequently assessed with ambiguous and controversial concept of statistical interaction give the practitioners limited and confused view of the nature of interaction in statistical world. We attempt here to elucidate some of the controversial issues surrounding the concept of statistical interaction with systematic introduction of definitions, methods and theorems to build the Neyman-Rubin's causality analysis for assessment of interaction effect of intercorrelation. The parametrization of constructing regression model formulated from a multivariate distribution brings a theoretical foundation in connecting a causal model with the interaction cause of intercorrelation parameter that allows us to measure the effect of intercorrelation. Hopefully this would be recognized to have made permanent contribution in assessment of statistical interaction. We have several further conclusions:

(a) Suppose that the conditional mean of the response variable y given

$X_1 = x_1$ and $X_2 = x_2$ is an exponential function as

$$\mu(x_1, x_2) = exp\{\beta_0^*(\theta) + \beta_1^*(\theta)x_1 + \beta_2^*(\theta)x_2 + \beta_{12}^*(\theta)x_1x_2\}$$

where regression parameters $\beta_j^*(\theta), j = 0, 1, 2$ and $\beta_{12}^*(\theta)$ are function of distributional parameters θ. Let us evaluate the approach of Ai and Norton (2003) considering the cross derivative as

$$\frac{\partial^2 \mu(x_1, x_2)}{\partial x_1 \partial x_2} = \mu(x_1, x_2)[\beta_{12}^*(\theta) + (\beta_1^*(\theta) + \beta_{12}^*(\theta)x_2)(\beta_2^*(\theta) + \beta_{12}^*(\theta)x_1)]$$

for assessing statistical interaction. This interaction parameter involves complicated function of distributional parameters θ that is entirely dependent on intercorrelation parameter only if all regression parameters depend on covariance between X_1 and X_2 are zeros. This is generally not true.

(b) Any utility function $g(Y)$ that its explicit form of conditional mean $\mu_g(x_1, x_2, \theta) = E(g(Y)|X_1 = x_1, X_2 = x_2)$ available serves an interaction parameter for assessing its statistical interaction. Simpler form of other categorization such as $I(Y \geq \lambda), I(\lambda_1 \leq Y \leq \lambda_2)$ and $I(Y \leq \lambda_1$ or $\geq \lambda_2)$ are candidates.

(c) (Prediction interval) Again, we are interesting in the prediction of response y_0 when $X_1 = x_{10}$ and $X_2 = x_{20}$ are specified. The interest of interaction parameter is the socalled naive coverage interval $(\mu_{norm}(x_{10}, x_{20}) - z_\gamma \sigma_{y|x_1,x_2}, \mu_{norm}(x_{10}, x_{20}) + z_\gamma \sigma_{y|x_1,x_2})$.

7. Appendix

Appendix A

(a1) The partial derivatives for elements of $\beta^+(\theta)$ w.r.t. θ:

$$\frac{\partial \beta_0^+(\theta)}{\partial \theta'} =$$
$$\begin{pmatrix} 0 & -\sigma_0^{-1}(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12}) + \frac{\sigma_{y1}}{\sigma_1^2} & \sigma_0^{-1}(\sigma_{y1}\sigma_{12} + \sigma_{y2}\sigma_1^2) + \frac{\sigma_{y2}}{\sigma_2^2} & 0 & d_{15} & d_{16} & d_{17} & d_{18} & d_{19} \end{pmatrix}$$

where

$$d_{15} = -\sigma_0^{-1}\sigma_{y2}\mu_2 + \sigma_0^{-2}[(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})\sigma_2^2\mu_1 + (-\sigma_{y1}\sigma_{12} + \sigma_{y2}\sigma_1^2)\sigma_2^2\mu_2] - \frac{\sigma_{y1}\mu_1}{\sigma_1^4}$$
$$d_{16} = -\sigma_0^{-1}\sigma_{y1}\mu_1 + \sigma_0^{-2}[(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})\sigma_1^2\mu_1 + (-\sigma_{y1}\sigma_{12} + \sigma_{y2}\sigma_1^2)\sigma_1^2\mu_2] - \frac{\sigma_{y2}\mu_2}{\sigma_2^4}$$
$$d_{17} = \sigma_0^{-1}[-\sigma_2^2\mu_1 + \sigma_{12}\mu_2] + \frac{\mu_1}{\sigma_1^2},\ d_{18} = \sigma_0^{-1}[\sigma_{12}\mu_1 - \sigma_1^2\mu_2] + \frac{\mu_2}{\sigma_2^2}$$
$$d_{19} = \sigma_0^{-1}(\sigma_{y2}\mu_1 + \sigma_{y1}\mu_2) - 2\sigma_0^{-2}[(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})\sigma_{12}\mu_1 + (-\sigma_{y1}\sigma_{12} + \sigma_{y2}\sigma_1^2)\sigma_{12}\mu_2]$$

$$\frac{\partial\beta_1^+(\theta)}{\partial\theta'} = (\,0 \quad 0 \quad 0 \quad 0 \quad -\sigma_0^{-2}(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})\sigma_2^2 + \tfrac{\sigma_{y1}}{\sigma_1^4} \quad d_{26} \quad d_{27} \quad d_{28} \quad d_{29}\,)$$

where

$$d_{26} = \sigma_0^{-1}\sigma_{y1} - \sigma_0^{-2}(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})\sigma_1^2,\ d_{27} = \sigma_0^{-1}\sigma_2^2 - \sigma_1^{-2}$$
$$d_{28} = -\sigma_0^{-1}\sigma_{12},\ d_{29} = -\sigma_0^{-1}\sigma_{y2} + 2\sigma_0^{-2}(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})\sigma_{12}$$

$$\frac{\partial\beta_2^+(\theta)}{\partial\theta'} = (\,0 \quad 0 \quad 0 \quad 0 \quad d_{35} \quad d_{36} \quad -\sigma_0^{-1}\sigma_{12} \quad \sigma_0^{-1}\sigma_1^2 - \sigma_2^{-2} \quad d_{39}\,)$$

where

$$d_{35} = \sigma_0^{-1}\sigma_{y2} + \sigma_0^{-2}(\sigma_{y1}\sigma_{12} - \sigma_{y2}\sigma_1^2)\sigma_2^2,\ d_{36} = \sigma_0^{-2}(\sigma_{y1}\sigma_{12} - \sigma_{y2}\sigma_1^2)\sigma_1^2 + \sigma_{y2}\sigma_2^{-2}$$
$$d_{39} = -\sigma_0^{-1}\sigma_{y1} - 2\sigma_0^{-2}(\sigma_{y1}\sigma_{12} - \sigma_{y2}\sigma_1^2)\sigma_{12}$$

(a2) The partial derivatives of elements of $\beta^+(\theta)$ w.r.t. σ_{12}: $\frac{\partial\beta^+(\theta|\sigma_{12})}{\partial\sigma_{12}} = \begin{pmatrix} 0_8' & \frac{\partial\beta_0^+(\theta)}{\partial\sigma_{12}} \\ 0_8' & \frac{\partial\beta_1^+(\theta)}{\partial\sigma_{12}} \\ 0_8' & \frac{\partial\beta_2^+(\theta)}{\partial\sigma_{12}} \end{pmatrix}$

(a3) The partial derivatives of elements of $\beta(\theta)$ w.r.t. θ: $\frac{\partial\beta(\theta|\sigma_{12})}{\partial\theta'} = \begin{pmatrix} G_1 \\ G_2 \\ G_3 \end{pmatrix}$

where

$$G_1 = (\,1 \quad -\tfrac{\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12}}{\sigma_0} \quad -\tfrac{\sigma_{y2}\sigma_1^2 - \sigma_{y1}\sigma_{12}}{\sigma_0} \quad 0 \quad g_5 \quad g_6 \quad g_7 \quad g_8 \quad g_9\,)$$
$$G_2 = (\,0 \quad 0 \quad 0 \quad 0 \quad h_5 \quad h_6 \quad \tfrac{\sigma_2^2}{\sigma_0} \quad -\tfrac{\sigma_{12}}{\sigma_0} \quad h_9\,)$$
$$G_3 = (\,0 \quad 0 \quad 0 \quad 0 \quad k_5 \quad k_6 \quad -\tfrac{\sigma_{12}}{\sigma_0} \quad \tfrac{\sigma_1^2}{\sigma_0} \quad k_9\,)$$

with

$$g_5 = -\frac{\sigma_{y2}\mu_2}{\sigma_0} + \frac{(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})\mu_1\sigma_2^2}{\sigma_0^2} + \frac{(\sigma_{y2}\sigma_1^2 - \sigma_{y1}\sigma_{12})\mu_2\sigma_2^2}{\sigma_0^2}$$

$$g_6 = -\frac{\sigma_{y1}\mu_1}{\sigma_0} + \frac{(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})\mu_1\sigma_1^2}{\sigma_0^2} + \frac{(\sigma_{y2}\sigma_1^2 - \sigma_{y1}\sigma_{12})\mu_2\sigma_1^2}{\sigma_0^2}$$

$$g_7 = \frac{\sigma_{12}\mu_2 - \sigma_2^2\mu_1}{\sigma_0},\ g_8 = \frac{\sigma_{12}\mu_1 - \sigma_1^2\mu_2}{\sigma_0},$$

$$g_9 = \frac{\sigma_{y2}\mu_1 + \sigma_{y1}\mu_2}{\sigma_0} - 2\frac{(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})\mu_1\sigma_{12} + (\sigma_{y2}\sigma_1^2 - \sigma_{y1}\sigma_{12})\mu_2\sigma_{12}}{\sigma_0^2}$$

$$h_5 = \frac{(\sigma_{y2}\sigma_{12} - \sigma_{y1}\sigma_2^2)\sigma_2^2}{\sigma_0^2},\ h_6 = \frac{\sigma_{y1}}{\sigma_0} + \frac{(\sigma_{y2}\sigma_{12} - \sigma_{y1}\sigma_2^2)\sigma_1^2}{\sigma_0^2}$$

$$h_9 = -\frac{\sigma_{y2}}{\sigma_0} + 2\frac{(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})\sigma_{12}}{\sigma_0^2}$$

$$k_5 = \frac{\sigma_{y2}}{\sigma_0} + \frac{(\sigma_{y1}\sigma_{12} - \sigma_{y2}\sigma_1^2)\sigma_2^2}{\sigma_0^2},\ k_6 = \frac{(\sigma_{y1}\sigma_{12} - \sigma_{y2}\sigma_1^2)\sigma_1^2}{\sigma_0^2}$$

$$k_9 = -\frac{\sigma_{y1}}{\sigma_0} + 2\frac{(\sigma_{y2}\sigma_1^2 - \sigma_{y1}\sigma_{12})\sigma_{12}}{\sigma_0^2}$$

(a4) The partial derivatives of elements of $\beta(\theta)$ w.r.t. θ' if $\sigma_{12} = 0$: $\frac{\partial\beta(\theta|\sigma_{12}=0)}{\partial\theta'} = \begin{pmatrix} H_1 \\ H_2 \\ H_3 \end{pmatrix}$ where

$$H_1 = \begin{pmatrix} 1 & -\frac{\sigma_{y1}}{\sigma_1^2} & -\frac{\sigma_{y2}}{\sigma_2^2} & 0 & \frac{\sigma_{y1}\mu_1}{\sigma_1^4} & \frac{\sigma_{y2}\mu_2}{\sigma_2^4} & -\frac{\mu_1}{\sigma_1^2} & -\frac{\mu_2}{\sigma_2^2} & 0 \end{pmatrix}$$

$$H_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & -\frac{\sigma_{y1}}{\sigma_1^4} & 0 & \frac{1}{\sigma_1^4} & 0 & 0 \end{pmatrix}$$

$$H_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -\frac{\sigma_{y2}}{\sigma_2^4} & 0 & \frac{1}{\sigma_2^4} & 0 \end{pmatrix}$$

Appendix B

(b1) The partial derivative of the causal effect for conditional variance w.r.t. θ', $\tau_{cv}^+(\theta)$, is $\frac{\partial\tau_{cv}^+(\theta)}{\partial\theta'} = (0, 0, 0, 0, t_5, t_6, \frac{2\sigma_{y1}}{\sigma_1^2} - \frac{2\sigma_{y1}\sigma_2^2 + 2\sigma_{y2}\sigma_{12}}{\sigma_0}, \frac{2\sigma_{y2}}{\sigma_2^2} - \frac{2\sigma_{y1}\sigma_{12} - 2\sigma_{y2}\sigma_1^2}{\sigma_0}, t_9)$ where

$$t_5 = -\frac{\sigma_{y1}^2}{\sigma_1^4} + \frac{\sigma_{y2}^2}{\sigma_0} + \frac{\sigma_{y1}^2\sigma_2^2 + 2\sigma_{y1}\sigma_{y2}\sigma_{12}\sigma_2^2 - \sigma_{y2}^2\sigma_1^2\sigma_2^2}{\sigma_0^2}$$

$$t_6 = -\frac{\sigma_{y2}^2}{\sigma_2^4} - \frac{\sigma_{y1}^2}{\sigma_0} + \frac{\sigma_{y1}^2\sigma_1^2\sigma_2^2 + 2\sigma_{y1}\sigma_{y2}\sigma_{12}\sigma_1^2 - \sigma_{y2}^2\sigma_1^4}{\sigma_0^2}$$

$$t_9 = -\frac{2\sigma_{y1}\sigma_{y2}}{\sigma_0} - \frac{2\sigma_{y1}^2\sigma_2^2\sigma_{12} + 4\sigma_{y1}\sigma_{y2}\sigma_{12}^2 - 2\sigma_{y2}^2\sigma_1^2\sigma_{12}}{\sigma_0^2}$$

(b2) The partial derivative of $\tau_{cv}(\theta)$ w.r.t. θ': $\frac{\partial \tau_{cv}(\theta|\sigma_{12})}{\partial \theta'} = (\, 0 \quad 0 \quad 0 \quad 1 \quad u_5 \quad u_6 \quad u_7 \quad u_8 \quad u_9 \,)$ with

$$
\begin{aligned}
u_5 &= -\frac{\sigma_{y2}^2}{\sigma_0} + \frac{\sigma_{y1}^2\sigma_2^2 - 2\sigma_{y1}\sigma_{y2}\sigma_{12}\sigma_2^2 + \sigma_{y2}^2\sigma_1^2\sigma_2^2}{\sigma_0^2} \\
u_6 &= -\frac{\sigma_{y1}^2}{\sigma_0} + \frac{\sigma_{y1}^2\sigma_1^2\sigma_2^2 - 2\sigma_{y1}\sigma_{y2}\sigma_{12}\sigma_1^2 + \sigma_{y2}^2\sigma_1^4}{\sigma_0^2} \\
u_7 &= 2\frac{\sigma_{y2}\sigma_{12} - \sigma_{y1}\sigma_2^2}{\sigma_0}, \ u_8 = 2\frac{\sigma_{y1}\sigma_{12} - \sigma_{y2}\sigma_1^2}{\sigma_0} \\
u_9 &= 2\frac{\sigma_{y1}\sigma_{y2}}{\sigma_0} - 2\frac{\sigma_{y1}^2\sigma_2^2\sigma_{12} - 2\sigma_{y1}\sigma_{y2}\sigma_{12}^2 + \sigma_{y2}^2\sigma_1^2\sigma_{12}}{\sigma_0^2}
\end{aligned}
$$

(b3) $\frac{\partial \tau_{cv}(\theta|\sigma_{12}=0)}{\partial \theta'} = (\, 0 \quad 0 \quad 0 \quad 1 \quad \frac{\sigma_{y1}^2}{\sigma_1^4} \quad \frac{\sigma_{y2}^2}{\sigma_2^4} \quad -2\frac{\sigma_{y1}}{\sigma_1^2} \quad -2\frac{\sigma_{y2}}{\sigma_2^2} \quad 0 \,)$

Appendix C

(c1) $\frac{\partial \beta^+(\gamma,\theta)}{\partial \theta'} = \frac{\partial \beta^+(\theta)}{\partial \theta'} + z_\gamma \begin{pmatrix} E \\ 0_9' \\ 0_9' \end{pmatrix}$ where

$$
E = (\, 0 \quad 0 \quad 0 \quad \tfrac{1}{2}(\tfrac{1}{(\tau_{cv}(\theta|\sigma_{12}))^{1/2}} - \tfrac{1}{(\tau_{cv}(\theta|\sigma_{12}=0))^{1/2}}) \quad e_5 \quad e_6 \quad e_7 \quad e_8 \quad e_9 \,)
$$

with

$$
\begin{aligned}
e_5 &= \frac{1}{2}(\frac{(\sigma_{y1}\sigma_2^2 - \sigma_{y2}\sigma_{12})^2}{(\tau_{cv}(\theta|\sigma_{12}))^{1/2}} - \frac{\sigma_{y1}^2}{\sigma_1^4(\tau_{cv}(\theta|\sigma_{12}=0))^{1/2}}) \\
e_6 &= \frac{1}{2}(\frac{(\sigma_{y1}\sigma_{12} - \sigma_{y2}\sigma_1^2)^2}{(\tau_{cv}(\theta|\sigma_{12}))^{1/2}} - \frac{\sigma_{y2}^2}{\sigma_2^4(\tau_{cv}(\theta|\sigma_{12}=0))^{1/2}}) \\
e_7 &= \frac{\sigma_{y2}\sigma_{12} - \sigma_{y1}\sigma_2^2}{\sigma_0(\tau_{cv}(\theta|\sigma_{12}))^{1/2}} + \frac{\sigma_{y1}}{\sigma_1^2(\tau_{cv}(\theta|\sigma_{12}=0))^{1/2}} \\
e_8 &= \frac{\sigma_{y1}\sigma_{12} - \sigma_{y2}\sigma_1^2}{\sigma_0(\tau_{cv}(\theta|\sigma_{12}))^{1/2}} + \frac{\sigma_{y2}}{\sigma_2^2(\tau_{cv}(\theta|\sigma_{12}=0))^{1/2}} \\
e_9 &= \frac{(\sigma_{y2}\sigma_{12} - \sigma_{y1}\sigma_2^2)(\sigma_{y1}\sigma_{12} - \sigma_{y2}\sigma_1^2)}{(\tau_{cv}(\theta|\sigma_{12}))^{1/2}}.
\end{aligned}
$$

(c2) Cramer-Rao lower bound of the regression parameters θ: $V_\theta = \begin{pmatrix} A_\theta & 0_{3\times 6} \\ 0_{6\times 3} & B_\theta \end{pmatrix}$

with 3×3 matrix $A_\theta = \begin{pmatrix} \sigma_y^2 & \sigma_{y1} & \sigma_{y2} \\ \sigma_{y1} & \sigma_1^2 & \sigma_{12} \\ \sigma_{y2} & \sigma_{12} & \sigma_2^2 \end{pmatrix}$ and 6×6 matrix

$$B_\theta =$$
$$\begin{pmatrix} 2\sigma_y^4 & 2\sigma_{y1}^2 & 2\sigma_{y2}^2 & 2\sigma_y^2\sigma_{y1} & 2\sigma_y^2\sigma_{y2} & 2\sigma_{y1}\sigma_{y2} \\ 2\sigma_{y1}^2 & 2\sigma_1^4 & 2\sigma_{12}^2 & 2\sigma_1^2\sigma_{y1} & 2\sigma_{y1}\sigma_{12} & 2\sigma_1^2\sigma_{12} \\ 2\sigma_{y2}^2 & 2\sigma_{12}^2 & 2\sigma_2^4 & 2\sigma_{y2}\sigma_{12} & 2\sigma_2^2\sigma_{y2} & 2\sigma_2^2\sigma_{12} \\ 2\sigma_y^2\sigma_{y1} & 2\sigma_1^2\sigma_{y1} & 2\sigma_{y2}\sigma_{12} & \sigma_y^2\sigma_1^2 + \sigma_{y1}^2 & \sigma_y^2\sigma_{12} + \sigma_{y1}\sigma_{y2} & \sigma_1^2\sigma_{y2} + \sigma_{y1}\sigma_{12} \\ 2\sigma_y^2\sigma_{y2} & 2\sigma_{y1}\sigma_{12} & 2\sigma_2^2\sigma_{y2} & \sigma_y^2\sigma_{12} + \sigma_{y1}\sigma_{y2} & \sigma_y^2\sigma_2^2 + \sigma_{y2}^2 & \sigma_2^2\sigma_{y1} + \sigma_{y2}\sigma_{12} \\ 2\sigma_{y1}\sigma_{y2} & 2\sigma_1^2\sigma_{12} & 2\sigma_2^2\sigma_{12} & \sigma_1^2\sigma_{y2} + \sigma_{y1}\sigma_{12} & \sigma_2^2\sigma_{y1} + \sigma_{y2}\sigma_{12} & \sigma_1^2\sigma_2^2 + \sigma_{12}^2 \end{pmatrix}.$$

REFERENCES

Abadie, A. (2005). Semiparametric difference-in differences estimators. *Review of Econometric Studies*, **72**, 1-19.

Ai, C. and Norton, E., C. (2003). Interaction terms in logic and probit models. *Economics Letters*, **80**,123-129.

Chen, et al. (2013). Statistical Isobole for Interaction Assessment. Submitted for possible publication.

Dielman, T. E. (1996). *Applied Regression Analysis*, Second edition, Duxbury Press: New York.

Greenland, S. (1993). Basic problems in interaction assessment. *Environmental Health Persepectives Supplements*, **101**, 59-66.

Greenland, S. (2009). Interactions in epidemiology: Relevance, identification, and estimation. *Epidemiology*, **20**, 14-17.

Heckman, J. J. (2008). Econometric causality. *International Statistical Review*, **76**, 1-27.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, **81**, 945-960.

Holland, P. W. and Rubin, D. B. (1980). Causal inference in prospective and retrospective studies. *American Statistical Association Annual Meeting*.

Irwin, J. R. and McClelland, G. H. (2003). Negative consequences of dichotomizing continuous predictor variables. *Journal of Marketing Research*, **40**, 366-371.

Koenker, R. and Bassett, G.J. (1978). Regression quantiles. *Econometrica* **46**, 33-50.

Loewe, S (1928). Die Quantitation probleme der pharmakologie. *Ergeb Physiol*, **27**, 47-187.

Loewe, S (1953). The problem of synergism and antagonism of combined drugs. *Arzneimittelforschung* **3**, 285-290.

MacCallum, R. C., Zhang, S., Preacher, K. J. and Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, **7**, 19-40.

Mantel, N., Brown, C. and Byar, D. P. (1977). Tests for homogeneity of effect in an epidemiologic investigation. *Annals of Journal Epidemiology*, **106**, 125-129.

Mauderly, J. L. and Samet, J. M. (2009). Is there evidence for synergy among air pollutants in causing health effects. *Environmental Health Perspectives*, **117**, 1-6.

Mullahy, J. (1999). Interaction effects and difference-in-difference estimation in loglinear models. NBER Technical Working Paper No. 245.

Neyman, J. (1990). On the application of probability theory to agricultural experiments: Essay on statistical principles, Section 9. Translated in *Statistical Science*, **5**, 465-480.

Pocock, S. J., Collier, T. J. et al. (2004). Issues in reporting of epidemiological studies: a survey of recent practice. *British Medical Journal*, 329: 883.

Prentice, R. L. and Kalbfleisch, J. L. (1988). Letter to editor. *Biometrics*, **44**, 1205.

Rothman, K. (1976). Causes. *American Journal of Apidemiology*, **104**, 587-592.

Rothman, K., Greenland, S. and Walker, A. M. (1980). Concepts of interaction. *American Journal of Apidemiology*, **112**, 467-470.

Rothman, K. and Greenland, S. (1998). *Modern Epidemiology*, Lippincott, Philadelphia.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, **66**, 688-701.

Rubin, D. B. (2006). *Matched Sampling for Causal Effects.* Cambridge, England: Cambridge University Press.

Russo F. (2011) *Explaining causal modelling. Or, what a causal model ought to explain*, In M. DAgostino, G. Giorello, F. Laudisa, T. Pievani and C. Sinigaglia (eds), New Essays in Logic and Philosophy of Science, SILF Series, Volume I, College Publications, London.

Sekhon, J. S. (2008). The Neyman-Rubin model of causal inference and estimation via matching methods. In Box-Steffensmeier, Janet, Henry Brady, and David Collier, eds., *The Oxford Handbook of Political Methodology* pages 271-299.

Suhnel, J. (1992). Comment on the paper: A three-dimensional model to analyze drug-drug interactions. *Antiviral Research*, **14**, 181-206.

VanderWeele, T. J. and Robin, J. M. (2008). Empirical and counterfactual conditions for sufficient cause interactions. *Biometrika*, **95**, 49-61.

VanderWeele, T. J. (2009). Sufficient cause interactions and statistical interactions. *Epidemiology*, **20**, 6-13.