



**LIBRARY**  
**Michigan State**  
**University**

This is to certify that the

dissertation entitled

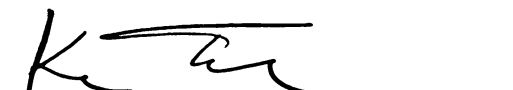
**THE DISTRIBUTION OF THE PRODUCT OF  
TWO DEPENDENT CORRELATION COEFFICIENTS  
WITH APPLICATIONS IN CAUSAL INFERENCE**

presented by

**Wei Pan**

has been accepted towards fulfillment  
of the requirements for

Ph.D. degree in Counseling,  
Educational Psychology, and Special Education

  
Major professor

Date 12-13-01

PLACE IN RETURN BOX to remove this checkout from your record.  
 TO AVOID FINES return on or before date due.  
 MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
DEC 12 2003 12522003		
NOV 12 2006 07 17 06		
JUN 26 2007 04 02 08		

THE DISTRIBUTION OF THE PRODUCT OF  
TWO DEPENDENT CORRELATION COEFFICIENTS  
WITH APPLICATIONS IN CAUSAL INFERENCE

By

Wei Pan

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Educational Psychology,  
and Special Education

2001

## ABSTRACT

### THE DISTRIBUTION OF THE PRODUCT OF TWO DEPENDENT CORRELATION COEFFICIENTS WITH APPLICATIONS IN CAUSAL INFERENCE

By

Wei Pan

Causal inference is an important, controversial topic in the social sciences, in which it is difficult to statistically control for all possible confounding variables. To address this concern, Frank (2000) derives an index, a product of two dependent correlation coefficients (between the confounding variable and the predictor of interest and between the confounding variable and the outcome), to express the sensitivity of regression inferences based on linear modeling to the impact of a confounding variable. Frank's index leads to a promising methodology by which we can inform causal knowledge to address the controversy in causal inference. However, the behavior of the distribution of the product of two dependent correlation coefficients is little known. Frank used a reference distribution generated through an approximation based on the Fisher  $z$  transformation, and then an approximation to the product of two normal variables; therefore, this doubly asymptotic result is tenuous. The present study advances Frank's approach and provides a direct and more accurate approximation to the reference distribution with a closed form—Pearson Type I (*Beta*) distribution. A simulation study is conducted to assess the accuracy of the approximation. With the more accurate

approximation to the reference distribution, we will have more confidence to conclude whether a causal interpretation of a given predictor is robust to confounding variables, that is, whether uncontrolled confounding variables are unlikely to have impacts great enough to alter an inference about a predictor of interest. This study also conveys the robustness into a probability scale, and guidance for interpreting the magnitude of the probability is given. Applications are illustrated with an example pertaining to educational attainment. The methodology discussed in this study would allow for multiple partial causes in the complex social phenomena that we study, informing causal inferences in the social sciences from statistical linear models. The findings in the present study may also be applicable to other methodological issues, such as indirect effects in path analysis.

## ACKNOWLEDGEMENTS

The dissertation would not be possible without the patient and thoughtful guidance of my advisor, Ken Frank. I am deeply grateful for his guidance from the beginning, listening, and supporting with rational care as I stood at the crossroads of countless decisions.

I wish also to recognize the contributions of my committee members. Betsy Becker, Robert Floden, and James Stapleton offered important comments and suggestions that made my dissertation strong both theoretically and practically.

Thanks also to my family. My parents, Yi-Nan Zhou and Zhan-Ping Pan, encouraged me from my childhood to study hard and to be a well-educated person. I am glad that I fulfilled their wish today. My wife, Li Zhang, who sacrificed her profession in our home country coming to the US with me, and my son, Bill, have given me constant strong support.

## TABLE OF CONTENTS

LIST OF TABLES .....	vii
LIST OF FIGURES .....	viii
CHAPTER 1	
INTRODUCTION .....	1
1.1 Causal Inference .....	1
1.2 Confounding and Frank's Index $k$ .....	3
1.3 Purpose of the Study .....	8
CHAPTER 2	
LITERATURE REVIEW .....	10
CHAPTER 3	
APPROXIMATION PROCEDURES .....	13
3.1 Moments of $r_{xc}r_{yc}$ .....	13
3.2 Pearson Distributions .....	23
3.3 Approximate Distribution of $r_{xc}r_{yc}$ .....	25
CHAPTER 4	
SIMULATION STUDY.....	28
4.1 Simulation Design .....	28
4.2 Simulation Results .....	38
CHAPTER 5	
APPROXIMATION CORRECTION .....	48
CHAPTER 6	
DISTRIBUTION COMPARISON .....	56
CHAPTER 7	
APPLICATIONS .....	62
7.1 An Index of the Robustness of a Causal Inference (IOROCI) .....	62
7.2 An Example .....	71

<b>CHAPTER 8</b>	
<b>DISCUSSION</b> .....	75
8.1 Conclusions and Limitations .....	75
8.2 Extensions .....	78
<b>APPENDIX A</b> .....	80
<b>REFERENCES</b> .....	83

## LIST OF TABLES

Table 1 – Coefficients of the Regressions .....	4
Table 2 – Parameter Specifications for the Simulation Study .....	29
Table 3 – Simulation Results for the Mean $\mu'_1$ ( $N = 28$ ) .....	32
Table 4 – Simulation Results for the Mean $\mu'_1$ ( $N = 84$ ) .....	32
Table 5 – Simulation Results for the Mean $\mu'_1$ ( $N = 783$ ) .....	33
Table 6 – Simulation Results for the Variance $\mu_2$ ( $N = 28$ ) .....	33
Table 7 – Simulation Results for the Variance $\mu_2$ ( $N = 84$ ) .....	34
Table 8 – Simulation Results for the Variance $\mu_2$ ( $N = 783$ ) .....	34
Table 9 – Simulation Results for the Third Moment $\mu_3$ ( $N = 28$ ) .....	35
Table 10 – Simulation Results for the Third Moment $\mu_3$ ( $N = 84$ ) .....	35
Table 11 – Simulation Results for the Third Moment $\mu_3$ ( $N = 783$ ) .....	36
Table 12 – Simulation Results for the Fourth Moment $\mu_4$ ( $N = 28$ ) .....	36
Table 13 – Simulation Results for the Fourth Moment $\mu_4$ ( $N = 84$ ) .....	37
Table 14 – Simulation Results for the Fourth Moment $\mu_4$ ( $N = 783$ ) .....	37
Table 15 – Estimated Values of the Third Moment Compared with the Simulated Values and the Approximated Values ( $N = 28$ ) .....	52
Table 16 – Estimated Values of the Third Moment Compared with the Simulated Values and the Approximated Values ( $N = 84$ ) .....	53
Table 17 – Estimated Values of the Fourth Moment Compared with the Simulated Values and the Approximated Values ( $N = 28$ ) .....	54
Table 18 – Estimated Values of the Fourth Moment Compared with the Simulated Values and the Approximated Values ( $N = 84$ ) .....	55
Table 19 – Detailed Counts and Percentages for $t_c < -2$ , $-2 < t_c < 2$ , and $2 < t_c$ Based on A Simulated Dataset with $N = 84$ and $\rho_{xy} = .30$ .....	67

## LIST OF FIGURES

Figure 1 – Specifications of $\kappa$ for Distinguishing the Types of Pearson Distributions .....	24
Figure 2 – The Distributions of the Product of Simulated Correlations $\tilde{r}_{xc}\tilde{r}_{yc}$ with Different Numbers of Replications for the Selected Cells .....	39
Figure 3 – The Trends of the Accuracy of the Approximated and Frank's Moments along One Parameter Holding the Other Parameters Constant, Including the Sample Size .....	42
Figure 4 – P-P Plots for the Selected Cells .....	58
Figure 5 – P-P Plots for the Selected Cells Using the Estimated Values for the Third and the Fourth Moments .....	59
Figure 6 – A P-P Plot with Frank's Distribution for $\rho_{xc} = \rho_{yc} = \rho_{xy} = .5$ and $N = 783$ ...	61
Figure 7 – The relationship of $t_c$ and $t_0$ .....	65
Figure 8 – The relationship of $t_c$ and $t_0$ for a simulated dataset with $N = 84$ and $\rho_{xy} = .30$ .....	66
Figure 9 – <i>Father's Education</i> as a Potential Confounding Variable for the Causal Relationship Between <i>Father's Occupation</i> and <i>Educational Attainment</i> ...	71
Figure 10 – Distribution function of $k = r_{xc}r_{yc}$ .....	73

## **Chapter 1**

### **INTRODUCTION**

#### **1.1 Causal Inference**

Causal inference is an important, controversial issue in most fields of the social sciences, such as educational research, behavioral research, psychometrics, econometrics, and sociology as well as epidemiology and biostatistics. In those fields, researchers routinely draw conclusions about causal relationships between dependent variables and independent variables from statistical linear models using data from observational studies (See Jacobs, Finken, Griffin, & Wright, 1998; Lee, 1999; Okagaki & Frensch, 1998; Portes, 1999, for examples). However, the usual statistical approaches may not lead to valid causal inferences, even if the models are supported by related theories and fully specified (Abbott, 1998; Cook & Campbell, 1979; Holland, 1986, 1988; McKim & Turner, 1997; Pearl, 2000; Rubin, 1974; Sobel, 1996, 1998). The problem mainly comes from the failure to control for all possible confounding variables for which the list cannot be exhausted.

For instance, Okagaki and Frensch (1998) examined the relationship between parenting and children's school performance for different ethnic groups, but did not control for the children's age, gender, or socioeconomic status. Jacobs, Finken, Griffin, and Wright (1998) examined the relationships between parent attitudes, intrinsic values

of science, peer support, available activities, and preference for future science career for science-talented, rural, adolescent females. However, they also failed to control for other demographics, such as age and socioeconomic status. A third example is that Lee (1999) examined the differences in children's views of the world after they personally experienced a natural disaster for various ethnic, socioeconomic status, and gender groups, but failed to control for pre-world-views. Still another, Portes (1999) examined the influence of various factors in immigrant students' school achievement and controlled for many demographic and sociopsychological covariates. Nevertheless, we still can ask: "Did he control for all possible sociopsychological factors?" Therefore, the conclusions in each case might not support causal relationships, although there are statistically significant results.

To accommodate this problem, the literature suggests the following options:

- (1) Use alternative models, e.g., randomized and well-controlled non-randomized studies (Rubin, 1974);
- (2) Try causal discovery algorithms which operate on statistical data sets to produce directed causal graphs (Spirtes, Glymour, & Scheines, 1993);
- (3) Abandon the use of causal language and emphasize the effects of causes rather than the causes of effects (Holland, 1986, 1988; Sobel, 1996);
- (4) Spend more effort on descriptive work (Abbott, 1998; Sobel, 1998).

If we are still interested in exploring causal relationships in the real world, options (3) and (4) will not work. As for option (1), random assignment is often impractical in the social sciences given logical, ethical, and political concerns. In addition, it is not

always possible to measure all confounding variables to be controlled for in statistical analyses. Thus, option (1) would be also inapplicable. In option (2), the causal graphs are generated by calculations of conditional statistical dependence or independence among pairs of variables, but in most cases, the assumptions under which the algorithms operate are not powerful enough to uniquely identify the real causal structure underlining correlational data rather than some set of statistically equivalent but genuinely alternative representations (Woodward, 1997). Thus, the soundness of the methodology of causal graphs is uncertain.

Is there a comparatively simple, feasible way to explore causal relationships using commonly used statistical linear models? The answer is optimistic. Although we are never able to find all possible causes of an outcome, we can statistically characterize the extent to which a causal inference regarding a given predictor is robust to the impacts of other minor causes of the outcome using Frank's (2000) index of the impact of a confounding variable. Before going further to discuss how to implement Frank's index, I will first explain below what Frank's index is as well as the concept of confounding.

## **1.2 Confounding and Frank's Index $k$**

A confounding variable is one related to both the predictor and the outcome and it is also assumed to occur causally prior to the predictor (Anderson, Auquier, Hauck, Oakes, Vandaele, & Weisberg, 1980; Cook & Campbell, 1979). If a confounding variable were introduced into a linear model, the effect of a predictor of interest on an outcome might be changed from statistically significant to not statistically significant.

For example, I took 990 schools that have complete data regarding the following three relevant variables from the NELS:88 data (National Center for Education Statistics, 1996) and found, using a simple regression, that *Teachers' Morale* (BYSC47G) of those schools has a significant effect ( $p < .001$ ) on their *Students' Academic Achievement* (mean score of F12XCOMP—standardized test composite of reading and math). Also, extensive research has shown that *School Socioeconomic Status* (mean score of BYSES) is related to both *Teachers' Morale* and *Students' Academic Achievement* (Chall, Jacobs, & Baldwin, 1991; Miller & et al., 1986; Solomon, Battistich, & Hom, 1996; Trusty, Peck, & Mathews, 1994, to name a few). Thus, *School Socioeconomic Status* is a potential confounding variable for *Teachers' Morale* on *Students' Academic Achievement*. After introducing *School Socioeconomic Status* into the regression model, I found that the effect of *Teachers' Morale* on their *Students' Academic Achievement* was no longer statistically significant ( $p > .448$ ) (see Table 1), supporting that *School Socioeconomic Status* is a confounding variable for *Teachers' Morale* on their *Students' Academic Achievement*.

Table 1  
*Coefficients of the Regressions<sup>a</sup>*

Model	Variable	Unstd. Coefficient	(s.e.)	Std. Coefficient	<i>t</i>	<i>p</i>
1	Intercept	47.567	(.842)		56.493	.000
	Teachers' Morale	.849	(.206)	.130	4.121	.000
2	Intercept	51.226	(.541)		94.607	.000
	Teachers' Morale	.100	(.132)	.015	.759	.448
	School SES	8.799	(.229)	.776	38.448	.000

*Note.*  $N = 990$ . <sup>a</sup>Dependent variable: F12XCOMP—standardized test composite (reading & math—school mean).

Usually, we do not always have measures of confounds, and cannot always control for them. However, we can ask: “How large must be the impact of a confounding variable to alter the inference?” Technically, the impact can be obtained by expressing a  $t$ -statistic for a regression coefficient in terms of zero-order correlations:

$$t = \frac{\hat{\beta}_x}{se(\hat{\beta}_x)} = \frac{r_{xy} - r_{xc}r_{yc}}{\sqrt{\frac{1 - r_{xy}^2 - r_{xc}^2 - r_{yc}^2 + 2r_{xy}r_{xc}r_{yc}}{n - q - 1}}}, \quad (1)$$

where

$X$  is the predictor of interest;

$Y$  is the outcome;

$C$  is the confounding variable;

$n$  is the sample size;

$q$  is the number of independent variables, e.g., 2 if are here  $X$  and  $C$ ;

$\hat{\beta}_x$  is an estimate of the regression coefficient of  $X$ ;

$se(\hat{\beta}_x)$  is the standard error of  $\hat{\beta}_x$ ;

$r_{xy}$ ,  $r_{xc}$ , and  $r_{yc}$  are the observed correlation coefficients between  $X$  and  $Y$ , between  $X$  and  $C$ , and between  $Y$  and  $C$ , respectively.

This formula holds if the regression model only has  $X$  and  $C$  as independent variables. For the case of more than two independent variables, the formula will be more complex (cf., Frank, 2000, p. 164). If the coefficient  $\beta_x$  changes from statistically significant to non-significant after including the confounding variable, the  $t$ -statistic (1) crosses the threshold of the critical value.

Unfortunately, in many cases, we do not have a measure of a confounding variable. In these cases, there is no absolute rebuttal to challenges of causal inference associated with confounding variables. Nonetheless, Frank (2000) quantifies the impact of a confounding variable on a regression coefficient as a product of the two correlation coefficients  $r_{xc}$  and  $r_{yc}$ :  $k = r_{xc} \times r_{yc}$ <sup>[1]</sup>. Given this constraint, the value of the  $t$ -statistic (1) achieves its minimum value when  $r_{xc}^2 = r_{yc}^2 = k$ . Thus, the  $t$ -value can be re-expressed as

$$t_{\min} = \frac{r_{xy} - k}{\sqrt{\frac{(1 + r_{xy} - 2k)(1 - r_{xy})}{n - q - 1}}}, \quad (2)$$

which affords the confounding variable the greatest impact on the inference regarding the regression coefficient of  $X$ .

The expression (2) reveals the relationship between the index  $k$  and the  $t$ -value. That is, we can express  $k$  as a function of an estimated correlation  $r_{xy}$  and the minimum  $t$ -value,  $t_{\min}$  (2). Therefore, if we set the minimum  $t$ -value equal to a critical value, we can obtain a magnitude of  $k$  that is necessary for a potential confounding variable to alter the inference regarding the predictor on the outcome, even if we are not able to measure the confounding variable. In other words, if we had a confounding variable, we know from the index  $k$  how strong the correlational relationships of the confounding variable with the predictor and with the outcome must be to alter the inference for the predictor.

Frank refers to the threshold at which the impact of a confounding variable would alter a statistical inference as the impact threshold for a confounding variable (ITCV).

---

<sup>[1]</sup> (a) In the case of more than two independent variables,  $k$  will be the product of two partial correlations (cf. Frank, 2000). (b) For the moment, assume  $k > 0$ . For the case of  $k < 0$ , see Frank (2000).

That is, if the index  $k$  of any potential confounding variable does not exceed the corresponding ITCV of a given predictor that is statistically significant, we can say that the causal inference about the predictor on the outcome is robust to the other causes or confounding variables. Therefore, we may not need to worry about the validity of the predictor as a major cause of the outcome, since we can argue for the validity in terms of the ITCV. Thus, Frank's methodology is a promising attempt to lessen the crisis in causal inference that was mentioned at the beginning of the chapter, by rephrasing the problem in terms of the sensitivity of a statistical inference to the impact of confounding variables.

Of course, confounding variables are usually unmeasured or immeasurable. While Frank's index quantifies the impact necessary to alter an inference, how do we know the likelihood that such an impact could or would alter the inference in the presence of the confounding variable? One response is to generate a reference distribution for the impact of the unmeasured confounding variable from the impacts of existing, measured covariates. The reference distribution takes the same form as that of index  $k^{[2]}$ : the product of two dependent correlation coefficients between the covariate and the predictor and between the covariate and the outcome. With this reference distribution, we can assess the likelihood that the causal interpretation of the predictor could be altered if a confounding variable with comparable impact were measured and controlled in the linear model.

---

<sup>[2]</sup> In the case of more than two independent variables,  $k$  is the product of two partial correlations (cf., Frank, 2000, p. 166).

Some researchers may be uncomfortable with the strategy of using measured covariates to generate a reference distribution for the impact of an unknown confounding variable. For example, a poorly chosen set of covariates will underestimate the impact of an important confounding variable. Thus, we acknowledge that this use of the reference distribution is only as valid as is the set of covariates on which it is based. In this sense, the problem is no different from any other associated with making an inference from a sample that must be representative of the population. In this light, the impact of existing covariates represents important information by which to assess the ITCV. For example, would it not be informative if the ITCV were much larger than the impact of any measured covariate? If one agrees, then the question is not whether to use the impacts of measured covariates, but how to use this information. This becomes the core of the present study.

### 1.3 Purpose of the Study

In order to utilize the reference distribution, we must understand the behavior of the distribution of the product of two dependent correlation coefficients. From Cohen and Cohen (1983, p. 280), we know that the product of two dependent correlation coefficients is constrained by the upper and lower limits, rather than just  $-1$  and  $1$ :

$$r_{xy} - \sqrt{(1 - r_{xc}^2)(1 - r_{yc}^2)} < r_{xc}r_{yc} < r_{xy} + \sqrt{(1 - r_{xc}^2)(1 - r_{yc}^2)}. \quad (3)$$

Thus, the distribution of the product is obviously not exactly normal. However, beyond this constraint, we know little.

Frank (2000) used an approximation based on Fisher  $z$  transformation, and then an approximation to the product of two normal variables, to assess the likelihood of an impact of a confounding variable greater than ITCV. Unfortunately, this doubly asymptotic result is tenuous. Therefore, the remaining task is to generate a more accurate and direct expression for the reference distribution. The purpose of this study is to obtain a more accurate approximation to the reference distribution that is the product of two dependent correlation coefficients. Using the reference distribution generated by the more accurate approximation method will give us a more valid conclusion about the robustness of a cause inference to the impact of a confounding variable.

## Chapter 2

### LITERATURE REVIEW

There are many papers in the literature about the distribution of a single correlation coefficient (Konishi, 1978; Konishi, 1979; Kraemer, 1973; Olkin, 1967; Olkin & Siotani, 1976) or the distribution of the difference of two correlation coefficients (Choi, 1977; Dunn & Clark, 1971; Meng, Rosenthal, & Rubin, 1992; Neill & Dunn, 1975; Olkin, 1967; Steiger, 1980; Wolfe, 1976), but few about the distribution of the product of two correlation coefficients. I will review the few papers below.

First, we know from asymptotic distribution theory that the distribution of a set of correlation coefficients approaches the multivariate normal distribution, as the sample size becomes very large (Steiger, 1980, p.246). Hence, the focal problem of this thesis could be solved by a straightforward approach that treats the product of two correlation coefficients as a product of two asymptotically normal variables and incorporates the findings in the literature about the distribution of the product of two normal variables (Aroian, 1947; Aroian, Taneja, & Cornwell, 1978; Cornwell, Aroian, & Taneja, 1978; Craig, 1936, 1942; Meeker, Cornwell, & Aroian, 1981; Meeker & Escobar, 1994; Springer, 1983; Wallgren, 1980). However, the asymptotic convergence of this product is very slow (Craig, 1936). Thus, this approach presents a severe limitation in application.

There are two other plausible approaches to obtain the distribution of the product of two correlation coefficients. Mathai and Saxena (1969) express the product of two correlation coefficients as a special case of the product of two generalized Mellin-Barnes functions or  $H$ -functions (Mathai & Saxena, 1978). However, the expression obtained for the distribution function is quite unwieldy. In addition, in their study, the two correlation coefficients are assumed independent, while the two relevant correlation coefficients defining the impact index  $k$  are associated with a common variable—the confounding variable or covariate. Therefore, the two correlation coefficients that we are interested in are not statistically independent (Boyer, Palachek, & Schucany, 1983; Choi, 1977; Dunn & Clark, 1969, 1971; May & Hittner, 1997a, b; Meng, Rosenthal, & Rubin, 1992; Neill & Dunn, 1975; Olkin & Finn, 1990; Steiger, 1980; Williams, 1959; Wolfe, 1976).

The other approach is described in Frank (2000) and mentioned above. Frank transforms the two correlation coefficients to two asymptotically, normally distributed Fisher  $z$ s, then uses Aroian's findings about the distribution of the product of two normal variables to obtain the distribution of the product of two Fisher  $z$ 's, instead of the two original correlation coefficients. A problem with this approach, however, is that we do not have a closed form function for the distribution. Furthermore, we do not know how stable the Fisher  $z$  transformation is. Moreover, we are interested in the product of two original correlation coefficients, while the behavior of the distribution of the product of two Fisher  $z$ 's is detached from that of two original correlation coefficients. That is, Frank's approach is relying on asymptotic theory for Fisher's  $z$ , compounded with the

approximation error associated with Aroian's approach, which results in very slow convergence.

## Chapter 3

### APPROXIMATION PROCEDURES

As stated in Chapter 2, the prior research does not provide sound methods for obtaining the distribution of the product of two dependent correlation coefficients. To address this limitation, in this Chapter, I will develop a more accurate approximation to, and a direct expression for, the distribution of the product of two dependent correlation coefficients. The approximation procedures have two steps. First, I obtain the first four moments of the product of  $r_{xc}r_{yc}$ . Then, I apply these moments to Pearson distribution family (Pearson, 1895), obtaining an approximate distribution of  $r_{xc}r_{yc}$  as a Pearson Type I (*Beta*) distribution.

#### 3.1 Moments of $r_{xc}r_{yc}$

Let  $\rho_{xy}$ ,  $\rho_{xc}$ , and  $\rho_{yc}$  be the population values of  $r_{xy}$ ,  $r_{xc}$ , and  $r_{yc}$ , respectively. Let  $\Delta r_{xy}$ ,  $\Delta r_{xc}$ , and  $\Delta r_{yc}$  be the deviations of the sample values from their population values. In particular,  $\Delta r_{xy} = r_{xy} - \rho_{xy}$ ,  $\Delta r_{xc} = r_{xc} - \rho_{xc}$ , and  $\Delta r_{yc} = r_{yc} - \rho_{yc}$ . Then, we have

$$r_{xc}r_{yc} = \rho_{xc}\rho_{yc} + \rho_{xc}\Delta r_{yc} + \rho_{yc}\Delta r_{xc} + \Delta r_{xc}\Delta r_{yc};$$

$$(r_{xc}r_{yc})^2 = \rho_{xc}^2\rho_{yc}^2 + \rho_{xc}^2(\Delta r_{yc})^2 + \rho_{yc}^2(\Delta r_{xc})^2 + (\Delta r_{xc})^2(\Delta r_{yc})^2 + 2\rho_{xc}^2\rho_{yc}\Delta r_{yc} \\ + 2\rho_{xc}\rho_{yc}^2\Delta r_{xc} + 2\rho_{yc}(\Delta r_{xc})^2\Delta r_{yc} + 2\rho_{xc}\Delta r_{xc}(\Delta r_{yc})^2 + 4\rho_{xc}\rho_{yc}\Delta r_{xc}\Delta r_{yc};$$

$$(r_{xc}r_{yc})^3 = \rho_{xc}^3\rho_{yc}^3 + \rho_{xc}^3(\Delta r_{yc})^3 + \rho_{yc}^3(\Delta r_{xc})^3 + (\Delta r_{xc})^3(\Delta r_{yc})^3 + 3\rho_{xc}^3\rho_{yc}^2\Delta r_{yc} \\ + 3\rho_{xc}^3\rho_{yc}(\Delta r_{yc})^2 + 3\rho_{xc}^2\rho_{yc}^3\Delta r_{xc} + 3\rho_{xc}\rho_{yc}^3(\Delta r_{xc})^2 + 3\rho_{xc}^2\Delta r_{xc}(\Delta r_{yc})^3 \\ + 3\rho_{xc}(\Delta r_{xc})^2(\Delta r_{yc})^3 + 3\rho_{yc}^3(\Delta r_{xc})^3\Delta r_{yc} + 3\rho_{yc}(\Delta r_{xc})^3(\Delta r_{yc})^2 \\ + 9\rho_{xc}^2\rho_{yc}^2\Delta r_{xc}\Delta r_{yc} + 9\rho_{xc}^2\rho_{yc}\Delta r_{xc}(\Delta r_{yc})^2 + 9\rho_{xc}\rho_{yc}^2(\Delta r_{xc})^2\Delta r_{yc} \\ + 9\rho_{xc}\rho_{yc}(\Delta r_{xc})^2(\Delta r_{yc})^2;$$

$$(r_{xc}r_{yc})^4 = \rho_{xc}^4\rho_{yc}^4 + \rho_{xc}^4(\Delta r_{yc})^4 + \rho_{yc}^4(\Delta r_{xc})^4 + (\Delta r_{xc})^4(\Delta r_{yc})^4 + 4\rho_{xc}^4\rho_{yc}^3\Delta r_{yc} \\ + 4\rho_{xc}^3\rho_{yc}^4\Delta r_{xc} + 4\rho_{yc}^3(\Delta r_{xc})^4\Delta r_{yc} + 4\rho_{xc}(\Delta r_{xc})^3(\Delta r_{yc})^4 + 4\rho_{xc}^4\rho_{yc}(\Delta r_{yc})^3 \\ + 4\rho_{xc}^3\Delta r_{xc}(\Delta r_{yc})^4 + 4\rho_{xc}\rho_{yc}^4(\Delta r_{xc})^3 + 4\rho_{yc}(\Delta r_{xc})^4(\Delta r_{yc})^3 + 6\rho_{xc}^4\rho_{yc}^2(\Delta r_{yc})^2 \\ + 6\rho_{xc}^2\rho_{yc}^4(\Delta r_{xc})^2 + 6\rho_{xc}^2(\Delta r_{xc})^2(\Delta r_{yc})^4 + 6\rho_{yc}^2(\Delta r_{xc})^4(\Delta r_{yc})^2 \\ + 16\rho_{xc}^3\rho_{yc}^3\Delta r_{xc}\Delta r_{yc} + 16\rho_{xc}^3\rho_{yc}\Delta r_{xc}(\Delta r_{yc})^3 + 16\rho_{xc}\rho_{yc}^3(\Delta r_{xc})^3\Delta r_{yc} \\ + 16\rho_{xc}\rho_{yc}(\Delta r_{xc})^3(\Delta r_{yc})^3 + 24\rho_{xc}^3\rho_{yc}^2\Delta r_{xc}(\Delta r_{yc})^2 + 24\rho_{xc}^2\rho_{yc}^3(\Delta r_{xc})^2\Delta r_{yc} \\ + 24\rho_{xc}^2\rho_{yc}(\Delta r_{xc})^2(\Delta r_{yc})^3 + 24\rho_{xc}\rho_{yc}^2(\Delta r_{xc})^3(\Delta r_{yc})^2 \\ + 36\rho_{xc}^2\rho_{yc}^2(\Delta r_{xc})^2(\Delta r_{yc})^2.$$

Dropping the terms of order higher than the fourth<sup>[3]</sup> and taking expectations of

$(r_{xc}r_{yc})^i$ ,  $i = 1, 2, 3, 4$ , give us the approximate first four non-central moments as follows:

---

<sup>[3]</sup> Since we only want to obtain the approximate first four moments, the terms of order higher than the fourth will have little effect on the approximation.

$$\begin{aligned}
\mu'_1 &= E(r_{xc}r_{yc}) = \rho_{xc}\rho_{yc} + \rho_{xc}E[(\Delta r_{yc})] + \rho_{yc}E[(\Delta r_{xc})] + E[(\Delta r_{xc}\Delta r_{yc})]; \\
\mu'_2 &= E[(r_{xc}r_{yc})^2] = \rho_{xc}^2\rho_{yc}^2 + \rho_{xc}^2E[(\Delta r_{yc})^2] + \rho_{yc}^2E[(\Delta r_{xc})^2] + E[(\Delta r_{xc})^2(\Delta r_{yc})^2] \\
&\quad + 2\rho_{xc}^2\rho_{yc}E[(\Delta r_{yc})] + 2\rho_{xc}\rho_{yc}^2E[(\Delta r_{xc})] + 2\rho_{yc}E[(\Delta r_{xc})^2\Delta r_{yc}] \\
&\quad + 2\rho_{xc}E[(\Delta r_{xc})(\Delta r_{yc})^2] + 4\rho_{xc}\rho_{yc}E[(\Delta r_{xc}\Delta r_{yc})]; \\
\mu'_3 &= E[(r_{xc}r_{yc})^3] = \rho_{xc}^3\rho_{yc}^3 + \rho_{xc}^3E[(\Delta r_{yc})^3] + \rho_{yc}^3E[(\Delta r_{xc})^3] + 3\rho_{xc}^3\rho_{yc}^2E[(\Delta r_{yc})] \\
&\quad + 3\rho_{xc}^3\rho_{yc}E[(\Delta r_{yc})^2] + 3\rho_{xc}^2\rho_{yc}^3E[(\Delta r_{xc})] + 3\rho_{xc}\rho_{yc}^3E[(\Delta r_{xc})^2] \\
&\quad + 3\rho_{xc}^2E[(\Delta r_{xc})(\Delta r_{yc})^3] + 3\rho_{yc}^2E[(\Delta r_{xc})^3\Delta r_{yc}] + 9\rho_{xc}^2\rho_{yc}^2E[(\Delta r_{xc}\Delta r_{yc})] \\
&\quad + 9\rho_{xc}^2\rho_{yc}E[(\Delta r_{xc})(\Delta r_{yc})^2] + 9\rho_{xc}\rho_{yc}^2E[(\Delta r_{xc})^2\Delta r_{yc}] \\
&\quad + 9\rho_{xc}\rho_{yc}E[(\Delta r_{xc})^2(\Delta r_{yc})^2]; \\
\mu'_4 &= E[(r_{xc}r_{yc})^4] = \rho_{xc}^4\rho_{yc}^4 + \rho_{xc}^4E[(\Delta r_{yc})^4] + \rho_{yc}^4E[(\Delta r_{xc})^4] + 4\rho_{xc}^4\rho_{yc}^3E[(\Delta r_{yc})] \\
&\quad + 4\rho_{xc}^3\rho_{yc}^4E[(\Delta r_{xc})] + 4\rho_{xc}^4\rho_{yc}E[(\Delta r_{yc})^3] + 4\rho_{xc}\rho_{yc}^4E[(\Delta r_{xc})^3] \\
&\quad + 6\rho_{xc}^4\rho_{yc}^2E[(\Delta r_{yc})^2] + 6\rho_{xc}^2\rho_{yc}^4E[(\Delta r_{xc})^2] + 16\rho_{xc}^3\rho_{yc}^3E[(\Delta r_{xc}\Delta r_{yc})] \\
&\quad + 16\rho_{xc}^3\rho_{yc}E[(\Delta r_{xc})(\Delta r_{yc})^3] + 16\rho_{xc}\rho_{yc}^3E[(\Delta r_{xc})^3\Delta r_{yc}] \\
&\quad + 24\rho_{xc}^3\rho_{yc}^2E[(\Delta r_{xc})(\Delta r_{yc})^2] + 24\rho_{xc}^2\rho_{yc}^3E[(\Delta r_{xc})^2\Delta r_{yc}] \\
&\quad + 36\rho_{xc}^2\rho_{yc}^2E[(\Delta r_{xc})^2(\Delta r_{yc})^2].
\end{aligned} \tag{4}$$

In order to obtain closed form expressions for the first four moments, we need to express  $E[(\Delta r_{xc})^i]$ ,  $i = 1, 2, 3, 4$ ,  $E[(\Delta r_{yc})^j]$ ,  $j = 1, 2, 3, 4$ , and  $E[(\Delta r_{xc})^k(\Delta r_{yc})^l]$ ,  $k, l = 1, 2$ , or  $3$ , in terms of  $\rho_{xc}$ ,  $\rho_{yc}$ , or  $\rho_{xy}$ . Before going further, we need to define some notation and get some preliminary results. Suppose the three initial variables  $X$ ,  $Y$ , and  $C$  follow a trivariate normal distribution. Then, following Ghosh (1966) and Hotelling (1936, 1940), the moments and the covariance of  $r_{xc}$  and  $r_{yc}$  can be approximately expressed as follows:

$$\begin{aligned}
\mu_{r_*} &= E(r_*) = \rho_* - \frac{\rho_*(1-\rho_*^2)}{2M} \left\{ 1 + \frac{9}{4M} (3 + \rho_*^2) + \frac{3}{8M^2} (121 + 70\rho_*^2 + 25\rho_*^4) \right. \\
&\quad + \frac{3}{64M^3} (6479 + 4923\rho_*^2 + 2925\rho_*^4 + 1225\rho_*^6) \\
&\quad \left. + \frac{3}{128M^4} (86341 + 77260\rho_*^2 + 58270\rho_*^4 + 38220\rho_*^6 + 19845\rho_*^8) \right\}; \\
\sigma_{r_*}^{(2)} &= Var(r_*) = E[(r_* - \mu_{r_*})^2] \\
&= \frac{(1-\rho_*^2)^2}{M} \left\{ 1 + \frac{1}{2M} (14 + 11\rho_*^2) + \frac{1}{2M^2} (98 + 130\rho_*^2 + 75\rho_*^4) \right. \\
&\quad + \frac{1}{8M^3} (2744 + 4645\rho_*^2 + 4422\rho_*^4 + 2565\rho_*^6) \\
&\quad \left. + \frac{1}{8M^4} (19208 + 37165\rho_*^2 + 44499\rho_*^4 + 40299\rho_*^6 + 26685\rho_*^8) \right\}; \\
\sigma_{r_*}^{(3)} &= E[(r_* - \mu_{r_*})^3] = -\frac{\rho_*(1-\rho_*^2)^3}{M^2} \left\{ 6 + \frac{1}{M} (69 + 88\rho_*^2) \right. \\
&\quad + \frac{3}{4M^2} (797 + 1691\rho_*^2 + 1560\rho_*^4) \\
&\quad \left. + \frac{3}{8M^3} (12325 + 33147\rho_*^2 + 488099\rho_*^4 + 44109\rho_*^6) \right\}; \\
\sigma_{r_*}^{(4)} &= E[(r_* - \mu_{r_*})^4] = \frac{3(1-\rho_*^2)^4}{M^2} \left\{ 1 + \frac{1}{M} (12 + 35\rho_*^2) \right. \\
&\quad + \frac{1}{4M^2} (436 + 2028\rho_*^2 + 3025\rho_*^4) \\
&\quad \left. + \frac{1}{4M^3} (3552 + 20009\rho_*^2 + 46462\rho_*^4 + 59751\rho_*^6) \right\}; \\
\sigma_{r_{xc}, r_{yc}} &= Cov(r_{xc}, r_{yc}) = E[(r_{xc} - \mu_{r_{xc}})(r_{yc} - \mu_{r_{yc}})] \\
&= \frac{1}{M} \left[ \rho_{xy} (1 - \rho_{xc}^2 - \rho_{yc}^2) - \frac{1}{2} \rho_{xc} \rho_{yc} (1 - \rho_{xc}^2 - \rho_{yc}^2 - \rho_{xy}^2) \right],
\end{aligned} \tag{5}$$

where  $M = N + 6$ ,  $N$  is the sample size; the subscript “•” represents  $x_c$  or  $y_c$ ; and the superscripts “(2)”, “(3)”, and “(4)” represent a variance, a third moment, and a fourth moment, respectively (as distinguished from quadratic, cubic, and quartic powers).

On the other hand, we have the expressions for  $E[(\Delta r_\bullet)^i]$ ,  $i = 1, 2, 3, 4$ , and  $E(\Delta r_{xc} \Delta r_{yc})$  as follows:

$$\begin{aligned}
 E(\Delta r_\bullet) &= E(r_\bullet - \rho_\bullet) = \mu_{r_\bullet} - \rho_\bullet, \text{ which will be referred to as } b_{r_\bullet}; \\
 E[(\Delta r_\bullet)^2] &= E[(r_\bullet - \rho_\bullet)^2] = E\{[(r_\bullet - \mu_{r_\bullet}) - (\mu_{r_\bullet} - \rho_\bullet)]^2\} \\
 &= E[(r_\bullet - \mu_{r_\bullet})^2] + (\mu_{r_\bullet} - \rho_\bullet)^2 = \sigma_{r_\bullet}^{(2)} + b_{r_\bullet}^2; \\
 E[(\Delta r_\bullet)^3] &= E[(r_\bullet - \rho_\bullet)^3] = E\{[(r_\bullet - \mu_{r_\bullet}) - (\mu_{r_\bullet} - \rho_\bullet)]^3\} \\
 &= E[(r_\bullet - \mu_{r_\bullet})^3] - 3(\mu_{r_\bullet} - \rho_\bullet)E[(r_\bullet - \mu_{r_\bullet})^2] - (\mu_{r_\bullet} - \rho_\bullet)^3 \\
 &= \sigma_{r_\bullet}^{(3)} - 3\sigma_{r_\bullet}^{(2)}b_{r_\bullet} - b_{r_\bullet}^3; \\
 E[(\Delta r_\bullet)^4] &= E[(r_\bullet - \rho_\bullet)^4] = E\{[(r_\bullet - \mu_{r_\bullet}) - (\mu_{r_\bullet} - \rho_\bullet)]^4\} \\
 &= E[(r_\bullet - \mu_{r_\bullet})^4] - 4(\mu_{r_\bullet} - \rho_\bullet)E[(r_\bullet - \mu_{r_\bullet})^3] \\
 &\quad + 6(\mu_{r_\bullet} - \rho_\bullet)^2 E[(r_\bullet - \mu_{r_\bullet})^2] + (\mu_{r_\bullet} - \rho_\bullet)^4 \\
 &= \sigma_{r_\bullet}^{(4)} - 4\sigma_{r_\bullet}^{(3)}b_{r_\bullet} + 6\sigma_{r_\bullet}^{(2)}b_{r_\bullet}^2 + b_{r_\bullet}^4; \\
 E(\Delta r_{xc} \Delta r_{yc}) &= E[(r_{xc} - \rho_{xc})(r_{yc} - \rho_{yc})] \\
 &= E\{[(r_{xc} - \mu_{r_{xc}}) - (\mu_{r_{xc}} - \rho_{xc})][(r_{yc} - \mu_{r_{yc}}) - (\mu_{r_{yc}} - \rho_{yc})]\} \\
 &= E[(r_{xc} - \mu_{r_{xc}})(r_{yc} - \mu_{r_{yc}})] + (\mu_{r_{xc}} - \rho_{xc})(\mu_{r_{yc}} - \rho_{yc}) \\
 &= \sigma_{r_{xc}, r_{yc}} + b_{r_{xc}} b_{r_{yc}}.
 \end{aligned} \tag{6}$$

We also need the expressions for the higher order product-moments,  $E[(\Delta r_{xc})^s(\Delta r_{yc})^t]$ ,  $s, t$

$= 1, 2$ , or  $3$ . After a few simplifications, we first have:

$$\begin{aligned}
E(\Delta r_i \Delta r_j \Delta r_k \Delta r_l) &= E[(r_i - \rho_i)(r_j - \rho_j)(r_k - \rho_k)(r_l - \rho_l)] \\
&= E\{[(r_i - \mu_{r_i}) + (\mu_{r_i} - \rho_i)][(r_j - \mu_{r_j}) + (\mu_{r_j} - \rho_j)] \\
&\quad \cdot [(r_k - \mu_{r_k}) + (\mu_{r_k} - \rho_k)][(r_l - \mu_{r_l}) + (\mu_{r_l} - \rho_l)]\} \\
&= E[(r_i - \mu_{r_i})(r_j - \mu_{r_j})(r_k - \mu_{r_k})(r_l - \mu_{r_l})] \\
&\quad + b_{r_i} E[(r_i - \mu_{r_i})(r_j - \mu_{r_j})(r_k - \mu_{r_k})] \\
&\quad + b_{r_k} E[(r_i - \mu_{r_i})(r_j - \mu_{r_j})(r_l - \mu_{r_l})] \\
&\quad + b_{r_j} E[(r_i - \mu_{r_i})(r_k - \mu_{r_k})(r_l - \mu_{r_l})] \\
&\quad + b_{r_l} E[(r_j - \mu_{r_j})(r_k - \mu_{r_k})(r_l - \mu_{r_l})] \\
&\quad + b_{r_k} b_{r_i} E[(r_i - \mu_{r_i})(r_j - \mu_{r_j})] + b_{r_j} b_{r_i} E[(r_i - \mu_{r_i})(r_k - \mu_{r_k})] \\
&\quad + b_{r_j} b_{r_k} E[(r_i - \mu_{r_i})(r_l - \mu_{r_l})] + b_{r_i} b_{r_l} E[(r_j - \mu_{r_j})(r_k - \mu_{r_k})] \\
&\quad + b_{r_i} b_{r_k} E[(r_j - \mu_{r_j})(r_l - \mu_{r_l})] + b_{r_i} b_{r_j} E[(r_k - \mu_{r_k})(r_l - \mu_{r_l})] \\
&\quad + b_{r_i} b_{r_j} b_{r_k} b_{r_l} \\
&= \sigma_{r_i, r_j} \sigma_{r_k, r_l} + \sigma_{r_i, r_k} \sigma_{r_j, r_l} + \sigma_{r_i, r_l} \sigma_{r_j, r_k} + \sigma_{r_i, r_j} b_{r_k} b_{r_l} + \sigma_{r_i, r_k} b_{r_j} b_{r_l} \\
&\quad + \sigma_{r_i, r_l} b_{r_j} b_{r_k} + \sigma_{r_j, r_k} b_{r_i} b_{r_l} + \sigma_{r_j, r_l} b_{r_i} b_{r_k} + \sigma_{r_k, r_l} b_{r_i} b_{r_j} \\
&\quad + b_{r_i} b_{r_j} b_{r_k} b_{r_l},
\end{aligned} \tag{7}$$

where  $i, j, k$ , and  $l$  can be  $xc$  or  $yc$ , and Anderson's (1958, p.39, Equations 25 & 26) formulae<sup>[4]</sup> were applied to the last equality in (7). By the same fashion, we further have

$$\begin{aligned} E(\Delta r_i \Delta r_j \Delta r_k) &= E[(r_i - \rho_i)(r_j - \rho_j)(r_k - \rho_k)] \\ &= \sigma_{r_i, r_j} b_{r_k} + \sigma_{r_i, r_k} b_{r_j} + \sigma_{r_j, r_k} b_{r_i} + b_{r_i} b_{r_j} b_{r_k}. \end{aligned} \quad (8)$$

Then, in (7) letting  $i = j = k = xc$  and  $l = yc$ , letting  $i = xc$  and  $j = k = l = yc$ , and letting  $i = j = xc$  and  $k = l = yc$ , respectively, give us the desired expressions for some of the higher order product-moments as follows:

$$\left. \begin{aligned} E[(\Delta r_{xc})^3 (\Delta r_{yc})] &= 3\sigma_{r_{xc}}^{(2)} \sigma_{r_{xc}, r_{yc}} + 3\sigma_{r_{xc}}^{(2)} b_{r_{xc}} b_{r_{yc}} + 3\sigma_{r_{xc}, r_{yc}} b_{r_{xc}}^2 + b_{r_{xc}}^3 b_{r_{yc}}; \\ E[(\Delta r_{xc}) (\Delta r_{yc})^3] &= 3\sigma_{r_{yc}}^{(2)} \sigma_{r_{xc}, r_{yc}} + 3\sigma_{r_{yc}}^{(2)} b_{r_{xc}} b_{r_{yc}} + 3\sigma_{r_{xc}, r_{yc}} b_{r_{yc}}^2 + b_{r_{xc}} b_{r_{yc}}^3; \\ E[(\Delta r_{xc})^2 (\Delta r_{yc})^2] &= \sigma_{r_{xc}}^{(2)} \sigma_{r_{yc}}^{(2)} + 2\sigma_{r_{xc}, r_{yc}}^2 + \sigma_{r_{xc}}^{(2)} b_{r_{yc}}^2 + \sigma_{r_{yc}}^{(2)} b_{r_{xc}}^2 \\ &\quad + 4\sigma_{r_{xc}, r_{yc}} b_{r_{xc}} b_{r_{yc}} + b_{r_{xc}}^2 b_{r_{yc}}^2. \end{aligned} \right\} \quad (9)$$

And, in (8) letting  $i = j = xc$  and  $k = yc$  and letting  $i = xc$  and  $j = k = yc$ , respectively, give us the desired expressions for the rest of the higher order product-moments as follows:

$$\left. \begin{aligned} E[(\Delta r_{xc})^2 (\Delta r_{yc})] &= \sigma_{r_{xc}}^{(2)} b_{r_{yc}} + 2\sigma_{r_{xc}, r_{yc}} b_{r_{xc}} + b_{r_{xc}}^2 b_{r_{yc}}; \\ E[(\Delta r_{xc}) (\Delta r_{yc})^2] &= \sigma_{r_{yc}}^{(2)} b_{r_{xc}} + 2\sigma_{r_{xc}, r_{yc}} b_{r_{yc}} + b_{r_{xc}} b_{r_{yc}}^2. \end{aligned} \right\} \quad (10)$$

<sup>[4]</sup> Anderson's equations were tactically used only for obtaining the higher order product-moments, although these equations are based on the normal distribution.

Applying (6), (9), and (10) to (4), I obtain the approximate first four non-central moments of  $r_{xc}r_{yc}$  as follows, in terms of  $\rho_{xc}$  and  $\rho_{yc}$  and the moments and the covariance of  $r_{xc}$  and  $r_{yc}$  that are the functions of  $\rho_{xc}$ ,  $\rho_{yc}$ , and  $\rho_{xy}$  (cf., Equation 5):

$$\begin{aligned}
\mu'_1 &= E(r_{xc}r_{yc}) = (b_{r_{xc}} + \rho_{xc})(b_{r_{yc}} + \rho_{yc}) + \sigma_{r_{xc}, r_{yc}}; \\
\mu'_2 &= E(r_{xc}r_{yc})^2 = [(b_{r_{xc}} + \rho_{xc})^2 + \sigma_{r_{xc}}^{(2)}][(b_{r_{yc}} + \rho_{yc})^2 + \sigma_{r_{yc}}^{(2)}] \\
&\quad + 4(b_{r_{xc}} + \rho_{xc})(b_{r_{yc}} + \rho_{yc})\sigma_{r_{xc}, r_{yc}} + 2\sigma_{r_{xc}, r_{yc}}^2; \\
\mu'_3 &= E(r_{xc}r_{yc})^3 = -b_{r_{yc}}^3\rho_{xc}^3 + b_{r_{xc}}^3(3b_{r_{yc}} - \rho_{yc})\rho_{yc}^2 + \rho_{xc}^3\rho_{yc}^3 + 3\rho_{xc}\rho_{yc}^3\sigma_{r_{xc}}^{(2)} + 3\rho_{xc}^3\rho_{yc}\sigma_{r_{yc}}^{(2)} \\
&\quad + 9\rho_{xc}\rho_{yc}\sigma_{r_{xc}}^{(2)}\sigma_{r_{yc}}^{(2)} + \rho_{yc}^3\sigma_{r_{xc}}^{(3)} + \rho_{xc}^3\sigma_{r_{yc}}^{(3)} + 9[\rho_{yc}^2\sigma_{r_{xc}}^{(2)} + \rho_{xc}^2(\rho_{yc}^2 + \sigma_{r_{yc}}^{(2)})]\sigma_{r_{xc}, r_{yc}} \\
&\quad + 18\rho_{xc}\rho_{yc}\sigma_{r_{xc}, r_{yc}}^2 + 3b_{r_{xc}}^2\rho_{xc}(\rho_{xc}^2\rho_{yc} + 3\rho_{yc}\sigma_{r_{xc}}^{(2)} + 3\rho_{xc}\sigma_{r_{xc}, r_{yc}}) \\
&\quad + 3b_{r_{xc}}^2\rho_{yc}[\rho_{xc}(3b_{r_{yc}}^2 + 3b_{r_{xc}}\rho_{yc} + \rho_{yc}^2 + 3\sigma_{r_{yc}}^{(2)}) + 3\rho_{yc}\sigma_{r_{xc}, r_{yc}}] \\
&\quad + 3b_{r_{yc}}\rho_{xc}[3\rho_{yc}^2\sigma_{r_{xc}}^{(2)} + \rho_{xc}^2(\rho_{yc}^2 - \sigma_{r_{yc}}^{(2)}) + 6\rho_{xc}\rho_{yc}\sigma_{r_{xc}, r_{yc}}] \\
&\quad + 3b_{r_{xc}}\{b_{r_{yc}}^3\rho_{xc}^2 + 3b_{r_{yc}}^2\rho_{xc}^2\rho_{yc} + 3b_{r_{yc}}[\rho_{yc}^2\sigma_{r_{xc}}^{(2)} + \rho_{xc}^2(\rho_{yc}^2 + \sigma_{r_{yc}}^{(2)}) + 4\rho_{xc}\rho_{yc}\sigma_{r_{xc}, r_{yc}}] \\
&\quad + \rho_{yc}[-\rho_{yc}^2\sigma_{r_{xc}}^{(2)} + \rho_{xc}^2(\rho_{yc}^2 + 3\sigma_{r_{yc}}^{(2)}) + 6\rho_{xc}\rho_{yc}\sigma_{r_{xc}, r_{yc}}]\}; \\
\mu'_4 &= E(r_{xc}r_{yc})^4 = b_{r_{yc}}^4\rho_{xc}^4 + 4b_{r_{yc}}^3(4b_{r_{xc}} - \rho_{xc})\rho_{xc}^3\rho_{yc} + b_{r_{xc}}^4\rho_{yc}^4 - 4b_{r_{xc}}^3\rho_{xc}\rho_{yc}^4 + \rho_{xc}^4\rho_{yc}^4 \\
&\quad + 6\rho_{xc}^2\rho_{yc}^4\sigma_{r_{xc}}^{(2)} + 6\rho_{xc}^4\rho_{yc}^2\sigma_{r_{yc}}^{(2)} + 36\rho_{xc}^2\rho_{yc}^2\sigma_{r_{xc}}^{(2)}\sigma_{r_{yc}}^{(2)} + 4\rho_{xc}\rho_{yc}^4\sigma_{r_{xc}}^{(3)} + 4\rho_{xc}^4\rho_{yc}\sigma_{r_{yc}}^{(3)} \\
&\quad + \rho_{yc}^4\sigma_{r_{xc}}^{(4)} + \rho_{xc}^4\sigma_{r_{yc}}^{(4)} + 16\rho_{xc}\rho_{yc}[3\rho_{yc}^2\sigma_{r_{xc}}^{(2)} + \rho_{xc}^2(\rho_{yc}^2 + 3\sigma_{r_{yc}}^{(2)})]\sigma_{r_{xc}, r_{yc}} \\
&\quad + \rho_{xc}^4(b_{r_{yc}}^4 + 6b_{r_{yc}}^2\sigma_{r_{yc}}^{(2)} - 4b_{r_{yc}}\sigma_{r_{yc}}^{(3)} + \sigma_{r_{yc}}^{(4)}) + 16\rho_{xc}^3\rho_{yc}^3(b_{r_{xc}}b_{r_{yc}} + \sigma_{r_{xc}, r_{yc}})
\end{aligned} \tag{11}$$

$$\begin{aligned}
& + 72\rho_{xc}^2\rho_{yc}^2\sigma_{r_{xc},r_{yc}}^2 + 6b_{r_{yc}}^2\rho_{xc}^2[6b_{r_{xc}}^2\rho_{yc}^2 + 4b_{r_{xc}}\rho_{xc}\rho_{yc}^2 + 6\rho_{yc}^2\sigma_{r_{xc}}^{(2)} + \rho_{xc}^2(\rho_{yc}^2 + \sigma_{r_{yc}}^{(2)}) \\
& + 8\rho_{xc}\rho_{yc}\sigma_{r_{xc},r_{yc}}] + 6b_{r_{xc}}^2\rho_{yc}^2[\rho_{yc}^2\sigma_{r_{xc}}^{(2)} + \rho_{xc}^2(\rho_{yc}^2 + 6\sigma_{r_{yc}}^{(2)}) + 8\rho_{xc}\rho_{yc}\sigma_{r_{xc},r_{yc}}] \\
& + 4b_{r_{xc}}\rho_{yc}^2[-3\rho_{xc}\rho_{yc}^2\sigma_{r_{xc}}^{(2)} + \rho_{xc}^3(\rho_{yc}^2 + 6\sigma_{r_{yc}}^{(2)}) - \rho_{yc}^2\sigma_{r_{xc}}^{(3)} + 12\rho_{xc}^2\rho_{yc}\sigma_{r_{xc},r_{yc}}] \\
& + 4b_{r_{yc}}\rho_{xc}\{4b_{r_{xc}}^3\rho_{yc}^3 + 6b_{r_{xc}}^2\rho_{xc}\rho_{yc}^3 + 4b_{r_{xc}}\rho_{yc}[3\rho_{yc}^2\sigma_{r_{xc}}^{(2)} + \rho_{xc}^2(\rho_{yc}^2 + 3\sigma_{r_{yc}}^{(2)}) \\
& + 9\rho_{xc}\rho_{yc}\sigma_{r_{xc},r_{yc}}] + \rho_{xc}[6\rho_{yc}^3\sigma_{r_{xc}}^{(2)} + \rho_{xc}^2(\rho_{yc}^3 - 3\rho_{yc}\sigma_{r_{yc}}^{(2)} - \sigma_{r_{yc}}^{(3)}) + 12\rho_{xc}\rho_{yc}^2\sigma_{r_{xc},r_{yc}}]\} \}.
\end{aligned} \tag{11'}$$

The derivations for the formulae (11) and (11') were done by *Mathematica* (Wolfram, 1999). We could further substitute (5) into (11) and (11') to obtain direct expressions for the first four moments of  $r_{xc}r_{yc}$  in terms of  $\rho_{xc}$ ,  $\rho_{yc}$ , and  $\rho_{xy}$  as well as the sample size  $N$ , but this is not advisable because the formulae would be much messier. For the central moments about the mean, substitution in Kendall and Stuart's equations (Kendall & Stuart, 1977, Equation 3.9, p. 58):

$$\begin{aligned}
\mu_2 &= \mu'_2 - \mu'_1{}^2 \\
&= (b_{r_{xc}} + \rho_{xc})^2\sigma_{r_{yc}}^{(2)} + \sigma_{r_{xc}}^{(2)}[(b_{r_{yc}} + \rho_{yc})^2 + \sigma_{r_{yc}}^{(2)}] \\
&\quad + 2(b_{r_{xc}} + \rho_{xc})(b_{r_{yc}} + \rho_{yc})\sigma_{r_{xc},r_{yc}} + \sigma_{r_{xc},r_{yc}}^2; \\
\mu_3 &= \mu'_3 - 3\mu'_1\mu'_2 + 2\mu'_1{}^3 \\
&= -2b_{r_{yc}}^3\rho_{xc}^3 - 3b_{r_{yc}}^3\rho_{xc}\sigma_{r_{xc}}^{(2)} - 6b_{r_{yc}}\rho_{xc}^3\sigma_{r_{yc}}^{(2)} - 3b_{r_{yc}}\rho_{xc}\sigma_{r_{xc}}^{(2)}\sigma_{r_{yc}}^{(2)} + 6\rho_{xc}\rho_{yc}\sigma_{r_{xc}}^{(2)}\sigma_{r_{yc}}^{(2)}
\end{aligned} \tag{12}$$



$$\begin{aligned}
& + 3b_{r_{yc}}^2 \rho_{xc} \sigma_{r_{xc}}^{(2)} \sigma_{r_{yc}}^{(2)} - 3b_{r_{yc}} \rho_{xc} \rho_{yc} \sigma_{r_{xc}}^{(2)} \sigma_{r_{yc}}^{(2)} - 6\rho_{xc} \rho_{yc}^2 \sigma_{r_{xc}}^{(2)} \sigma_{r_{yc}}^{(2)} - b_{r_{yc}} \rho_{yc}^3 \sigma_{r_{xc}}^{(3)} - 2\rho_{yc}^4 \sigma_{r_{xc}}^{(3)} \\
& - b_{r_{yc}} \rho_{yc}^3 \sigma_{r_{xc}}^{(3)} - 2\rho_{yc}^4 \sigma_{r_{xc}}^{(3)} - b_{r_{yc}} \rho_{xc}^3 \sigma_{r_{yc}}^{(3)} - \rho_{xc}^3 \rho_{yc} \sigma_{r_{yc}}^{(3)} + 3[2b_{r_{yc}}^3 \rho_{xc}^2 - 6b_{r_{yc}}^2 \rho_{xc}^2 \rho_{yc} \\
& + b_{r_{yc}}^3 \sigma_{r_{xc}}^{(2)} + 3b_{r_{yc}}^2 \rho_{yc} \sigma_{r_{xc}}^{(2)} - 3b_{r_{yc}} \rho_{yc}^2 \sigma_{r_{xc}}^{(2)} - \rho_{yc}^3 \sigma_{r_{xc}}^{(2)} - (b_{r_{yc}} + \rho_{yc})(3\rho_{xc}^2 - \sigma_{r_{xc}}^{(2)}) \sigma_{r_{yc}}^{(2)}] \sigma_{r_{xc}, r_{yc}} \\
& + 3\rho_{xc} (8b_{r_{yc}}^2 - 2b_{r_{yc}} \rho_{yc} - 4\rho_{yc}^4 + \sigma_{r_{yc}}^{(2)}) \sigma_{r_{xc}, r_{yc}}^2 + 9(b_{r_{yc}} + \rho_{yc}) \sigma_{r_{xc}, r_{yc}}^2 \} - 4b_{r_{yc}} [2\rho_{xc}^4 \sigma_{r_{yc}}^{(3)} \\
& + 3\rho_{xc}^3 \sigma_{r_{yc}}^{(2)} \sigma_{r_{xc}, r_{yc}} + 3\rho_{yc} \sigma_{r_{xc}}^{(2)} \sigma_{r_{xc}, r_{yc}}^2 + 6\rho_{xc}^2 \rho_{yc} (\sigma_{r_{xc}}^{(2)} \sigma_{r_{yc}}^{(2)} + 2\sigma_{r_{xc}, r_{yc}}^2) + \rho_{xc} (\rho_{yc}^3 \sigma_{r_{xc}}^{(3)} \\
& + 9\rho_{yc}^2 \sigma_{r_{xc}}^{(2)} \sigma_{r_{xc}, r_{yc}} - 3\sigma_{r_{xc}}^{(2)} \sigma_{r_{yc}}^{(2)} \sigma_{r_{xc}, r_{yc}} - 9\sigma_{r_{xc}, r_{yc}}^3)].
\end{aligned} \tag{12''}$$

Note that  $\mu_1 = \int_{-\infty}^{\infty} (r_{xc} r_{yc} - \mu'_1) dF = \int_{-\infty}^{\infty} [r_{xc} r_{yc} - E(r_{xc} r_{yc})] dF = \int_{-\infty}^{\infty} (r_{xc} r_{yc}) dF - E(r_{xc} r_{yc})$

$= 0.$

### 3.2 Pearson Distributions

The four moments only give us a general idea about the characteristics of the distribution of the product of two dependent correlation coefficients. To better understand the distribution of the product of two dependent correlation coefficients, we need to explore the shape of the distribution. Since the Pearson distribution family provides approximations to a wide variety of observed distributions using the only first four moments, in this thesis the Pearson distribution family is employed to obtain an approximate distribution for the product of the two dependent correlation coefficients.

There are three main types and some other uncommon types of frequency curves in the Pearson distribution family, which are characterized by  $\beta$ -coefficients:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}, \beta_2 = \frac{\mu_4}{\mu_2^2}. \quad (13)$$

In particular, after obtaining  $\beta_1$  and  $\beta_2$  via the first four moments, we can plot the couplet  $(\beta_1, \beta_2)$  on the  $(\beta_1, \beta_2)$  plane illustrated in Pearson and Hartley (1972, p. 78). From the  $(\beta_1, \beta_2)$  plane we know the type to which the observed distribution belongs. Instead of referencing the  $(\beta_1, \beta_2)$  plane, we can also distinguish the types of Pearson distributions by evaluating the criterion  $\kappa$  (*kappa*) (Kendall & Stuart, 1977, Equation 6.10):

$$\kappa = \frac{\beta_1(\beta_2 + 3)^2}{4(4\beta_2 - 3\beta_1)(2\beta_2 - 3\beta_1 - 6)}, \quad (14)$$

with the specifications illustrated in Figure 1 (Elderton & Johnson, 1969, p.49).

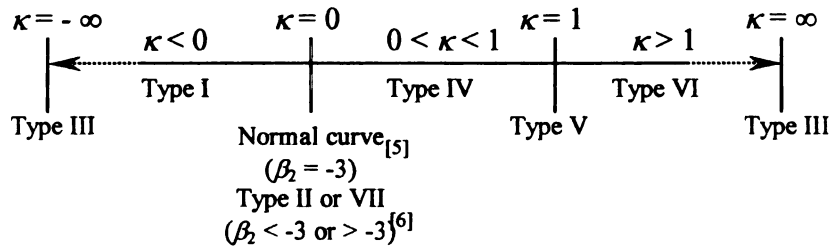


Figure 1. Specifications of  $\kappa$  for distinguishing the types of Pearson distributions.

<sup>[5], [6]</sup> Originally, they are " $\beta_2 = 3$ " and " $\beta_2 < 3$  or  $> 3$ " in Elderton and Johnson (1969, p.49), but according to the expression (14) above (from Kendall & Stuart, 1977, Equation 6.10) and the expression (4) in Elderton and Johnson (1969, p.41), they should be " $\beta_2 = -3$ " and " $\beta_2 < -3$  or  $> -3$ ", respectively.

Type I distributions are *beta distributions* and type III distributions are *gamma distributions*. Pearson and Hartley (1972, p. 261-285) have tables for percentage points of Pearson curves for given  $\beta_1$  and  $\beta_2$ . Elderton and Johnson (1969) provide mathematical expressions for Pearson distributions with formulae for the parameters in terms of the first four moments.

### 3.3 Approximate Distribution of $r_{xc}r_{yc}$

In order to apply the first four moments of  $r_{xc}r_{yc}$  to Pearson distributions, I begin by noting that there are three conditions for the dependent population correlations  $\rho_{xc}$ ,  $\rho_{yc}$ , and  $\rho_{xy}$  in this particular study:

(i)  $\rho_{xc}$  and  $\rho_{yc}$  are dependent and, from (3),  $\rho_{xy}$  is correspondingly constrained by:

$$\rho_{xc}\rho_{yc} \pm \sqrt{(1 - \rho_{xc}^2)(1 - \rho_{yc}^2)} ;$$

(ii)  $\rho_{xc}$ ,  $\rho_{yc}$ , and  $\rho_{xy} \neq -1, 0$ , or  $1$ ;

(iii)  $\rho_{xc}\rho_{yc}\rho_{xy} > 0$ .

It is obvious that condition (i) is supported by Cohen and Cohen's constraint in equation (3) above. Since  $-1, 0, 1$  are extreme cases and trivial, condition (ii) is sensible. As for condition (iii), I will deal only with cases in which this condition holds for issues of confounding and indirect effects. The general interpretation of confounding applies when the impact of the confounding variable,  $K = \rho_{xc}\rho_{yc}$ , takes the same sign as the

relationship between  $X$  and  $Y$ ,  $\rho_{xy}$ . Therefore, the product of the two components,  $\rho_{xc}\rho_{yc}$  and  $\rho_{xy}$ , must be positive, showing that condition (iii) applies.

Under these three conditions, we might be able to theoretically prove that  $\kappa < 0$ , but for simplicity, I only numerically evaluate the criterion  $\kappa$ . Under the conditions above, I evaluated  $\kappa$  value for all possible conditional triplets of  $\rho_{xc}$ ,  $\rho_{yc}$ , and  $\rho_{xy}$  with an increment of .10 for each correlation coefficient and I found that the larger  $N$ , the more values of  $\kappa$  are negative. When  $N > 300$ <sup>[7]</sup>, all values of  $\kappa$  are negative under the three conditions. Note that the values of  $\kappa$  are approximate, since the  $\beta$ -coefficients are evaluated by the approximate moments. As  $N$  approaches infinity, the approximate values of  $\kappa$  will approach the true value that will be negative, as observed by examining the numerical trend of  $\kappa$  as  $N$  increases. Thus, by looking at Figure 1, we can conclude that the distribution of the product of two dependent correlation coefficients,  $r_{xc}r_{yc}$ , can be approximated by a Pearson Type I distribution. Correspondingly, the density function is (Elderton and Johnson, 1969; Kendall & Stuart, 1977):

$$f(k) = f_0 \left(1 + \frac{k}{a_1}\right)^{m_1} \left(1 - \frac{k}{a_2}\right)^{m_2}, \quad -a_1 \leq k \leq a_2; \quad \frac{m_1}{a_1} = \frac{m_2}{a_2},$$

where  $k = r_{xc}r_{yc}$ , and

$$f_0 = \frac{a_1^{m_1} a_2^{m_2}}{(a_1 + a_2)^{m_1 + m_2 + 1}} \cdot \frac{1}{B(m_1 + 1, m_2 + 1)},$$

---

<sup>[7]</sup> For  $N < 300$ ,  $\rho_{xc}$  or  $\rho_{yc}$  must be greater than .10 so that  $\kappa$  becomes negative. For smaller correlations, e.g.,  $\rho_{xc}$  and  $\rho_{yc} < .10$ , a rare case in confounding,  $0 < \kappa < 1$ ; and the distribution of the product of two dependent correlation coefficients,  $r_{xc}r_{yc}$ , can be approximated by a Pearson Type IV distribution.

$$a_1 + a_2 = \frac{1}{2} \sqrt{\mu_2 [\beta_1 (s+2)^2 + 16(s+1)]},$$

$$s = \frac{6(\beta_2 - \beta_1 - 1)}{6 + 3\beta_1 - 2\beta_2},$$

and  $m_1$  and  $m_2$  are given by

$$\frac{s-2}{2} \pm \frac{s(s+2)}{2} \sqrt{\frac{\beta_1}{\beta_1 (s+2)^2 + 16(s+1)}},$$

with  $m_1 < m_2$  if  $\mu_3 > 0$  and  $m_1 > m_2$  if  $\mu_3 < 0$ .

Note that here we got an approximate distribution for the product of two dependent zero-order correlation coefficients. The results hold for the product of two dependent partial correlation coefficients. It is evidenced by Fisher's (1924) founding that the distribution of the sample partial correlation is that of the zero-order correlation coefficient.

## Chapter 4

### SIMULATION STUDY

Equation (4) in Chapter 3 gives us the expressions for the first four moments of  $r_{xc}r_{yc}$  in terms of the moments and product-moments of the original correlation coefficients,  $r_{xc}$  and  $r_{yc}$ . Then, by applying approximations to the moments and product-moments of the original correlation coefficients, this current study obtains approximate expressions for the non-central moments or central moments of the product,  $r_{xc}r_{yc}$ . However, we do not know how accurate the approximation is. Therefore, a simulation study is conducted to check the accuracy of the approximate moments against the moments calculated from the simulated data.

#### 4.1 Simulation Design

The parameters in this simulation study are the sample size,  $N$ , and the population correlations,  $\rho_{xc}$ ,  $\rho_{yc}$ , and  $\rho_{xy}$ . For  $\rho_{xc}$ ,  $\rho_{yc}$ , and  $\rho_{xy}$ , I choose .10, .30, and .50 as small, medium, and large correlations (Cohen, 1988); and for the sample size, I select 28, 84, and 783, which correspond to a statistical power of .80 for the small, medium, and large correlations (Cohen & Cohen, 1983). Table 2 shows the parameter specifications for this simulation study.

Table 2  
*Parameter Specifications for the Simulation Study*

$\rho_{xc}$	$\rho_{yc}$	$\rho_{xy}$	$N$		
			28	84	783
.10	.10	.10			
		.30			
		.50			
	.30	.10			
		.30			
		.50			
	.50	.10			
		.30			
		.50			
.30	.30	.10			
		.30			
		.50			
	.50	.10			
		.30			
		.50			
.50	.50	.10			
		.30			
		.50			

As can be seen in Table 2, we do not need to include every possible combination of .10, .30, and .50, because  $\rho_{xc}$  and  $\rho_{yc}$  are symmetric in the mathematical expressions in Chapter 3. Therefore, I have removed the duplicate cases. In addition, under the condition where  $\rho_{xc}\rho_{yc}\rho_{xy} > 0$ , we can always change the sign of the relevant variable to have all three correlations positive. Thus, there are only 18 positive correlation triplets needed for this simulation study. Multiplied by the three magnitudes of  $N$ , we have  $18 \times 3 = 54$  cells to simulate.

For each of 54 cells (i.e., for each set of  $N$ ,  $\rho_{xc}$ ,  $\rho_{yc}$ ,  $\rho_{xy}$ ), by Cholesky factorization, I generate  $X$ ,  $Y$ , and  $C$  with the specified population correlations (cf., Browne, 1968) and with the sample size as  $N$ . For simplicity, I generate  $(X, Y, C)$  as a trivariate standard

normal distribution given values of  $\rho_{xc}$ ,  $\rho_{yc}$ , and  $\rho_{xy}$ . Then, I compute the pair of estimated correlation coefficients, which are defined as  $\tilde{r}_{xc}$  and  $\tilde{r}_{yc}$ . Within each cell, the same procedure is replicated 1000 times, resulting in 1000 pairs of  $\tilde{r}_{xc}$  and  $\tilde{r}_{yc}$  values. Directly multiplying  $\tilde{r}_{xc}$  and  $\tilde{r}_{yc}$  gives us a simulated distribution of the product  $r_{xc}r_{yc}$ , which serves as the true distribution of  $r_{xc}r_{yc}$ , with which the approximate moments<sup>[8]</sup> and distribution can be compared. Tables 3 to 14 show the comparisons of the first four moments.

To quantify the accuracy of the approximation method in this paper, by following Kendall and Stuart (1977, p. 247), the standard errors of the first four moments,  $\tilde{\mu}'_1$ ,  $\tilde{\mu}_2$ ,  $\tilde{\mu}_3$ , and  $\tilde{\mu}_4$ , of the simulated distribution of the product  $r_{xc}r_{yc}$  are also computed (see the numbers in the parentheses in Tables 3 to 14) by the following formulae:

$$\left. \begin{aligned} s.e.(\tilde{\mu}'_1) &= \sqrt{\frac{1}{1000} \mu_2} ; \\ s.e.(\tilde{\mu}_2) &= \sqrt{\frac{1}{1000} (\mu_4 - \mu_2^2)} ; \\ s.e.(\tilde{\mu}_3) &= \sqrt{\frac{1}{1000} (\mu_6 - \mu_3^2 - 6\mu_4\mu_2 + 9\mu_2^3)} ; \\ s.e.(\tilde{\mu}_4) &= \sqrt{\frac{1}{1000} (\mu_8 - \mu_4^2 - 8\mu_5\mu_3 + 16\mu_2\mu_3^2)} . \end{aligned} \right\} \quad (15)$$

---

<sup>[8]</sup> The approximate moments were calculated by SPSS (See Appendix A for the SPSS code). To make the code clearer, I computed Equations (6), (9), and (10) substituted them into (4), instead of using the messy Equations (12), (12'), and (12'').

here I simply use the formulae (15) above to calculate the standard errors instead.

Actually, the two approaches yield similar results, except for  $\tilde{\mu}_3$  for which the standard errors were a little overestimated by the formula in (15). For instance, for the cell of  $N = 84$ ,  $\rho_{xc} = .30$ ,  $\rho_{yc} = .30$ , and  $\rho_{xy} = .50$ , by the formulae (15) the standard errors of  $\tilde{\mu}'_1$ ,  $\tilde{\mu}_2$ ,  $\tilde{\mu}_3$ , and  $\tilde{\mu}_4$  are .001632, .000137, .000018, and .000004, respectively, while they are .001542, .000127, .000014, and .000003, respectively, if we generate 1000 of each of those moments. Thus, Tables 3 to 14 display in the parentheses the standard errors calculated only by the formulae (15).

In addition, to help to have a good look at how far the approximated values are from the simulated values relative to the standard errors (*s.e.*) of the simulated values, the standardized differences (Std. Diff. =  $\frac{\text{Approximated} - \text{Simulated}}{s.e.}$ ) are also listed in

Tables 3 to 14.

Table 3

*Simulation Results for the Mean  $\mu'_1$  ( $N = 28$ )*

$\rho_{xc}$	$\rho_{yc}$	$\rho_{xy}$	Simulated (s.e.)	Approximated (Std. Diff.)
.10	.10	.10	.016237540 (.001416792)	.015379623 ( -.605534834 )
		.30	.022705850 (.001482690)	.021156094 (-1.045232565 )
		.50	.029056640 (.001561022)	.026944329 (-1.353158933 )
	.30	.10	.035789020 (.002169760)	.034214339 ( -.725739594 )
		.30	.041859490 (.002253031)	.039543751 (-1.027832596 )
		.50	.047860880 (.002344481)	.044908457 (-1.259307843 )
	.50	.10	.054815860 (.003166590)	.053042589 ( -.559993936 )
		.30	.059943050 (.003241347)	.057454354 ( -.767796859 )
		.50	.065037510 (.003320096)	.061924942 ( -.937493265 )
	.30	.10	.093832050 (.002612237)	.091340180 ( -.953922018 )
		.30	.099503540 (.002755509)	.096269592 (-1.173629978 )
		.50	.105077490 (.002891617)	.101304886 (-1.304669237 )
.30	.50	.10	.151518190 (.003232683)	.148917726 ( -.804428991 )
		.30	.156342730 (.003393298)	.152976550 ( -.992008421 )
		.50	.161155260 (.003549875)	.157211844 (-1.110860634 )
	.50	.10	.248737050 (.003346673)	.245704306 ( -.906196698 )
		.30	.252917640 (.003566791)	.248939600 (-1.115299353 )
		.50	.257125970 (.003767061)	.252469012 (-1.236231048 )

Table 4

*Simulation Results for the Mean  $\mu'_1$  ( $N = 84$ )*

$\rho_{xc}$	$\rho_{yc}$	$\rho_{xy}$	Simulated (s.e.)	Approximated (Std. Diff.)
.10	.10	.10	.010398070 (.000618555)	.010916407 ( .837980770 )
		.30	.012648370 (.000655401)	.013098629 ( .686998108 )
		.50	.014908830 (.000694075)	.015285296 ( .542399656 )
	.30	.10	.030208880 (.001164637)	.030509873 ( .258443555 )
		.30	.032315040 (.001213070)	.032523206 ( .171602560 )
		.50	.034499610 (.001262882)	.034549873 ( .039800247 )
	.50	.10	.050552070 (.001821664)	.050096700 ( -.249974738 )
		.30	.052358840 (.001865880)	.051763367 ( -.319137803 )
		.50	.054340980 (.001914181)	.053452256 ( -.464284133 )
	.30	.10	.089548730 (.001436461)	.089522706 ( -.018116748 )
		.30	.091304370 (.001536766)	.091384928 ( .052420470 )
		.50	.093066460 (.001632057)	.093287151 ( .135222607 )
.30	.50	.10	.149496750 (.001868874)	.148693461 ( -.429825126 )
		.30	.150922470 (.001980651)	.150226795 ( -.351235456 )
		.50	.152451530 (.002090734)	.151826795 ( -.298811281 )
	.50	.10	.248913370 (.001942712)	.247612886 ( -.669416767 )
		.30	.249983780 (.002086890)	.248835108 ( -.550422870 )
		.50	.251040100 (.002226773)	.250168441 ( -.391444850 )

Table 5

*Simulation Results for the Mean  $\mu'_1$  ( $N = 783$ )*

$\rho_{xc}$	$\rho_{yc}$	$\rho_{xy}$	Simulated (s.e.)	Approximated (Std. Diff.)
.10	.10	.10	.010037140 (.000167601 )	.010105409 (.407331195 )
		.30	.010288380 (.000181604 )	.010354331 (.363158371 )
		.50	.010545220 (.000194859 )	.010603761 (.300427607 )
	.30	.10	.030017610 (.000358608 )	.030060721 (.120217478 )
		.30	.030253530 (.000376577 )	.030290379 (.097852579 )
		.50	.030497610 (.000393853 )	.030521557 (.060801909 )
	.50	.10	.049967830 (.000568102 )	.050015047 (.083113581 )
		.30	.050166610 (.000584003 )	.050205161 (.066011599 )
		.50	.050371640 (.000599416 )	.050397810 (.043659134 )
.30	.30	.10	.089681690 (.000442595 )	.089953037 (.613082460 )
		.30	.089872890 (.000482732 )	.090165458 (.606067374 )
		.50	.090092450 (.000520759 )	.090382441 (.556862130 )
	.50	.10	.149624530 (.000572730 )	.149862675 (.415806530 )
		.30	.149788470 (.000616166 )	.150037580 (.404290690 )
		.50	.149987730 (.000658331 )	.150220089 (.352951511 )
.50	.50	.10	.249354340 (.000582864 )	.249745934 (.671844984 )
		.30	.249408930 (.000638882 )	.249885351 (.745710667 )
		.50	.249522490 (.000695198 )	.250037442 (.740727301 )

Table 6

*Simulation Results for the Variance  $\mu_2$  ( $N = 28$ )*

$\rho_{xc}$	$\rho_{yc}$	$\rho_{xy}$	Simulated (s.e.)	Approximated (Std. Diff.)
.10	.10	.10	.002007300 (.000145714 )	.002082915 ( .518928191 )
		.30	.002198370 (.000168953 )	.002259304 ( .360657337 )
		.50	.002436790 (.000193817 )	.002502991 ( .341563738 )
	.30	.10	.004707860 (.000262925 )	.004732986 ( .095563454 )
		.30	.005076150 (.000288296 )	.005094097 ( .062251932 )
		.50	.005496590 (.000319352 )	.005514970 ( .057553959 )
	.50	.10	.010027290 (.000479899 )	.010002292 ( -.052090086 )
		.30	.010506330 (.000504464 )	.010463308 ( -.085282518 )
		.50	.011023040 (.000542946 )	.010970180 ( -.097357668 )
.30	.30	.10	.006823780 (.000324890 )	.006634122 ( -.583760512 )
		.30	.007592830 (.000360409 )	.007529351 ( -.176130254 )
		.50	.008361450 (.000398414 )	.008493985 ( .332656208 )
	.50	.10	.010450240 (.000464233 )	.010310752 ( -.300469693 )
		.30	.011514470 (.000494732 )	.011511735 ( -.005528245 )
		.50	.012601610 (.000532917 )	.012800062 ( .372388203 )
.50	.50	.10	.011200220 (.000472091 )	.010941769 ( -.547459824 )
		.30	.012722000 (.000519374 )	.012522676 ( -.383777661 )
		.50	.014190750 (.000561380 )	.014271177 ( .143266703 )

Table 7  
Simulation Results for the Variance  $\mu_2$  ( $N = 84$ )

$\rho_{xc}$	$\rho_{yc}$	$\rho_{xy}$	Simulated (s.e.)	Approximated (Std. Diff.)
.10	.10	.10	.000382610 (.000030085)	.000394662 (.400597850)
		.30	.000429550 (.000036380)	.000447068 (.481529421)
		.50	.000481740 (.000042849)	.000509135 (.639344415)
	.30	.10	.001356380 (.000077782)	.001320810 (-.457303157)
		.30	.001471540 (.000088283)	.001447716 (-.269860865)
		.50	.001594870 (.000099699)	.001583651 (-.112528167)
	.50	.10	.003318460 (.000161163)	.003138425 (-1.117101076)
		.30	.003481510 (.000173131)	.003308192 (-1.001078030)
		.50	.003664090 (.000186859)	.003485889 (-.953663857)
.30	.30	.10	.002063420 (.000104842)	.001977716 (-.817455669)
		.30	.002361650 (.000121537)	.002314607 (-.387068319)
		.50	.002663610 (.000137272)	.002665894 (.016638551)
	.50	.10	.003492690 (.000163370)	.003225462 (-1.635717828)
		.30	.003922980 (.000183839)	.003683807 (-1.300989915)
		.50	.004371170 (.000204267)	.004167092 (-.999076307)
	.50	.10	.003774130 (.000171079)	.003398698 (-2.194491155)
		.30	.004355110 (.000195961)	.004005468 (-1.784244005)
		.50	.004958520 (.000220055)	.004670806 (-1.307466027)

Table 8  
Simulation Results for the Variance  $\mu_2$  ( $N = 783$ )

$\rho_{xc}$	$\rho_{yc}$	$\rho_{xy}$	Simulated (s.e.)	Approximated (Std. Diff.)
.10	.10	.10	.000028090 (.000001459)	.000028980 (.609886393)
		.30	.000032980 (.000001747)	.000034073 (.625794230)
		.50	.000037970 (.000002035)	.000039300 (.653585335)
	.30	.10	.000128600 (.000006030)	.000130418 (.301481748)
		.30	.000141810 (.000006616)	.000144278 (.373008067)
		.50	.000155120 (.000007225)	.000158337 (.445249186)
	.50	.10	.000322740 (.000014903)	.000328215 (.367377356)
		.30	.000341060 (.000015682)	.000347269 (.395920301)
		.50	.000359300 (.000016454)	.000366650 (.446687221)
	.30	.10	.000195890 (.000008773)	.000202016 (.698282409)
		.30	.000233030 (.000010558)	.000240276 (.686304995)
		.50	.000271190 (.000012356)	.000279452 (.668649487)
.30	.30	.10	.000328020 (.000014938)	.000336743 (.583938073)
		.30	.000379660 (.000017344)	.000389197 (.549883427)
		.50	.000433400 (.000019839)	.000443997 (.534151916)
	.50	.10	.000339730 (.000015007)	.000353338 (.906795524)
		.30	.000408170 (.000018179)	.000422995 (.815510421)
		.50	.000483300 (.000021619)	.000499029 (.727550288)

Table 9

*Simulation Results for the Third Moment  $\mu_3$  ( $N = 28$ )*

$\rho_{xc}$	$\rho_{yc}$	$\rho_{xy}$	Simulated (s.e.)	Approximated (Std. Diff.)
.10	.10	.10	.000102612 (.000022926)	.000081765 ( -.909316933)
		.30	.000158858 (.000030373)	.000086309 (-2.388601719)
		.50	.000213329 (.000037334)	.000088324 (-3.348288423)
	.30	.10	.000229896 (.000040731)	.000240334 ( .256266726)
		.30	.000342621 (.000049056)	.000347121 ( .091731898)
		.50	.000460741 (.000058132)	.000460429 ( -.005367096)
	.50	.10	.000312498 (.000091352)	.000289363 ( -.253251160)
		.30	.000528974 (.000102736)	.000538633 ( .094017676)
		.50	.000761741 (.000122471)	.000801008 ( .320622841)
	.30	.10	.000400153 (.000055523)	.000532525 ( 2.384093079)
		.30	.000514982 (.000066161)	.000707575 ( 2.910974743)
		.50	.000637833 (.000078976)	.000911491 ( 3.465077998)
.30	.50	.10	.000332125 (.000086245)	.000526148 ( 2.249672445)
		.30	.000484532 (.000095090)	.000778373 ( 3.090135661)
		.50	.000656781 (.000107726)	.001073164 ( 3.865204315)
	.10	.10	.000186486 (.000080016)	.000490178 ( 3.795390922)
		.30	.000223523 (.000089843)	.000713757 ( 5.456563116)
		.50	.000281911 (.000100228)	.000994611 ( 7.110787405)

Table 10

*Simulation Results for the Third Moment  $\mu_3$  ( $N = 84$ )*

$\rho_{xc}$	$\rho_{yc}$	$\rho_{xy}$	Simulated (s.e.)	Approximated (Std. Diff.)
.10	.10	.10	.000010500 (.000002313)	.000009451 ( -.453523562)
		.30	.000015611 (.000003300)	.000012091 (-1.066666667)
		.50	.000020994 (.000004420)	.000014954 (-1.366515837)
	.30	.10	.000032053 (.000008501)	.000026099 ( -.700388190)
		.30	.000047655 (.000011095)	.000040805 ( -.617395223)
		.50	.000064406 (.000014054)	.000056834 ( -.538778995)
	.50	.10	.000051326 (.000020603)	.000030707 (-1.000776586)
		.30	.000082326 (.000024813)	.000062103 ( -.815016322)
		.50	.000116846 (.000029601)	.000095498 ( -.721191852)
	.30	.10	.000052947 (.000011332)	.000055051 ( .185668902)
		.30	.000072472 (.000014632)	.000077684 ( .356205577)
		.50	.000093117 (.000017853)	.000104548 ( .640284546)
.30	.50	.10	.000054061 (.000019433)	.000051763 ( -.118252457)
		.30	.000075321 (.000024430)	.000083043 ( .316086779)
		.50	.000098774 (.000029508)	.000120181 ( .725464281)
	.10	.10	.000027162 (.000017725)	.000045863 ( 1.055063470)
		.30	.000031018 (.000022389)	.000072871 ( 1.869355487)
		.50	.000041201 (.000027059)	.000107492 ( 2.449868805)

Table 11  
Simulation Results for the Third Moment  $\mu_3$  ( $N = 783$ )

$\rho_{xc}$	$\rho_{yc}$	$\rho_{xy}$	Simulated (s.e.)	Approximated (Std. Diff.)
.10	.10	.10	.000000108 (.000000020)	.000000111 (.150000000)
		.30	.000000151 (.000000026)	.000000155 (.153846154)
		.50	.000000197 (.000000032)	.000000205 (.250000000)
	.30	.10	.000000127 (.000000123)	.000000299 (1.398373984)
		.30	.000000288 (.000000142)	.000000485 (1.387323944)
		.50	.000000461 (.000000164)	.000000692 (1.408536585)
	.50	.10	-.000000203 (.000000467)	.000000348 (1.179871520)
		.30	.000000128 (.000000503)	.000000732 (1.200795229)
		.50	.000000506 (.000000542)	.000001144 (1.177121771)
.30	.30	.10	.000000605 (.000000214)	.000000615 (.046728972)
		.30	.000000801 (.000000293)	.000000895 (.320819113)
		.50	.000000992 (.000000373)	.000001231 (.640750670)
	.50	.10	.000000137 (.000000456)	.000000565 (.938596491)
		.30	.000000275 (.000000573)	.000000945 (1.169284468)
		.50	.000000389 (.000000700)	.000001401 (1.445714286)
.50	.50	.10	.000000545 (.000000449)	.000000490 (-.122494432)
		.30	.000000710 (.000000606)	.000000814 (.171617162)
		.50	.000000774 (.000000798)	.000001235 (.577694236)

Table 12  
Simulation Results for the Fourth Moment  $\mu_4$  ( $N = 28$ )

$\rho_{xc}$	$\rho_{yc}$	$\rho_{xy}$	Simulated (s.e.)	Approximated (Std. Diff.)
.10	.10	.10	.000025395 (.000005791)	.000000932 (-4.224313590)
		.30	.000033559 (.000008163)	-.000001358 (-4.277471518)
		.50	.000043744 (.000010421)	-.000004857 (-4.663755878)
	.30	.10	.000091691 (.000012082)	.000038147 (-4.431716603)
		.30	.000109363 (.000015133)	.000036404 (-4.821185489)
		.50	.000132794 (.000018452)	.000028269 (-5.664697594)
	.50	.10	.000332066 (.000036967)	.000252585 (-2.150052750)
		.30	.000366214 (.000042815)	.000267453 (-2.306691580)
		.50	.000417883 (.000053603)	.000271646 (-2.728149544)
.30	.30	.10	.000152673 (.000018407)	.000105481 (-2.563807247)
		.30	.000188228 (.000022664)	.000125898 (-2.750176491)
		.50	.000229483 (.000028618)	.000141015 (-3.091341114)
	.50	.10	.000325797 (.000034324)	.000300419 (-.739366041)
		.30	.000378532 (.000038363)	.000369983 (-.222844929)
		.50	.000444161 (.000044656)	.000440353 (-.085274095)
.50	.50	.10	.000349379 (.000031879)	.000374644 (.792527996)
		.30	.000432846 (.000036435)	.000486989 (1.486016193)
		.50	.000517932 (.000041225)	.000619684 (2.468211037)

Table 13

*Simulation Results for the Fourth Moment  $\mu_4$  ( $N = 84$ )*

$\rho_{xc}$	$\rho_{yc}$	$\rho_{xy}$	Simulated (s.e.)	Approximated (Std. Diff.)
.10	.10	.10	.000001057 (.000000269)	.000000160 (-3.334572491)
		.30	.000001517 (.000000437)	.000000094 (-3.256292906)
		.50	.000002080 (.000000652)	-.000000048 (-3.263803681)
	.30	.10	.000007925 (.000001627)	.000004301 (-2.227412415)
		.30	.000010005 (.000002306)	.000004813 (-2.251517780)
		.50	.000012543 (.000003136)	.000005068 (-2.383609694)
	.50	.10	.000037124 (.000005280)	.000027981 (-1.731628788)
		.30	.000042258 (.000006742)	.000030662 (-1.719964402)
		.50	.000048534 (.000008487)	.000032974 (-1.833392247)
	.30	.10	.000015310 (.000002288)	.000010897 (-1.928758741)
		.30	.000020430 (.000003221)	.000014492 (-1.843526855)
		.50	.000026042 (.000004214)	.000018342 (-1.827242525)
.30	.50	.10	.000039028 (.000004820)	.000030707 (-1.726348548)
		.30	.000049363 (.000006550)	.000039828 (-1.455725191)
		.50	.000061049 (.000008456)	.000050142 (-1.289853359)
	.50	.10	.000043661 (.000004284)	.000035330 (-1.944677871)
		.30	.000057562 (.000005858)	.000048855 (-1.486343462)
		.50	.000073253 (.000007544)	.000065800 (-.987937434)

Table 14

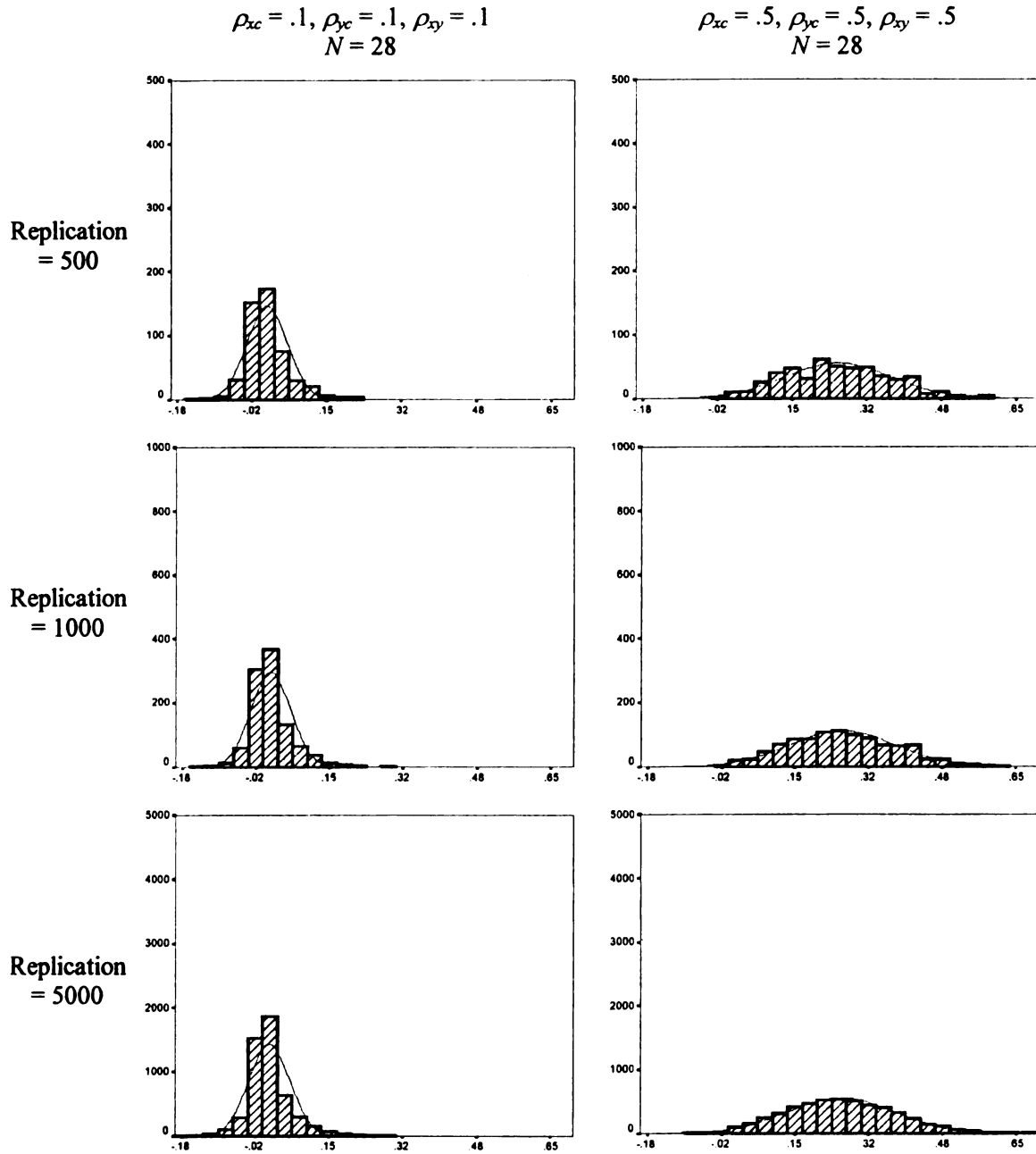
*Simulation Results for the Fourth Moment  $\mu_4$  ( $N = 783$ )*

$\rho_{xc}$	$\rho_{yc}$	$\rho_{xy}$	Simulated (s.e.)	Approximated (Std. Diff.)
.10	.10	.10	.000000003 (.000000001)	.000000002 (-1.000000000)
		.30	.000000004 (.000000001)	.000000003 (-1.000000000)
		.50	.000000006 (.000000001)	.000000004 (-2.000000000)
	.30	.10	.000000053 (.000000006)	.000000050 (-.500000000)
		.30	.000000064 (.000000007)	.000000061 (-.428571429)
		.50	.000000077 (.000000008)	.000000072 (-.625000000)
	.50	.10	.000000327 (.000000033)	.000000321 (-.181818182)
		.30	.000000364 (.000000037)	.000000359 (-.135135135)
		.50	.000000401 (.000000041)	.000000399 (-.048780488)
	.30	.10	.000000116 (.000000012)	.000000122 (.500000000)
		.30	.000000166 (.000000019)	.000000171 (.263157895)
		.50	.000000227 (.000000025)	.000000230 (.120000000)
.30	.50	.10	.000000332 (.000000033)	.000000340 (.242424242)
		.30	.000000447 (.000000044)	.000000453 (.136363636)
		.50	.000000583 (.000000058)	.000000589 (.103448276)
	.50	.10	.000000342 (.000000032)	.000000375 (1.031250000)
		.30	.000000499 (.000000047)	.000000538 (.829787234)
		.50	.000000703 (.000000069)	.000000748 (.652173913)

## 4.2 Simulation results

The simulation results are listed in Tables 3 through 14, and the first set of the results is the estimated first four moments of the product of two simulated correlations,  $\tilde{r}_{xc} \tilde{r}_{yc}$  (Under the “simulated” column in Tables 3 to 14). To establish whether 1000 replications (numbers of simulated correlations per cell) are enough, Figure 2 displays the distributions of the simulated product  $\tilde{r}_{xc} \tilde{r}_{yc}$  for selected cells with different replications. From the figure 2, we can see that there is not much difference in the shapes of the distributions across the different numbers of replications, verifying that 1000 replications are adequate for this simulation design. On the other hand, as the population correlations,  $\rho$ s, become bigger, the distributions are more spread, while as the sample size  $N$  becomes larger, the distributions are less spread.

Now, we turn to the comparison of approximated values and those obtained via Frank’s method for the first four moments against the simulated values. Comparing to the simulated values, Figure 3 shows the most noticeable fact that the approximated values are much more accurate than Frank’s values which consistently, heavily overestimate all four moments. For the approximated values themselves, most of them are within or very close to one standard error of the simulated values, except for the third and the fourth moments when  $N$  is 28 or 84. The inaccuracy might come from the lower-order approximation of  $\sigma_{r_{xc}, r_{yc}}$  (cf., Equation 5) which is heavily involved in the third and the fourth moments. In Chapter 5, I will introduce a regression approach to correct the inaccuracy for the third and the fourth moment when  $N = 28$  or  $84$ .



*Figure 2.* The distributions of the product of simulated correlations  $\tilde{r}_{xc}\tilde{r}_{yc}$  with different numbers of replications for the selected cells.

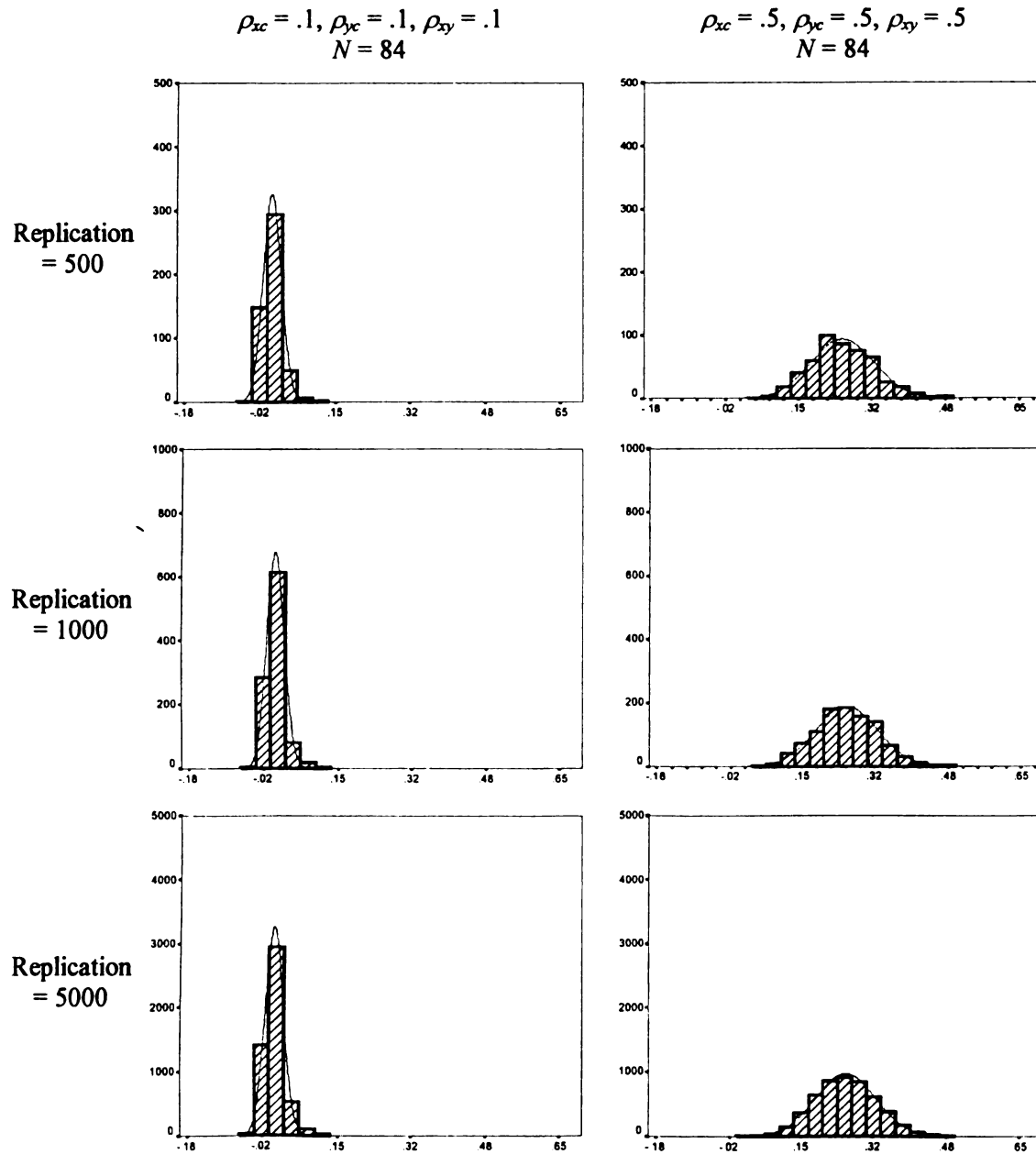


Figure 2 (cont'd). The distributions of the product of simulated correlations  $\tilde{r}_{xc}\tilde{r}_{yc}$  with different numbers of replications for the selected cells.

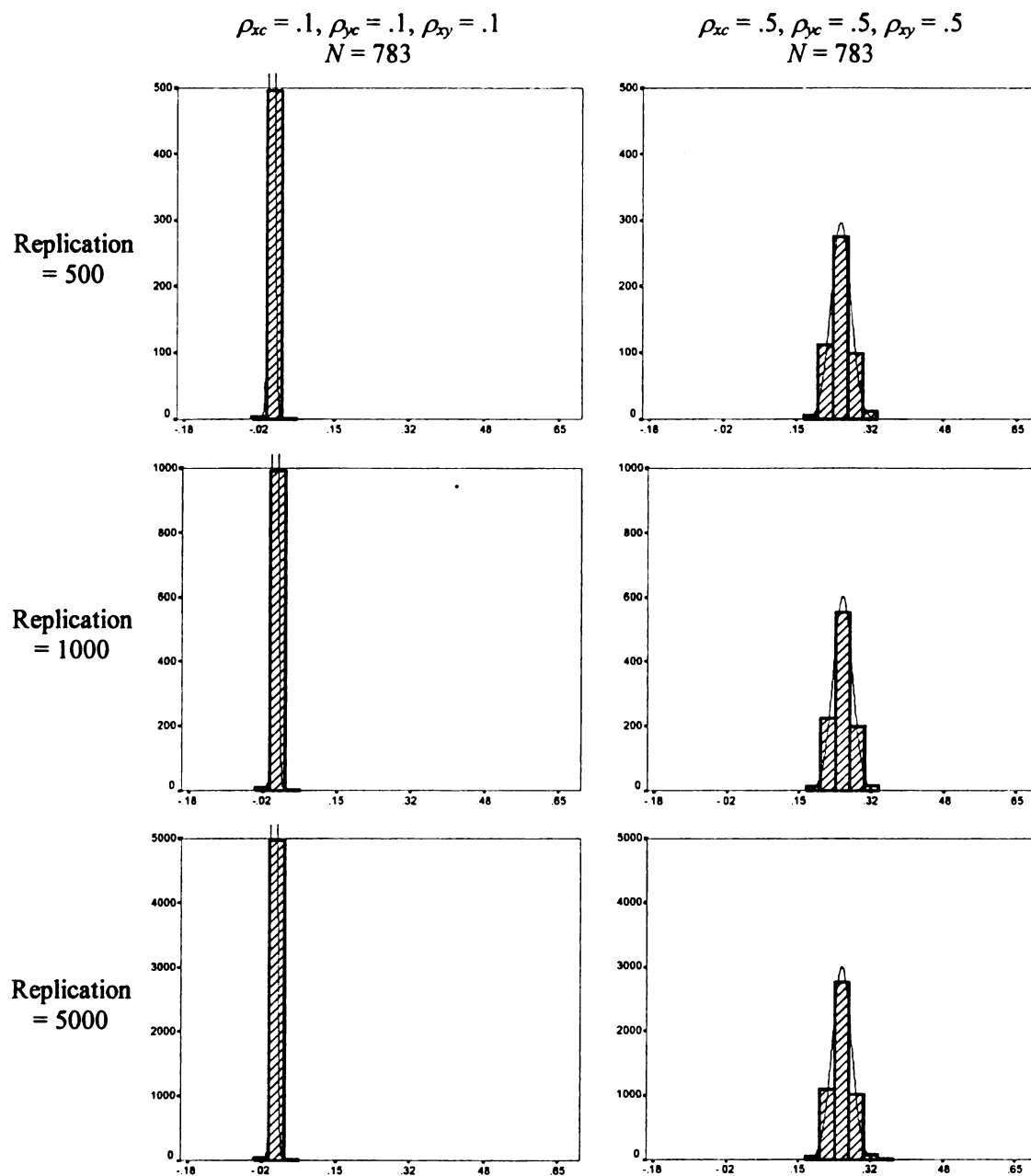
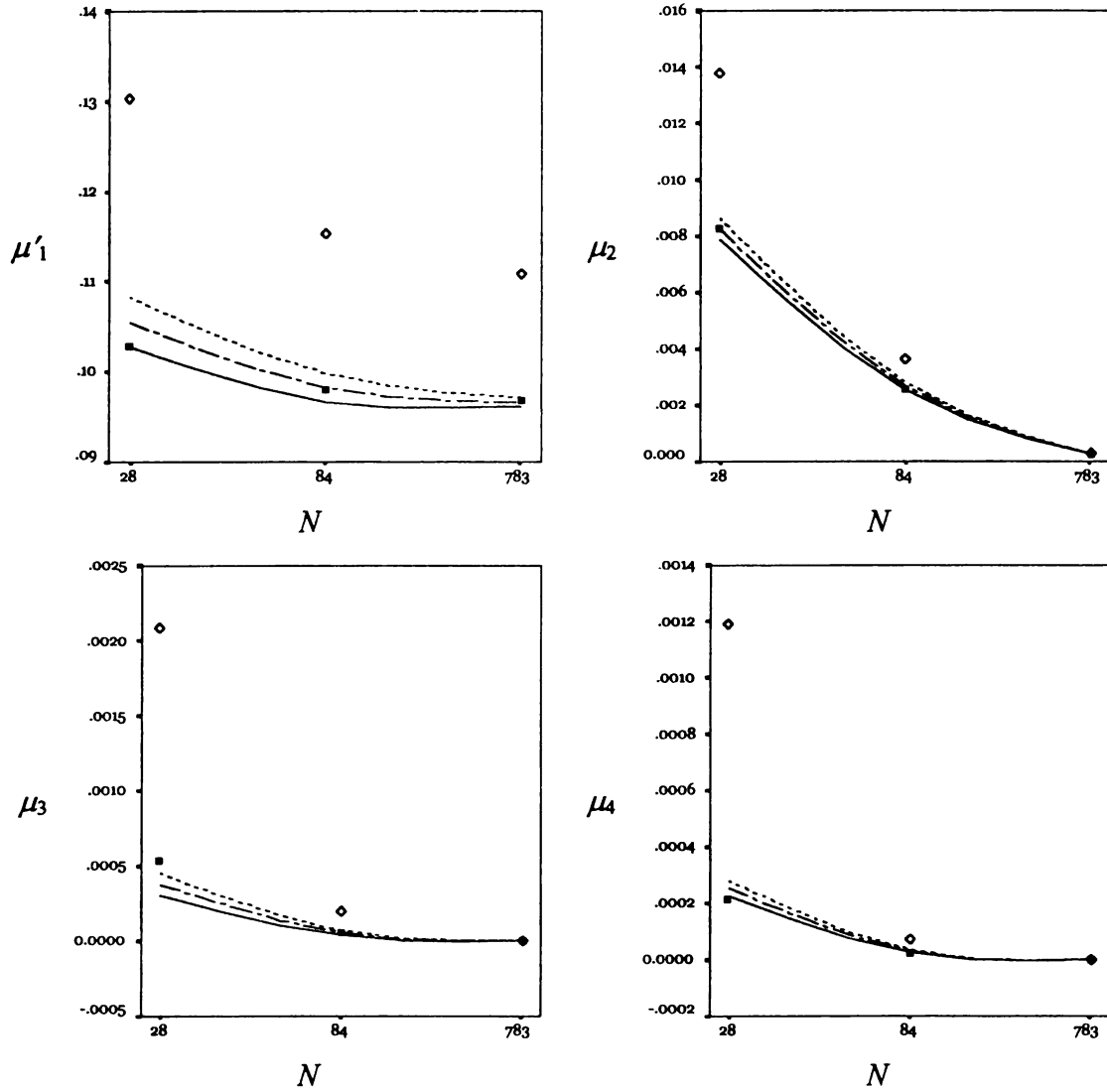


Figure 2 (cont'd). The distributions of the product of simulated correlations  $\tilde{r}_{xc} \tilde{r}_{yc}$  with different numbers of replications for the selected cells.



*Figure 3.* The trends of the accuracy of the approximated and Frank's moments along one parameter holding the other parameters constant. ■ = approximated value; ◇ = Frank's value; - - - - simulated value; - - - - one-standard-error upper bound; — one-standard-error lower bound.

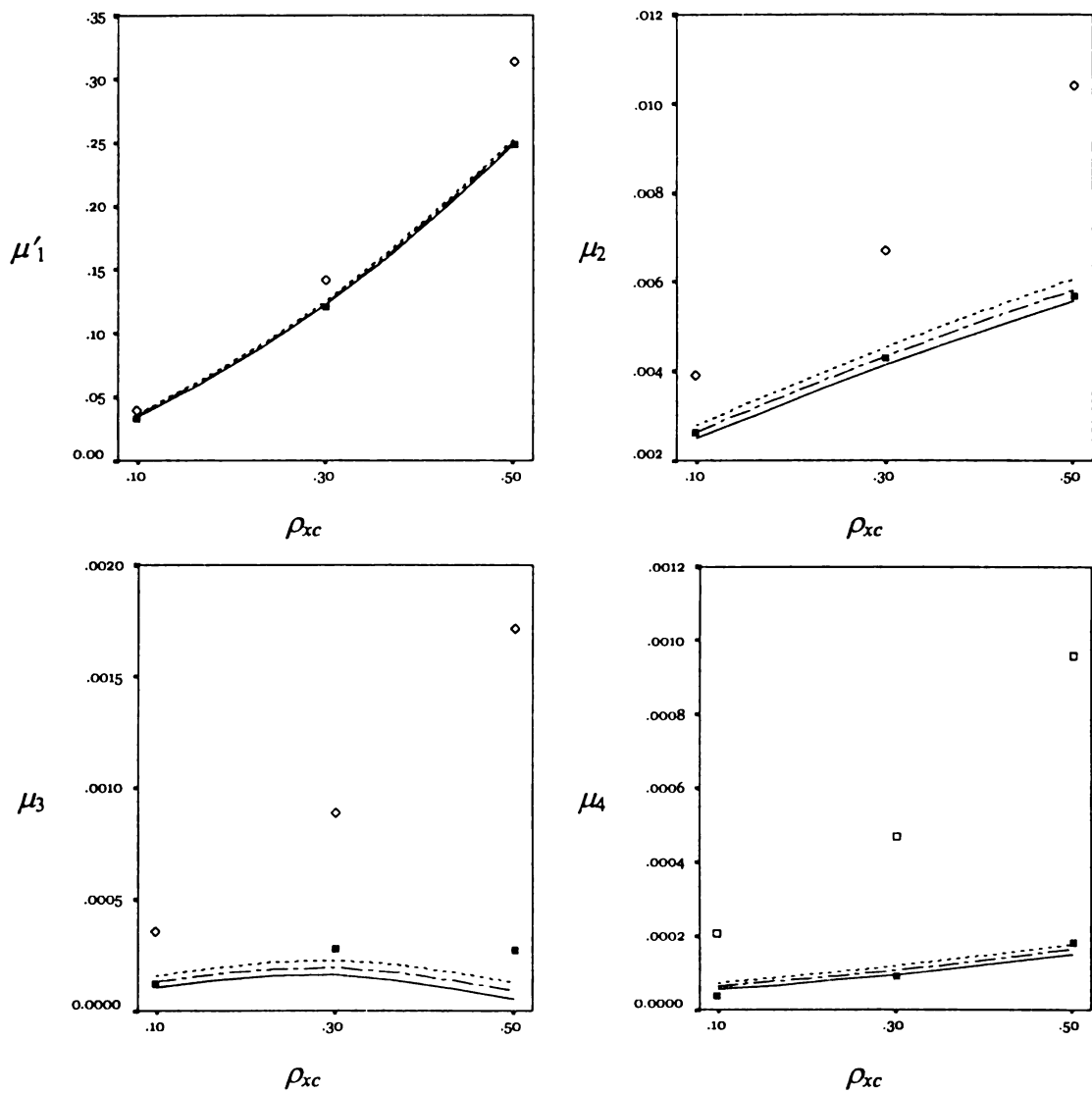


Figure 3 (cont'd). The trends of the accuracy of the approximated and Frank's moments along one parameter holding the other parameters constant, including the sample size. ■ = approximated value; ◇ = Frank's value; --- simulated value; ----- one-standard-error upper bound; \_\_\_\_\_ one-standard-error lower bound.

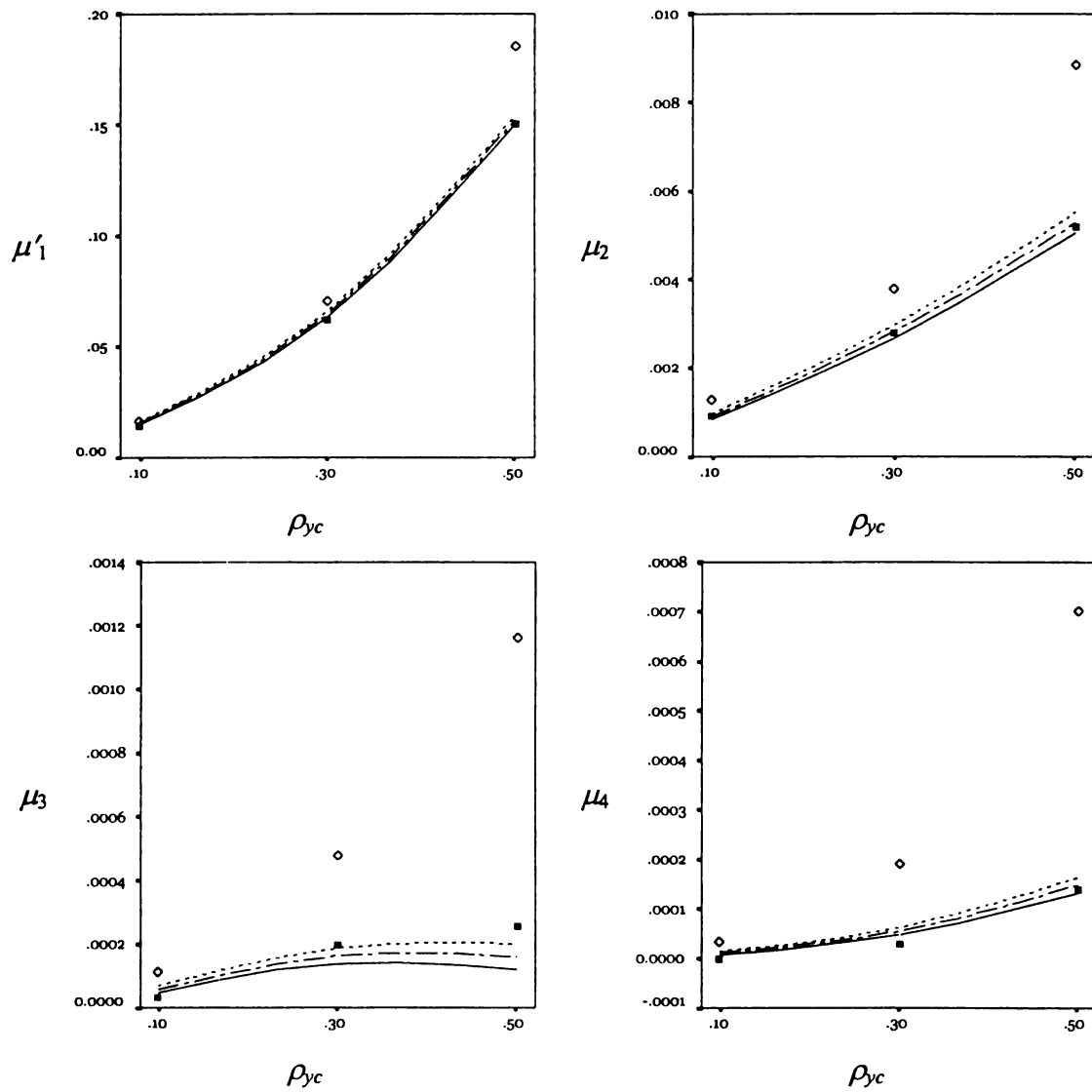


Figure 3 (cont'd). The trends of the accuracy of the approximated and Frank's moments along one parameter holding the other parameters constant, including the sample size. ■ = approximated value; ◇ = Frank's value; ----- simulated value; ----- one-standard-error upper bound; ----- one-standard-error lower bound.

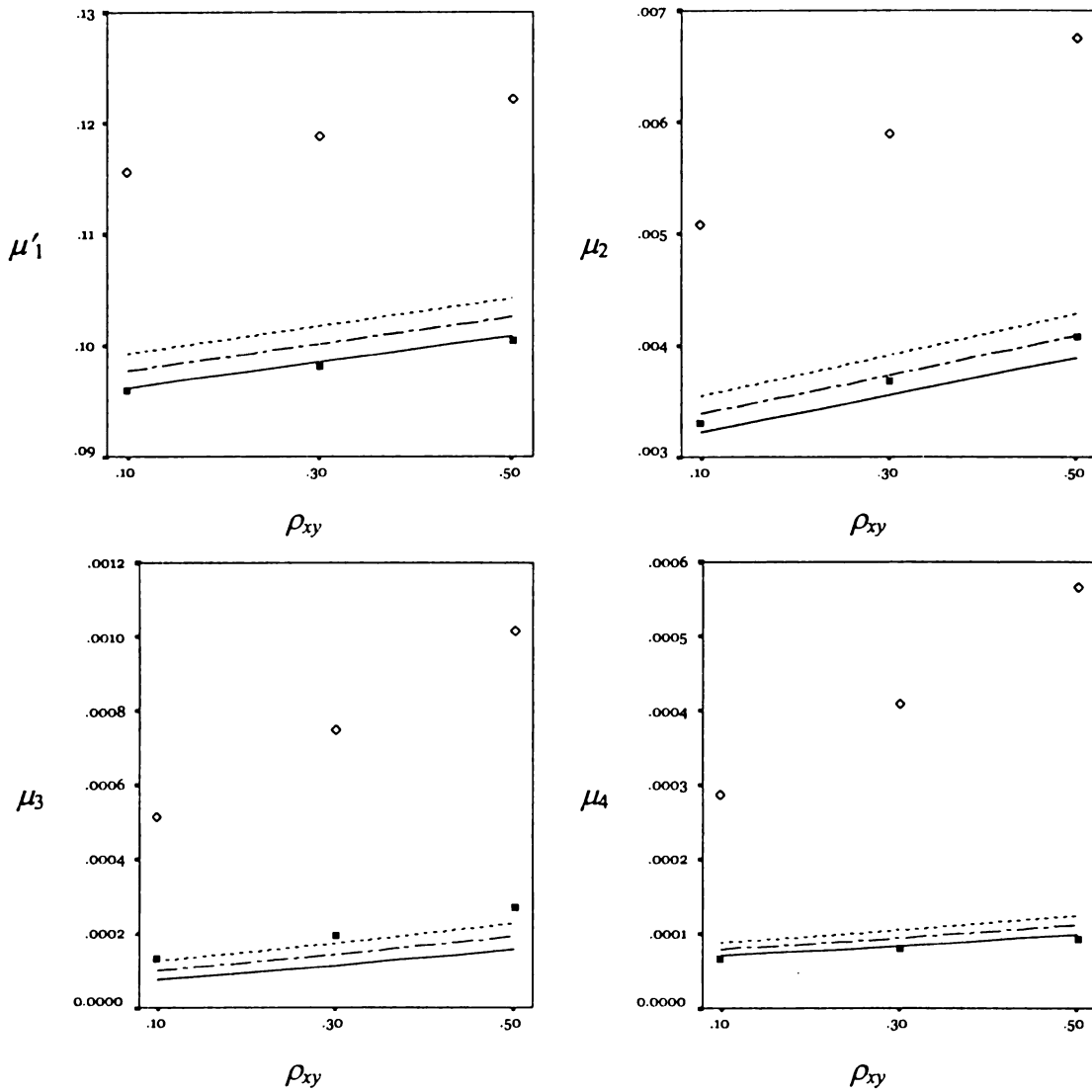


Figure 3 (cont'd). The trends of the accuracy of the approximated and Frank's moments along one parameter holding the other parameters constant, including the sample size. ■ = approximated value; ◇ = Frank's value; --- = simulated value; ..... = one-standard-error upper bound; — = one-standard-error lower bound.

Now we examine the trends of the accuracy of the approximated and Frank's values.

Figure 3 shows the trends of the accuracy of the approximated and Frank's values of the four moments along one parameter,  $N$ ,  $\rho_{xc}$ ,  $\rho_{yc}$ , or  $\rho_{xy}$ , holding the other parameters constant. As the sample size  $N$  becomes larger, all of the approximated values and most of Frank's values become more accurate, except for Frank's mean. This is not an unexpected result for Frank's mean in that it is obvious that the mean of the Fisher  $z$ ,

$$E[z(r_*)] = \frac{1}{2} \ln \left( \frac{1 + \rho_*}{1 - \rho_*} \right), \text{ does not equal the mean of the original correlation, } E(r_*) = \rho_* -$$

$O\left(\frac{1}{M}\right)$  (cf., Equation 5); and therefore, the abnormal trend regarding  $N$  for Frank's mean is predictable. It also shows the inadequacy of Frank's approach as an approximation to the distribution of the product of the two dependent correlation coefficients.

Frank might argue that he was trying to get a  $p$ -value through Aroian's procedure, not approximate the distribution. However, the P-P plots in Chapter 6 also will show that Frank's quantiles are not comparable to the extremely accurate approximated quantiles obtained by the approximation procedure described in this current study (cf., Figure 6).

As for the trends of the approximated and Frank's values regarding  $\rho_s$ , there is no clear trend in the approximated values, except for the third moment in which the approximated values increasingly overestimate the third moment as  $\rho_s$  become larger. As stated above, I will fix this problem by a regression approach based on simulated values. For Frank's values, they consistently overestimate all of the moments when any of the  $\rho_s$  become large, since the magnitude of the corresponding Fisher  $z$ 's increase much more than do those of the original correlations. It is yet another evidence of the

detachment of the distribution of the product of two Fisher  $z$ 's from that of the two original correlation coefficients.

## Chapter 5

### APPROXIMATION CORRECTION

As we can see in Tables 3 through 14 and Figure 3, the approximations in this study to the first four moments are very accurate, acknowledging a little discrepancy for the third and the fourth moments when  $N$  is small. To account for this discrepancy, we can borrow the principle of regression analysis in which we use predictors to explain as much variation in an outcome as possible through a linear model. In this light, we can model the deviations of the simulated values from the approximated values of the moments as a function of  $\rho_{xc}$ ,  $\rho_{yc}$ ,  $\rho_{xy}$ ,  $\frac{1}{N^2}$ , and up to 4-way interactions. By the nature of the formulae in Equations (12), (12'), and (12''), a regression model without an intercept is employed. To keep the work for the approximation correction minimum, we may need to model the discrepancy for only the third and the fourth moments when  $N = 28$  and  $84$ . For the other cases, we do not need to model the deviations, since the approximations are accurate enough and the deviations do not have much variation to be explained in the regression model. For instance, for the very accurate approximation to the variance  $\mu_2$ , only 41% of the outcome variance could be explained by all predictors in the model, whereas as shown below for the third moment  $\mu_3$ , more than 99% of the outcome variance can be explained by the predictors.

Since we only have 54 cells of records, a stepwise regression is used to obtain more statistical power; and the stepwise regression results in a simpler model for the third moment when  $N = 28$  and  $84$ . In particular, the model is

$$D = S - A = \beta_1\left(\frac{\rho_{xc}}{N^2}\right) + \beta_2\left(\frac{\rho_{yc}}{N^2}\right) + \beta_3\left(\frac{\rho_{xy}}{N^2}\right) + \beta_4\left(\frac{\rho_{xc} \times \rho_{yc}}{N^2}\right) + \beta_5\left(\frac{\rho_{xc} \times \rho_{xy}}{N^2}\right) + \beta_6\left(\frac{\rho_{yc} \times \rho_{xy}}{N^2}\right) + \varepsilon, \quad (16)$$

where  $D$  is the deviation between the simulated value and the approximated value;

$S$  is the simulated value;

$A$  is the approximated value;

$\beta_i$  is the regression coefficient;

$\varepsilon$  is the error.

We now can use the predicted value of the deviation,  $\hat{D}$ , to correct the inaccuracy of the approximation to the third moment. The corrected or estimated value of the approximation, denoted as  $\tilde{A} = A + \hat{D}$ , is theoretically assumed to be more accurate than the approximated value,  $A$ . The virtue of this methodology is to utilize all the available information borne in the simulation data across  $\rho_{xc}$ ,  $\rho_{yc}$ ,  $\rho_{xy}$ , and  $N$  to correct the inaccuracy of the approximation. This method is consistent with the general principle of statistical analysis; and the method works because the range of parameters in the experimental design is limited by theory, constraining  $\rho$ s according to what are defined as small, medium, and large effects.

After model (16) was fitted<sup>[9]</sup> to the data in Tables 9 and 10, I obtained the predictive equation as follows ( $R^2 = .991$ ):

$$\begin{aligned}\hat{D}_{\mu_3} = & -.176408\left(\frac{\rho_{xc}}{N^2}\right) + .116374\left(\frac{\rho_{yc}}{N^2}\right) + .447522\left(\frac{\rho_{xy}}{N^2}\right) - .550128\left(\frac{\rho_{xc} \times \rho_{yc}}{N^2}\right) \\ & - 1.659398\left(\frac{\rho_{xc} \times \rho_{xy}}{N^2}\right) - .805672\left(\frac{\rho_{yc} \times \rho_{xy}}{N^2}\right).\end{aligned}\quad (17)$$

And, the estimated values,  $\tilde{A}$ , for the third moments, along with the simulated values,  $S$ , as well as the approximated values,  $A$ , are listed in Tables 15 and 16 that show more accurate results for the third moment than those in Tables 9 and 10.

For the fourth moments when  $N = 28$  and  $84$ , a stepwise regression ended up with the following model:

$$\begin{aligned}D = S - A = & \beta_1\left(\frac{\rho_{xc}}{N^2}\right) + \beta_2\left(\frac{\rho_{yc}}{N^2}\right) + \beta_3\left(\frac{\rho_{xy}}{N^2}\right) + \beta_4\left(\frac{\rho_{xc} \times \rho_{yc}}{N^2}\right) + \beta_5\left(\frac{\rho_{xc} \times \rho_{xy}}{N^2}\right) \\ & + \beta_6\left(\frac{\rho_{yc} \times \rho_{xy}}{N^2}\right) + \beta_7\left(\frac{\rho_{xc} \times \rho_{yc} \times \rho_{xy}}{N^2}\right) + \varepsilon.\end{aligned}\quad (18)$$

After model (18) was fitted (cf., Footnote 9) to the data in Tables 12 and 13, I obtained the predictive equation as follows ( $R^2 = .981$ ):

$$\begin{aligned}\hat{D}_{\mu_4} = & .097559\left(\frac{\rho_{xc}}{N^2}\right) + .113876\left(\frac{\rho_{yc}}{N^2}\right) - .084159\left(\frac{\rho_{xy}}{N^2}\right) - .425381\left(\frac{\rho_{xc} \times \rho_{yc}}{N^2}\right) \\ & + 1.090953\left(\frac{\rho_{xc} \times \rho_{xy}}{N^2}\right) + .564063\left(\frac{\rho_{yc} \times \rho_{xy}}{N^2}\right) - 3.620234\left(\frac{\rho_{xc} \times \rho_{yc} \times \rho_{xy}}{N^2}\right).\end{aligned}\quad (19)$$

---

<sup>[9]</sup> Here a regular (unweighted) regression was conducted. A WLS regression weighted by the sample sizes was also tried, but the result was not evidently better than that of the unweighted regression.

The estimated values,  $\tilde{A}$ , for the fourth moments, along with the simulated values,  $S$ , as well as the approximated values,  $A$ , are listed in Tables 17 and 18 that show more accurate results for the fourth moment than those in Tables 12 and 13.

To determine whether the correction is sample dependent, a cross validation is analyzed by fitting the models (16) and (18) to the full sample in Tables 9 through 14. Extremely good approximations were obtained. Specifically, compared with ones in Equations (17) and (19), only two coefficients have differences in  $10^{-6}$ .

Table 15

*Estimated Values of the Third Moment Compared with the Simulated Values and the Approximated Values (N = 28)*

$\rho_{xc}$	$\rho_{xy}$	$S$ (s.e.)	$A$ (Std. Diff.)	$\tilde{A}$ (Std. Diff.)
.10	.10	.000102612 ( .000022926 )	.000081765 ( -.909316933 )	.000092543 ( -.439195673 )
	.30	.000158858 ( .000030373 )	.000086309 ( -2.388601719 )	.000148366 ( -.345438383 )
	.50	.000213329 ( .000037334 )	.000088324 ( -3.348288423 )	.000201660 ( -.312556919 )
.30	.10	.000229896 ( .000040731 )	.000240334 ( .256266726 )	.000245837 ( .391372665 )
	.30	.000342621 ( .000049056 )	.000347121 ( .091731898 )	.000362797 ( .411285062 )
	.50	.000460741 ( .000058132 )	.000460429 ( -.005367096 )	.000486279 ( .439310535 )
.50	.10	.000312498 ( .000091352 )	.000289363 ( -.253251160 )	.000289591 ( -.250755320 )
	.30	.000528974 ( .000102736 )	.000538633 ( .094017676 )	.000507929 ( -.204845429 )
	.50	.000761741 ( .000122471 )	.000801008 ( .320622841 )	.000739371 ( -.182655486 )
.30	.10	.000400153 ( .000055523 )	.000532525 ( 2.384093079 )	.000408592 ( .151991067 )
	.30	.000514982 ( .000066161 )	.000707575 ( 2.910974743 )	.000509153 ( -.088103263 )
	.50	.000637833 ( .000078976 )	.000911491 ( 3.465077998 )	.000638579 ( .009445908 )
.50	.10	.000332125 ( .000086245 )	.000526148 ( 2.249672445 )	.000368873 ( .426088469 )
	.30	.000484532 ( .000095090 )	.000778373 ( 3.090135661 )	.000505502 ( .220527921 )
	.50	.000656781 ( .000107726 )	.001073164 ( 3.865204315 )	.000684698 ( .259148209 )
.50	.10	.000186486 ( .000080016 )	.000490178 ( 3.795390922 )	.000175400 ( -.138547291 )
	.30	.000223523 ( .000089843 )	.000713757 ( 5.456563116 )	.000198721 ( -.276059348 )
	.50	.000281911 ( .000100228 )	.000994611 ( 7.110787405 )	.000279316 ( -.025890969 )

Note.  $\tilde{A} = A + \hat{D}$  = the estimated values;  $\hat{D}$  = the predicted deviations between  $S$  and  $A$ ;  $S$  = the simulated values;  $A$  = the approximated values.

Table 16

*Estimated Values of the Third Moment Compared with the Simulated Values and the Approximated Values (N = 84)*

$\rho_{xc}$	$\rho_{xy}$	$S$ (s.e.)	$A$ (Std. Diff.)	$\tilde{A}$ (Std. Diff.)
.10	.10	.000010500 ( .000002313 )	.000009451 ( -.453523562 )	.000012335 ( .793341980 )
	.30	.000015611 ( .000003300 )	.000012091 ( -1.066666667 )	.000020673 ( 1.533939394 )
	.50	.000020994 ( .000004420 )	.000014954 ( -1.366515837 )	.000029233 ( 1.864027149 )
.30	.10	.000032053 ( .000008501 )	.000026099 ( -.700388190 )	.000031771 ( -.033172568 )
	.30	.000047655 ( .000011095 )	.000040805 ( -.517395223 )	.000047607 ( -.004326273 )
	.50	.000064406 ( .000014054 )	.000056834 ( -.538778995 )	.000064766 ( .025615483 )
.50	.10	.000051326 ( .000020603 )	.000030707 ( -1.000776586 )	.000039167 ( -.590156773 )
	.30	.000082326 ( .000024813 )	.000062103 ( -.815016322 )	.000067126 ( -.612582114 )
	.50	.000116846 ( .000029601 )	.000095498 ( -.721191852 )	.000097083 ( -.667646363 )
.30	.10	.000052947 ( .000011332 )	.000055051 ( .185668902 )	.000046342 ( -.582862690 )
	.30	.000072472 ( .000014632 )	.000077684 ( .356205577 )	.000060697 ( -.804743029 )
	.50	.000093117 ( .000017853 )	.000104548 ( .640284546 )	.000079286 ( -.774715734 )
.50	.10	.000054061 ( .000019433 )	.000051763 ( -.118252457 )	.000042722 ( -.583491998 )
	.30	.000075321 ( .000024430 )	.000083043 ( .316086779 )	.000061159 ( -.579697094 )
	.50	.000098774 ( .000029508 )	.000120181 ( .725464281 )	.000085452 ( -.451470788 )
.50	.10	.000027162 ( .000017725 )	.000045863 ( 1.055063470 )	.000019322 ( -.442313117 )
	.30	.000031018 ( .000022389 )	.000072871 ( 1.869355487 )	.000024079 ( -.309928983 )
	.50	.000041201 ( .000027059 )	.000107492 ( 2.449868805 )	.000036449 ( -.175616246 )

Note.  $\tilde{A} = A + \hat{D}$  = the estimated values;  $\hat{D}$  = the predicted deviations between  $S$  and  $A$ ;  $S$  = the simulated values;  $A$  = the approximated values.

Table 17

*Estimated Values of the Fourth Moment Compared with the Simulated Values and the Approximated Values (N = 28)*

$\rho_{xc}$	$\rho_{xy}$	$S$ (s.e.)	$A$ (Std. Diff.)	$\tilde{A}$ (Std. Diff.)
.10	.10	.000025395 ( .000005791 )	.000000932 ( -4.224313590 )	.000028233 ( .490070800 )
	.30	.000033559 ( .000008163 )	-.000001358 ( -4.277471518 )	.000037458 ( .477643023 )
	.50	.000043744 ( .000010421 )	-.000004857 ( -4.663755878 )	.000045474 ( .166010939 )
.30	.10	.000091691 ( .000012082 )	.000038147 ( -4.431716603 )	.000088800 ( -.239281576 )
	.30	.000109363 ( .000015133 )	.000036404 ( -4.821185489 )	.000108880 ( -.031917003 )
	.50	.000132794 ( .000018452 )	.000028269 ( -5.664697594 )	.000122569 ( -.554140473 )
.50	.10	.000332066 ( .000036967 )	.000252585 ( -2.150052750 )	.000326591 ( -.148105067 )
	.30	.000366214 ( .000042815 )	.000267453 ( -2.306691580 )	.000373590 ( .172276071 )
	.50	.000417883 ( .000053603 )	.000271646 ( -2.728149544 )	.000409915 ( -.148648397 )
.30	.10	.000152673 ( .000018407 )	.000105481 ( -2.563807247 )	.000148592 ( -.221709132 )
	.30	.000188228 ( .000022664 )	.000125898 ( -2.750176491 )	.000191081 ( .125882457 )
	.50	.000229483 ( .000028618 )	.000141015 ( -3.091341114 )	.000228270 ( -.042385911 )
.50	.10	.000325797 ( .000034324 )	.000300419 ( -.739366041 )	.000326709 ( .026570330 )
	.30	.000378532 ( .000038363 )	.000369983 ( -.222844929 )	.000391712 ( .343560201 )
	.50	.000444161 ( .000044656 )	.000440353 ( -.085274095 )	.000457521 ( .299175923 )
.50	.10	.000349379 ( .000031879 )	.000374644 ( .792527996 )	.000353217 ( .120392735 )
	.30	.000432846 ( .000036435 )	.000486989 ( 1.486016193 )	.000424310 ( -.234280225 )
	.50	.000517932 ( .000041225 )	.000619684 ( 2.468211037 )	.000515752 ( -.052880534 )

Note.  $\tilde{A} = A + \hat{D}$  = the estimated values;  $\hat{D}$  = the predicted deviations between  $S$  and  $A$ ;  $S$  = the simulated values;  $A$  = the approximated values.

Table 18

*Estimated Values of the Fourth Moment Compared with the Simulated Values and the Approximated Values (N = 84)*

$\rho_x$	$\rho_y$	$\rho_{xy}$	$S$ (s.e.)	$A$ (Std. Diff.)	$\tilde{A}$ (Std. Diff.)
.10	.10	.10	.000001057 ( .000000269 )	.000000160 ( -3.334572491 )	.000001597 ( 2.005576208 )
		.30	.000001517 ( .000000437 )	.000000094 ( -3.256292906 )	.000002204 ( 1.570938215 )
		.50	.000002080 ( .000000652 )	-.000000048 ( -3.263803681 )	.000002772 ( 1.061349693 )
.30	.30	.10	.000007925 ( .000001627 )	.000004301 ( -2.227412415 )	.000009929 ( 1.231714813 )
		.30	.000010005 ( .000002306 )	.000004813 ( -2.251517780 )	.000012866 ( 1.240676496 )
		.50	.000012543 ( .000003136 )	.000005068 ( -2.383609694 )	.000015546 ( .957589286 )
.50	.50	.10	.000037124 ( .000005280 )	.000027981 ( -1.731628788 )	.000036204 ( -.174242424 )
		.30	.000042258 ( .000006742 )	.000030662 ( -1.719964402 )	.000042455 ( .029219816 )
		.50	.000048534 ( .000008487 )	.000032974 ( -1.833392247 )	.000048337 ( -.023211971 )
.30	.30	.10	.000015310 ( .000002288 )	.000010897 ( -1.928758741 )	.000015687 ( .164772727 )
		.30	.000020430 ( .000003221 )	.000014492 ( -1.843526855 )	.000021735 ( .405153679 )
		.50	.000026042 ( .000004214 )	.000018342 ( -1.827242525 )	.000028037 ( .473421927 )
.50	.50	.10	.000039028 ( .000004820 )	.000030707 ( -1.726348548 )	.000033628 ( -1.120331950 )
		.30	.000049363 ( .000006550 )	.000039828 ( -1.455725191 )	.000042243 ( -1.087022901 )
		.50	.000061049 ( .000008456 )	.000050142 ( -1.289853359 )	.000052050 ( -1.064214759 )
.50	.50	.10	.000043661 ( .000004284 )	.000035330 ( -1.944677871 )	.000039539 ( -0.962231559 )
		.30	.000057562 ( .000005858 )	.000048855 ( -1.486343462 )	.000050268 ( -1.245134858 )
		.50	.000073253 ( .000007544 )	.000065800 ( -.987937434 )	.000065102 ( -1.080408271 )

Note.  $\tilde{A} = A + \hat{D}$  = the estimated values;  $\hat{D}$  = the predicted deviations between  $S$  and  $A$ ;  $S$  = the simulated values;  $A$  = the approximated values.

## Chapter 6

### DISTRIBUTION COMPARISON

Up to now, we have compared the approximate distribution of the product of two dependent correlation coefficients with the simulated distribution in terms of the first four moments. However, we usually use  $p$ -values obtained from the shape of a distribution, rather than the moments, to do analysis. Thus, it is more desirable to compare the shapes of the approximate distribution with the simulated distribution. After obtaining the distribution function for the approximate distribution of the product of two dependent correlation coefficients as a Pearson Type I distribution, the distribution comparison becomes achievable. A P-P plot is employed for this comparison.

A P-P plot, a probability plot, is a graphical tool for assessing the fit of data to a theoretical distribution (Rice, 1995, p. 321). Specifically, for a given sample data  $X_1, \dots, X_n$ , we plot

$$X_{(i)} \quad \text{vs.} \quad F^{-1}\left(\frac{i}{n+1}\right)$$

where  $X_{(i)}$ ,  $i = 1, \dots, n$ , are the order statistics of  $X_1, \dots, X_n$ , and  $F$  is the cumulative distribution function of the theoretical distribution.

Although we do not know the theoretical distribution of the product of two dependent correlation coefficients, we have the approximated Pearson Type I distribution. Then, we

can compare the simulated data with the approximated Pearson Type I distribution to validate the approximation of the distribution of the product of two dependent correlation coefficients to the Pearson Type I distribution. Thus, here in the P-P plot, I will plot the quantiles of the Pearson Type I distribution against those of the simulated distribution of the product of two dependent correlation coefficients. Due to the complexity of the calculations of its mathematical function for obtaining every quantile of the Pearson Type I distribution, I used the common 15 percentiles, .25%, .5%, 1%, 2.5%, 5%, 10%, 25%, 50%, 75%, 90%, 95%, 97.5%, 99%, 99.5%, and 99.75%, which can be looked up in Pearson and Hartley's table (Pearson & Hartley, 1972, Table 32). Setting  $\frac{i}{n+1} = .25\%$ , .5%, 1%, 2.5%, 5%, 10%, 25%, 50%, 75%, 90%, 95%, 97.5%, 99%, 99.5%, or 99.75%, and noting that  $n = 1000$  in this (simulation) study, I obtained the corresponding  $i$ th order statistic  $X_{(i)}$  in the simulated data. Therefore, in our P-P plot there are only 15 points, instead of 1000 points, each corresponding to one of the 15 percentiles above.

Figure 4 displays some P-P plots for selected cells. From the P-P plots, we can see that when the correlations are larger and  $N$  becomes bigger, the simulated data fit the Pearson Type I distribution better, which is consistent with the conclusion in Chapter 3 (Approximation Procedures). I argued there that if the correlations are larger than a small size, say .10, and  $N$  is bigger than 300, the distribution of the product of two correlations can be approximated as the Pearson Type I distribution. Also, Figure 4 shows the unmatched low ends for the cases of smaller correlations, which results from the inaccuracy of the approximations to the fourth moment (cf., Tables 12 and 13) that regulates the tails of the distribution. However, for smaller correlations, these cases are less interesting to us in that the impact of the confounding variable would be so small that we would not need to assess its impact on the causal inference about the predictor of interest.

To address the concern above, I also created P-P plots (see Figure 5) using the estimated values, via the regression approach, for the third and the fourth moments. Figure 5 shows a better fit of the simulated data to the Pearson Type I distribution using the estimated third and fourth moments. The nicer results come from the better

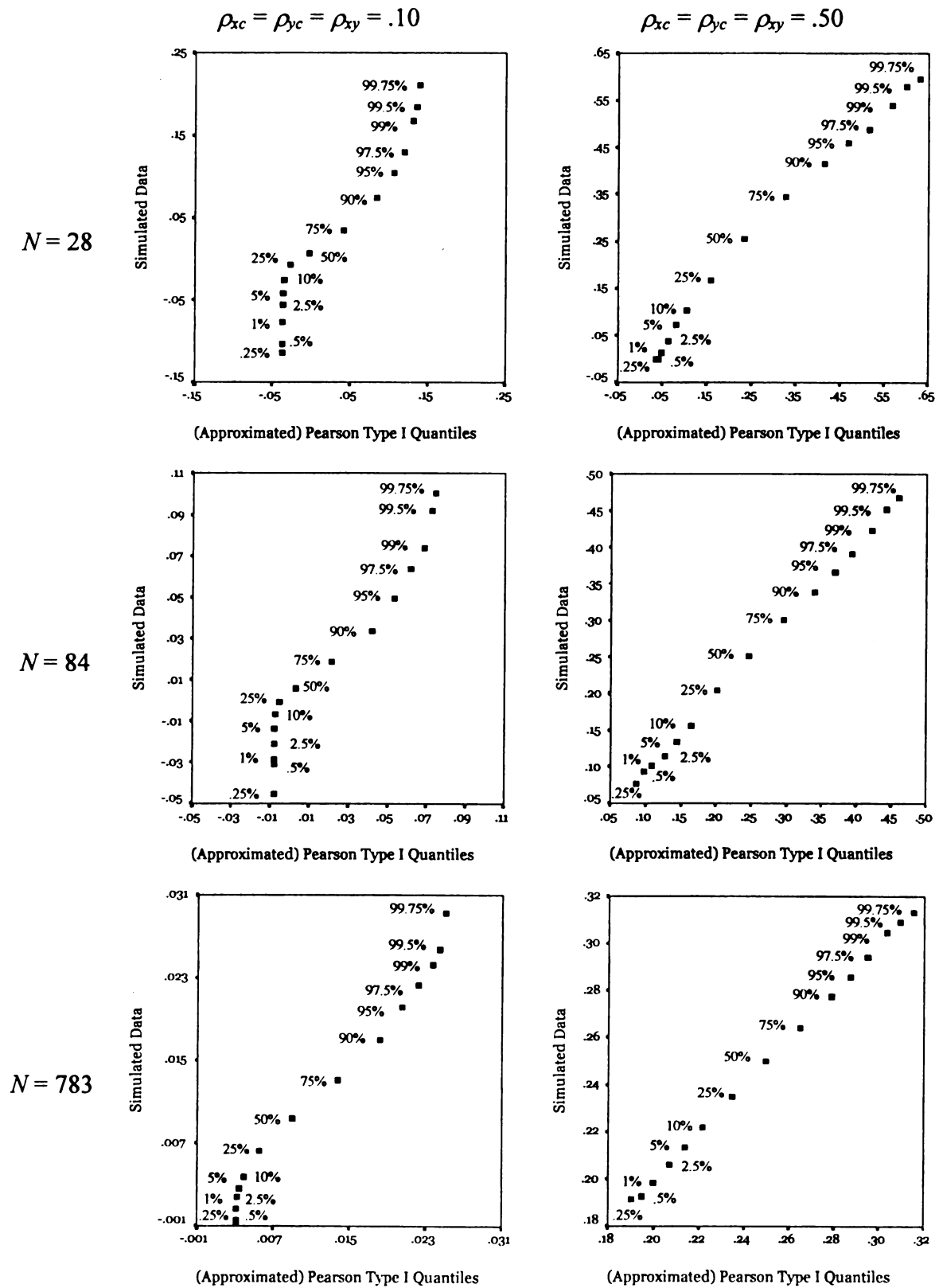
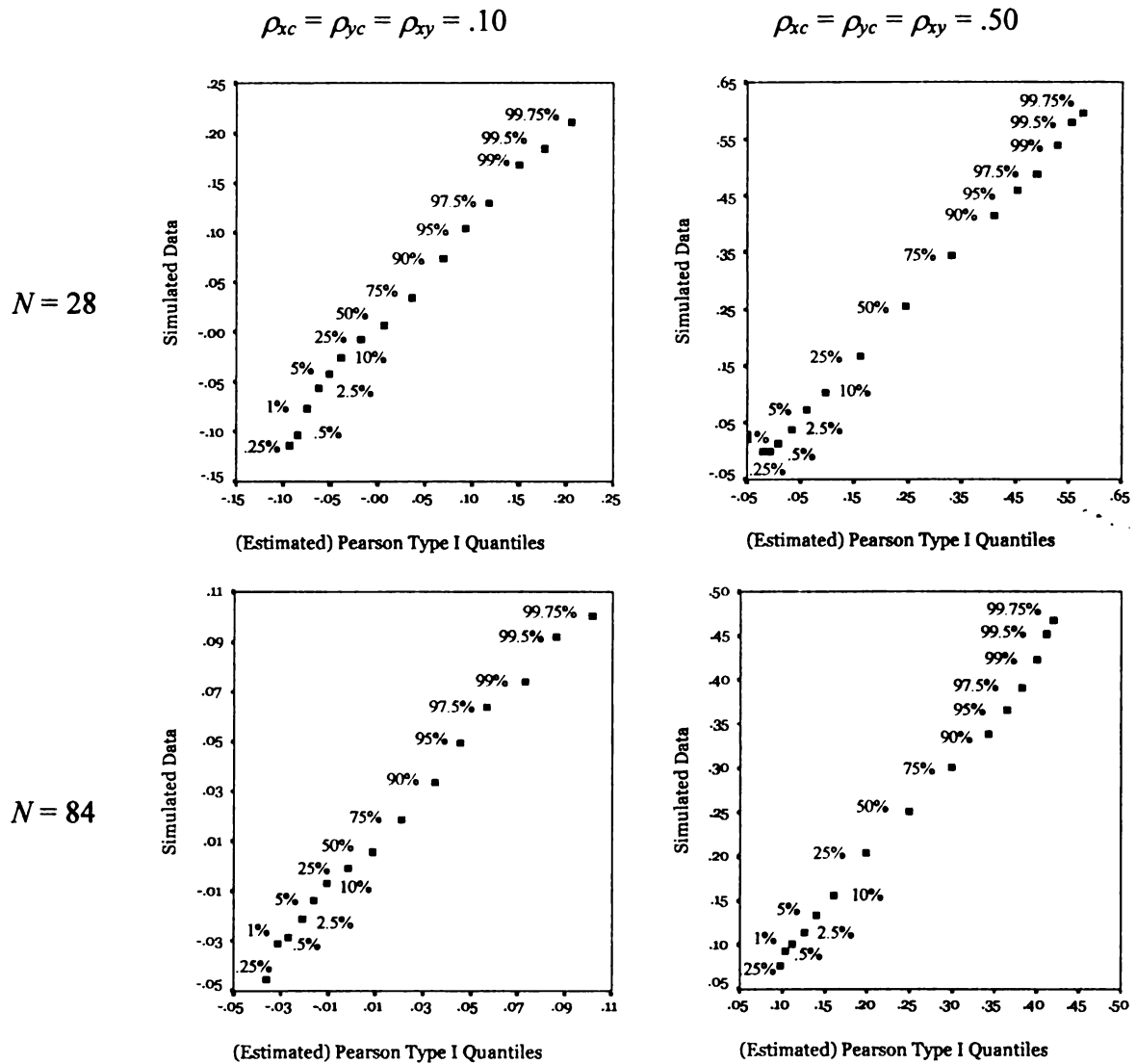


Figure 4. P-P plots for the selected cells.



*Figure 5.* P-P plots for the selected cells using the estimated values for the third and the fourth moments.

approximation of the estimated values of the moments to the simulated ones (relative to that of the approximated values). Thus, the better results validate the regression approach in Chapter 5 as a helpful methodology to correct the inaccuracy of the approximated values.

Figure 5 also shows that the estimated values for the third and the fourth moments did a good correction job, and the P-P plots indicate a better fit of the simulated values to Pearson Type I distribution for small correlations.

To compare the approximate distribution—Pearson Type I distribution—with the Frank's distribution, Figure 6 is a P-P plot including the two distributions. According to Frank's approach, Fisher z's,  $z(r_{xc})$  and  $z(r_{yc})$ , were obtained from the simulated correlations and standardized by the variances. Then, the quantiles of the standardized product  $z(r_{xc}) \times z(r_{yc})$  were looked up in Meeker, Cornwell, and Aroian's tables (1981) (see Frank, 2000, p.174 for the detailed steps). Figure 6 shows that the approximated distribution is fitted much better by the simulated data than is Frank's distribution. In addition, Frank's distribution is too conservative. For example, the quintile of .30 is corresponding to the percentile of 99% for the approximate distribution, but only 50% for Frank's distribution. Therefore, it provides evidence that the approximation procedures using the Pearson Type I distribution are more advanced and accurate than Frank's approach for approximating the distribution of the product of two dependent correlation coefficients.



## Chapter 7

### APPLICATIONS

In this Chapter, through a real example I will illustrate an application in causal inference by using the methodology with the approximate distribution of the product of two dependent correlation coefficients discussed in the current study.

#### 7.1 An Index of the Robustness of a Causal Inference (IOROCI)

Suppose a causal inference about a predictor of interest,  $X$ , on an outcome,  $Y$ , was made through the following regression model

$$Y = \beta_0 + \beta_x X + \varepsilon, \quad (20)$$

and the corresponding hypothesis regarding the coefficient of  $X$ ,  $\beta_x$ , was

$$H_0: \beta_x = 0 \text{ versus } H_1: \beta_x \neq 0.$$

We know that a  $t$ -value under the null hypothesis  $H_0$  is

$$t_0 = \frac{r_{xy}}{\sqrt{\frac{1 - r_{xy}^2}{n - 2}}}.$$

Suppose we also have the second model (21) that includes a potential confounding variable,  $C$ :

$$Y = \beta_0 + \beta_x X + \beta_c C + \varepsilon. \quad (21)$$

Then, the  $t$ -value under the null hypothesis  $H_0$  with respect to the model (21) is

$$t_c = \frac{r_{xy} - r_{xc}r_{yc}}{\sqrt{\frac{1 - r_{xy}^2 - r_{xc}^2 - r_{yc}^2 + 2r_{xy}r_{xc}r_{yc}}{n - 3}}}.$$

We want to know the likelihood that we will retain the primary statistical inference that rejects  $H_0$ , when  $C$  is in the model (21). That is, for a particular study with an observed  $t_0$  which is larger than  $t_\alpha$ , we are interested in the following probability function  $W$ :

$$W = W(t_\alpha | t_0) = P(t_c > t_\alpha | T = t_0), \text{ for } t_0 > t_\alpha^{[10]}, \quad (22)$$

where  $t_\alpha$  is a  $t$ -critical value at level  $\alpha$ .  $W$  is a likelihood that we will reject the null hypothesis  $H_0$ , when the potential confounding variable is in the model (21), for an observed  $t_0$  from model (20) which is larger than  $t_\alpha$ .

In practice, we make use of the all information we have for assessing the robustness of causal inference. In view of the unobserved confounding variable  $C$ , the information about

---

<sup>[10]</sup> (a) Without loss of generality,  $t_0 > 0$  is assumed, because the case of  $t_0 < 0$  is symmetrical. (b) Different  $t_\alpha$ 's might be used for the two models (20) and (21), because the degrees of freedom are different. One is  $n - 2$ , and  $n - 3$  for the other. But, the two  $t_\alpha$ 's are very close, and when  $n$  fairly large, they are almost identical. For simplicity, we used the same symbol for the both  $t_\alpha$ 's.

$r_{xc}$  and  $r_{yc}$  can be acquired from the distribution of observed covariates. We can use information about the distribution of the covariates to estimate  $W$  in the following manner.

Suppose we have  $m$  covariates  $Z_1, \dots, Z_m$ . The models (20) and (21) become as follows, respectively:

$$Y = \beta_0 + \beta_x X + \beta_{z_1} Z_1 + \dots + \beta_{z_m} Z_m + \varepsilon, \quad (23)$$

$$Y = \beta_0 + \beta_x X + \beta_c C + \beta_{z_1} Z_1 + \dots + \beta_{z_m} Z_m + \varepsilon. \quad (24)$$

The impact of the confounding variable is the product of two partial correlations,

$r_{xc \bullet z_1 \dots z_m} \times r_{yc \bullet z_1 \dots z_m}$  (cf., the footnote [1], p. 6). If we use the distribution of the covariates to

acquire the information about confounding variable  $C$ , then the impact of the confounding

variable,  $r_{xc \bullet z_1 \dots z_m} \times r_{yc \bullet z_1 \dots z_m}$ , can be estimated from the impacts of the covariates,

$r_{xz_i \bullet z_1 \dots z_{i-1} z_{i+1} \dots z_m} \times r_{yz_i \bullet z_1 \dots z_{i-1} z_{i+1} \dots z_m}$ ,  $i = 1, \dots, m$ . Consequently, we can estimate the

probability function  $W$  (22) as follows:

$$\hat{W} = \hat{W}(t_\alpha | t_0, \bar{r}_{xz \bullet z}, \bar{r}_{yz \bullet z}) = P(t_c > t_\alpha | T = t_0, \rho_{xc} = \bar{r}_{xz \bullet z}, \rho_{yc} = \bar{r}_{yz \bullet z}), \text{ for } t_0 > t_\alpha, \quad (25)$$

where  $\bar{r}_{xz \bullet z} = \frac{1}{m} \sum_{i=1}^m r_{xz_i \bullet z_1 \dots z_{i-1} z_{i+1} \dots z_m}$  and  $\bar{r}_{yz \bullet z} = \frac{1}{m} \sum_{i=1}^m r_{yz_i \bullet z_1 \dots z_{i-1} z_{i+1} \dots z_m}$ .

Intuitively,  $t_c$  could be any value when  $t_0 > t_\alpha$ , but  $t_c$  and  $t_0$  are related. For example, in the case of confounding ( $\rho_{xc}\rho_{yc}\rho_{xy} > 0$ ),  $|t_c| < t_\alpha$ ; in the case of suppression,  $t_c < -t_\alpha$ ; and in the case of robust,  $t_c > t_\alpha$ . Note that ITCV is not defined when  $t_0 < t_\alpha$ . Following Frank's Figure 1 (2000, p. 156), Figure 7 shows a general relationship between  $t_c$  and  $t_0$ .

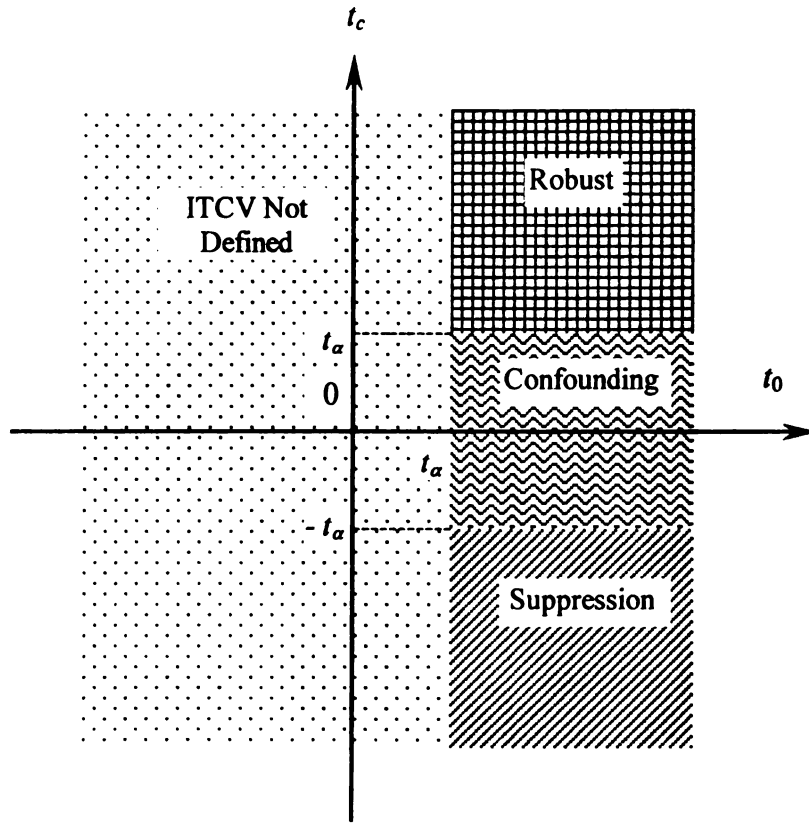


Figure 7. The relationship between  $t_c$  and  $t_0$ .

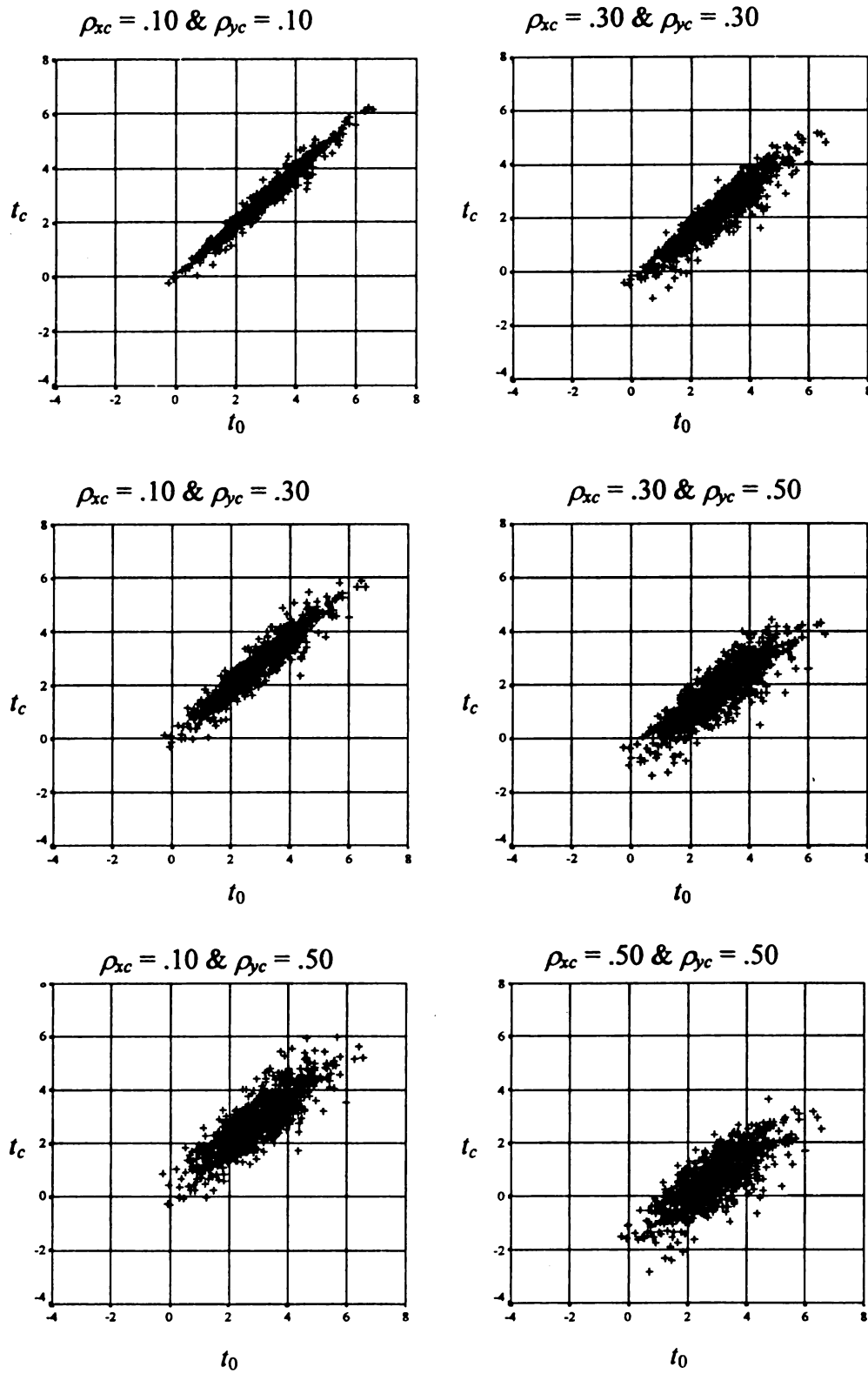


Figure 8. The relationship between  $t_c$  and  $t_0$  for a simulated dataset with  $N = 84$  and  $\rho_{xy} = 30$ .

Table 19  
Detailed Counts and Percentages for  $t_c < -2$ ,  $-2 < t_c < 2$ , and  $2 < t_c$  Based on A  
Simulated Dataset with  $N = 84$  and  $\rho_{xy} = .30$

	$k$	$t_c < -2$	$-2 < t_c < 2$	$2 < t_c$
$\rho_{xc} = .10$ & $\rho_{yc} = .10$	.01			
$-2 < t_0 < 2$			180(95.7%)	8(4.3%)
$2 < t_0 < 4$			14(2.1%)	648(97.9%)
$4 < t_0 < 6$				147(100%)
$6 < t_0$				3(100%)
$\rho_{xc} = .10$ & $\rho_{yc} = .30$	.03			
$-2 < t_0 < 2$			169(89.9%)	19(10.1%)
$2 < t_0 < 4$			56(8.5%)	606(91.5%)
$4 < t_0 < 6$				147(100%)
$6 < t_0$				3(100%)
$\rho_{xc} = .10$ & $\rho_{yc} = .50$	.05			
$-2 < t_0 < 2$			139(73.9%)	49(26.1%)
$2 < t_0 < 4$			73(11%)	589(89%)
$4 < t_0 < 6$			1(.7%)	146(99.3%)
$6 < t_0$				3(100%)
$\rho_{xc} = .30$ & $\rho_{yc} = .30$	.09			
$-2 < t_0 < 2$			188(100%)	
$2 < t_0 < 4$			222(33.5%)	440(66.5%)
$4 < t_0 < 6$			1(.7%)	146(99.3%)
$6 < t_0$				3(100%)
$\rho_{xc} = .30$ & $\rho_{yc} = .50$	.15			
$-2 < t_0 < 2$			188(100%)	
$2 < t_0 < 4$			416(62.8%)	246(37.2%)
$4 < t_0 < 6$			13(8.8%)	134(91.2%)
$6 < t_0$				3(100%)
$\rho_{xc} = .50$ & $\rho_{yc} = .50$	.25			
$-2 < t_0 < 2$		4(2.1%)	184(97.9%)	
$2 < t_0 < 4$			638(96.4%)	24(3.6%)
$4 < t_0 < 6$			93(63.9%)	54(36.7%)
$6 < t_0$				3(100%)

Figure 8 displays a sample relationship between  $t_c$  and  $t_0$  for the 1000 simulated data sets with  $N = 84$  and  $\rho_{xy} = .30$ . We can see in Figure 8 that  $t_c$  and  $t_0$  are positively related. When  $t_0$  becomes large, most of  $t_c$ 's also become large. We also notice that there are some  $t_c$ 's scattered around, but when  $t_0 > 2$  (a  $t_\alpha$  for  $N = 84$  and  $\alpha = .05$ ) they are never

below -2, which is the suppression case (cf., Figure 7). In addition, Figure 8 shows that more points fall into the robust region as  $\rho_{xc}$  or  $\rho_{yc}$  become smaller and that more points fall into the confounding region as  $\rho_{xc}$  and  $\rho_{yc}$  become larger. This information is also illustrated in Table 19 with the detailed counts and percentages for  $t_c < -2$ ,  $-2 < t_c < 2$ , and  $2 < t_c$ . This emphasizes the importance of using  $\bar{r}_{xc}$  and  $\bar{r}_{yc}$  which are employed in (25) through their estimates as the means of the partial correlations with respect to the observed covariates. When the impacts of covariates are generally large, the primary inferences are less robust because we anticipate that there would be a hypothetical confounding variable comparable to the covariates in terms of the (large) impact on the primary inferences for the predictor of interest. This is based on the assumption that the impact of the unmeasured covariates would behave as do the impacts of the measured covariates.

One way to obtain the value of  $\hat{W}$  in (25) is through the impact threshold of the confounding variable, ITCV. When  $t_c > t_\alpha$ , the coefficient of  $X$ ,  $\beta_x$ , is significant in the model (20) or (23), which means that the impact of the confounding variable does not exceed the impact threshold of the confounding variable, ITCV. Therefore, the value of  $\hat{W}$  can be obtained as follows in terms of ITCV:

$$\hat{W} = P(K < \text{ITCV}) = \int_{-\infty}^{\text{ITCV}} f(k) dk, \text{ for } t_0 > t_\alpha, \quad (26)$$

where  $K$  is the impact of the confounding variable with the reference distribution  $f(k)$ .

Note that the condition  $T = t_0$  in (25) is implied by ITCV because ITCV is a function of  $t_0$

(cf., Equation 18 in Frank, 2000, p. 167), and the reference distribution captures  $\rho_{xc} =$

$$\bar{r}_{xz \bullet z} \text{ and } \rho_{yc} = \bar{r}_{yz \bullet z}.$$

Of course, we cannot always measure confounding variables, but under the assumption that the impacts of existing covariates represent the impact of the confounding variable, we can use the impacts of existing covariates to represent the impact of the confounding variable through estimating the means for  $\rho_{xc} = \bar{r}_{xz \bullet z}$  and  $\rho_{yc} = \bar{r}_{yz \bullet z}$ . It is for this reference distribution that I have obtained the approximated Pearson distribution in the preceding chapters. Thus, we still can obtain the value of  $\hat{W}$  as described in (26) by using the reference distribution generated from the measured covariates for the impact of the unmeasured confounding variable.

For the impacts of the measured covariates to have a tractable distribution that is representative of the impact of the unmeasured confounding variable, we assume homogenous impacts of the covariates. When the empirical distribution of the impacts is heterogeneous, researchers need to evaluate the sources of impacts according to substantive theory. Concerns about the heterogeneity of the impacts of the covariates would be greater if the impacts of the covariates were obtained from multiple estimated models accounting for different sets of covariates. The heterogeneity could be assessed by generating a P-P plot of the observed impacts against the theoretical distribution.

In addition, one may be also concerned about the influence of small values of population correlations on obtaining the value of  $\hat{W}$  through (26). The fact that my approximation to the distribution of the product of two dependent correlations is

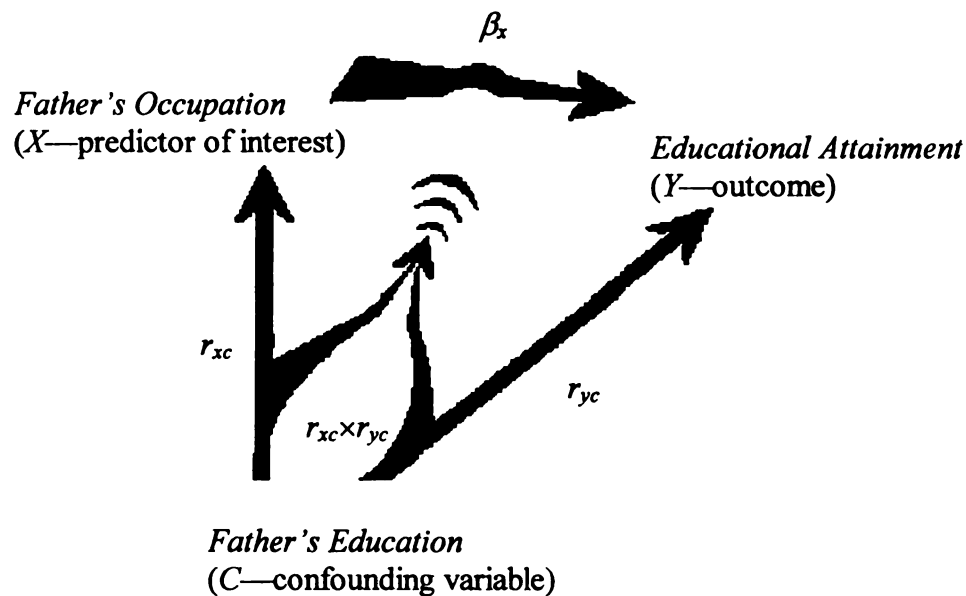
comparatively poor for small values of  $\rho_{xc}$  and  $\rho_{yc}$  becomes more of a concern if the partial correlations of observed covariates with the outcome and with the predictor of interest are smaller than their zero-order correlations, which usually happens in the social sciences because the covariates are often correlated with one another. On the other hand, in the case of confounding, given the generally negative relationship between the  $t$ -value for the predictor of interest and the product of the correlations with respect to the covariates, when the impacts of the covariates are small, the inferences about the predictor of interest through the  $t$ -test are more likely to be robust to the small impacts of the covariates. Thus, we are more likely to retain the primary inference due to the small impacts of covariates, although we may have some difficulty in characterizing the distribution for the small, partialled impacts of covariates. In other words, the poor approximation for small correlations would only result in more conservative decision.

The following is a guideline for interpreting the value of  $\hat{W}$ . If  $\hat{W} > .95$ , this means that the probability of sustaining the original inference is large and we can say that the statistical inference is very robust with respect to concerns about confounding variables. If  $.8 < \hat{W} \leq .95$ , the statistical inference is fairly robust, but we may still need to check some possible confounding variables, and we should interpret the causal inference regarding  $X$  with caution. If  $\hat{W} \leq .8$ , we may want to say that the statistical inference is not robust and we need to consider the possibility of a confounding variable, that is, we cannot make a causal inference regarding  $X$  from the linear model. Thus,  $\hat{W}$  serves as an index of the robustness of a causal inference (IOROCI) to a confounding variable. Please note that here .95 and .8 for IOROCI are arbitrary, as is .05 for the significance level or .2, .5, and .8 for small, medium, and large effect sizes. Researchers can make their own

judgments based on what is studied. The example below will show how exactly to implement this technique.

## 7.2 An Example

In the following text, I will use an example pertaining to educational attainment to illustrate the applications of the methodology delineated in this study.



*Figure 9. Father's Education as a potential confounding variable for the causal relationship between Father's Occupation and Educational Attainment.*

From a general linear model, Featherman and Hauser (1976) concluded that family background, e.g., *Father's Occupation*, has an effect on *Educational Attainment*, but Sobel (1998) argued that both family background and *Educational Attainment* are affected by *Father's Education* which was not controlled for in the analysis. That is,

*Father's Education* is a potential confounding variable for the causal relationship between *Father's Occupation* and *Educational Attainment* (cf., Figure 9). Now, I use the methodology described in this study to assess how robust Featherman and Hauser's causal inference about *Father's Occupation* is to the impact of *Father's Education*.

In order to compare my results with Frank's results (Frank, 2000), I will use the same data set extracted from Featherman and Hauser (1976) and Duncan, Featherman, and Duncan (1972). Let  $X$  be *Father's Occupation*,  $Y$  be *Educational Attainment*, and  $C$  be any covariate (we have 14 covariates and  $N = 10,567$  in the data). First, we need to obtain the reference distribution from the 14 covariates. We could estimate  $\beta$ -coefficients for the reference distribution directly from the sample moments of  $r_{xc}r_{yc}$ . But, we only have 14 covariates and the sample error for the sample moments would be too large. Thus, I will estimate the  $\beta$ -coefficients for the reference distribution from the formulae (5), (12), (12'), (12''), and (13) in Chapter 3.

First, I got the estimated population correlations  $\rho_{xc}$ ,  $\rho_{yc}$ , and  $\rho_{xy}$  as follows:

$$\begin{aligned}\hat{\rho}_{xc} &= \frac{1}{14} \sum_{i=1}^{14} r_{xc_i \cdot c_1 \dots c_{i-1} c_{i+1} \dots c_{14}} = .235; \\ \hat{\rho}_{yc} &= \frac{1}{14} \sum_{i=1}^{14} r_{yc_i \cdot c_1 \dots c_{i-1} c_{i+1} \dots c_{14}} = .260; \\ \hat{\rho}_{xy} &= r_{xy \cdot c_1 \dots c_{14}} = .325.\end{aligned}$$

Substituting those estimated population values into Formulae (5), (12), (12'), (12''), and

(13) gives us the estimated values of  $\beta_1$  and  $\beta_2$  as  $b_1 = .0073$  and  $b_2 = 2.995^{[11]}$ . Also, from (14), I obtained  $\hat{\kappa} = -.1726 < 0$ , verifying that the reference distribution can be approximated by a Pearson Type I distribution (cf. Figure 1). From the formulae at the end of Chapter 3, I obtained the distribution function  $f(k)$  for  $r_{xc}r_{yc}$  as follows:

$$f(k) = 33.18 \left(1 + \frac{k}{.14}\right)^{87.28} \left(1 - \frac{k}{.28}\right)^{180.99}, \quad -.14 \leq k \leq .28.$$

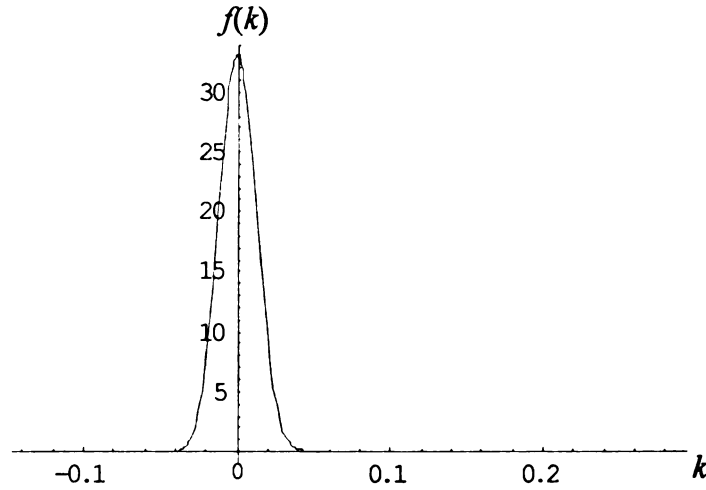


Figure 10. Distribution function of  $k = r_{xc}r_{yc}$ .

Given the ITCV for *Father's Occupation* as .228 (Frank, 2000), I obtained<sup>[12]</sup>

$$\text{IOROCI} = \int_{-\infty}^{\text{ITCV}} f(k) dk = \int_{-\infty}^{.228} f(k) dk = .99999 \text{ via a numerical integration using}$$

<sup>[11]</sup> It is not necessary to use the corrected  $\mu_3$  and  $\mu_4$  for this example, because the sample size  $N$  is very big, but the corrected  $\mu_3$  and  $\mu_4$  would be applied for small  $N$ , say  $N < 90$ : First, one would obtain the estimated mean values of  $\rho_{xc}$  and  $\rho_{yc}$ ; second, substitute these into Equations (17) and (19) to obtain  $\hat{D}_{\mu_3}$  and  $\hat{D}_{\mu_4}$ ; and last, from  $\tilde{A}_{\mu_3} = A_{\mu_3} + \hat{D}_{\mu_3}$  and  $\tilde{A}_{\mu_4} = A_{\mu_4} + \hat{D}_{\mu_4}$  one would obtain the corrected  $\mu_3$  and  $\mu_4$ .

<sup>[12]</sup> There are also two other methods to obtain IOROCI: (a) Programming using Bowman and Shenton's (1979) approach; and (b) looking up the probability value in Pearson and Hartley's table (1972) with interpolation.

*Mathematica*. According to the guidance of the interpretation for IOROCI in the previous section, we would like to say that the inference regarding *Father's Occupation* on *Educational Attainment* is very robust and that it is very unlikely for the impact of a covariate to alter the inference. To the extent that the impact of *Father's Education* is represented by the impacts of the other covariates, we could then conclude that it is very unlikely that the impact of *Father's Education* will alter our inference about *Father's Occupation*.

## **Chapter 8**

### **DISCUSSION**

#### **8.1 Conclusions and Limitations**

Causal inference is a controversial topic in the social sciences, where we are often not able to conduct a randomized experiment or statistically control for all possible confounding variables. In the literature, some have attempted to deal with the crisis in causal inference, but most approaches have practical or theoretical limitations. Frank's (2000) index is the soundest approach that leads to a very promising methodology.

Frank's index is composed of the product of two dependent correlation coefficients: the correlation between the predictor of interest and the outcome and the correlation between the confounding variable and the outcome. The index is most informative when evaluated against a reference distribution defined by the impacts of existing covariates, due to the immeasurability of confounding variables. Thus, Frank used the reference distribution to statistically assess how robust the causal inference for a given predictor on an outcome is to the impacts of uncontrolled confounding variables. Unfortunately, when he generated the reference distribution, Frank used an approximation based on approximately normally distributed Fisher  $z$  for each  $r$ , and then another approximation to

the product of two normal variables; and therefore, this doubly asymptotic result is very rough.

The current study provided a much more accurate approximation to the reference distribution with a closed form—Pearson Type I (*Beta*) distribution than Frank's. With the more accurate approximation to the reference distribution, we can make a more valid conclusion about whether the causal inference for a given predictor on an outcome is robust to the impacts of other possible confounding variables, that is, all uncontrolled confounds are unlikely to have an impact great enough to alter the inference. This methodology would allow for multiple partial causes in the complex social phenomenon that we study. Therefore, we are able to inform the controversy about causal inference that arises from the use of statistical linear models in the social sciences.

As stated in Chapter 1, some researchers may be uncomfortable with the use of measured covariates to generate a reference distribution for the impact of an unknown confounding variable. But, we acknowledge that this use of the reference distribution is only as valid as is the set of covariates on which it is based, which is no different from any other inference from a sample that must be representative of the population. In this light, the impact of existing covariates represents important information by which to assess the ITCV. Ultimately, this approach also allows us to rescale ITCV to a probability scale, IOROCI, that can be represented by the existing covariates.

We hope to have homogenous impacts of covariates, because it will produce a desirable reference distribution that is representative of the unknown confounding variable. We can use Q-statistic, like one in meta-analysis, or P-P plot, to test the

homogeneity. In case of having heterogeneous impacts of covariates, it is reasonable, but arguable, to use the maximum impact to get a reference distribution. Concerns about heterogeneity would be greater if the impacts of covariates were obtained from multiple estimated models accounting for different sets of covariates. Researchers should evaluate the likelihood of the observed extreme impacts according to theoretical and statistical criteria. In any case, the solution that is provided in the current study is technical, only provides a quantitative discussion, and not solves the problem of causal inference.

A simulation study was conducted to check the accuracy of the approximation method in this study. The simulation results (Tables 3 to 14), along with the line charts (Figure 3) and the P-P plot (Figure 6), indicate that the approximation technique is generally much better than Frank's (2000). We also understand that the approximation is not very favorable with respect to the third and the fourth moments when  $N$  is small. The problem may come from the lower-order approximation to the covariance of  $r_{xc}$  and  $r_{yc}$  (cf., the expression of  $\sigma_{r_{xc}, r_{yc}}$  in Equation 5). Although the approximation correction by the regression approach nicely solved the problem, it would be worth finding a better approximation to the covariance of  $r_{xc}$  and  $r_{yc}$  through further study.

For researchers in the field of the social sciences, this study provides IOROCI to assess the robustness of causal inferences drawn from correlational data via statistical linear models. Thus, IOROCI provides a statistical index for researchers to evaluate the robustness of causal inferences.

Note that the current study is based on linear modeling which assumes that all dependent and independent variables are measured without measurement error. This

concern does not apply directly to the confounding variable which is assumed to be perfectly measured to maximize impact. But, it does apply to the distribution of impacts of covariates used to generate the reference distribution. To the extent that the covariates are unreliably measured, their impacts will underestimate their true impacts. When reliabilities are known, a correlation disattenuation is recommended. That is, one can conduct all analyses on a correlation matrix that has been adjusted for attenuation. It is especially important to use a correlation disattenuation when the impact of covariates comprises two partial correlations, because partial correlations produce an underestimated small impact.

## **8.2 Extensions**

In addition to the applications of the distribution of the product of two dependent correlation coefficients to causal inference in educational research, the distribution of the product of two dependent correlation coefficients may be also applicable to other issues in the social sciences. For example, the distribution of the product could be applied to assessing indirect effects with respect to mediating processes in path analysis. This is because the indirect effect can be considered as a product of two correlation coefficients. Thus, if we know the distribution of the product, we can better understand the likelihood of observing an indirect effect of a given size.

Another task would be to extend the simulation range for the correlation coefficient beyond .50. Although .50 is a large size for correlation coefficients in the social sciences (Cohen & Cohen, 1983), in social research on real data we often see correlation

coefficients larger than .50. Thus, it would be a valuable exploration to see if the conclusions of the current study still hold for correlation coefficients larger than .50. That is, it is important to know whether the Pearson Type I distribution still can be a reasonable approximate distribution for the product of two dependent correlation coefficients when some population correlations are larger than .50. It would also be valuable to conduct a larger number of replications for small  $\rho_{xc}$  or  $\rho_{yc}$ , where approximation was poor, or to obtain better estimates of small  $p$ -values.

Note that the approximation method in this paper is based on the assumption that the three initial variables  $X$ ,  $Y$ , and  $C$  follow a trivariate normal distribution. However, it is not necessary for predictors and covariates to be normally distributed in linear models. Hence, for an extension of this study, it would be interesting to find an approximate distribution of the product of two dependent correlation coefficients for non-normal, categorical or mixed initial variables  $X$ ,  $Y$ , and  $C$ .

## APPENDIX A

Below is the SPSS code for calculating the approximate moments for  $N = 28$  or  $M = N + 6 = 34$  as an example, where

$$\begin{aligned}
 b_{xc} &= b_{r_{xc}}; & s2xc0 &= \sigma_{r_{xc}}^{(2)}; & s3xc0 &= \sigma_{r_{xc}}^{(3)}; & s4xc0 &= \sigma_{r_{xc}}^{(4)}; \\
 b_{yc} &= b_{r_{yc}}; & s2yc0 &= \sigma_{r_{yc}}^{(2)}; & s3yc0 &= \sigma_{r_{yc}}^{(3)}; & s4yc0 &= \sigma_{r_{yc}}^{(4)}; \\
 p_{xc} &= \rho_{xc}; & s2xc &= E[(\Delta r_{xc})^2]; & s3xc &= E[(\Delta r_{xc})^3]; & s4xc &= E[(\Delta r_{xc})^4]; \\
 p_{yc} &= \rho_{yc}; & s2yc &= E[(\Delta r_{yc})^2]; & s3yc &= E[(\Delta r_{yc})^3]; & s4yc &= E[(\Delta r_{yc})^4]; \\
 p_{xy} &= \rho_{xy}; & sxcyc0 &= \sigma_{r_{xc}, r_{yc}}; & sxcyc &= E(\Delta r_{xc} \Delta r_{yc}); \\
 ekl &= E[(\Delta r_{xc})^k (\Delta r_{yc})^l], \quad k, l = 1, 2, 3, \text{ or } 4; & m_i &= \mu'_i, \quad m_{ui} = \mu_i, \quad i = 1, 2, 3, \text{ or } 4.
 \end{aligned}$$

```

*****
*
*   The SPSS program for calculating approximate moments using SPSS
*   data "PCM.sav". (PCM = Population correlation matrix)
*
*****

GET
  FILE='C:\My Documents\Academics\Dissertation\Smulation\PCM.sav'.
EXECUTE .

**** N = 28 ****.

COMPUTE bxc = (-pxc*(1 - pxc ** 2)/(2*34)*(1 + 9*(3 + pxc**2)/(4*34)
+ 3*(121 + 70*pxc**2 + 25*pxc**4)/(8*34**2)
+ 3*(6479+4923*pxc**2+2925*pxc**4+1225*pxc**6)/(64*34**3)
+ 3*(86341+77260*pxc**2+58270*pxc**4+38220*pxc**6
+ 19845*pxc**8)/(128*34**4)) .
EXECUTE .

COMPUTE byc = (-pyc * (1 - pyc ** 2) / (2*34)) * (1+9*(3+pyc**2)/(4*34)
+ 3*(121+70*pyc**2+25*pyc**4)/(8*34**2)
+ 3*(6479+4923*pyc**2+2925*pyc**4+1225*pyc**6)/(64*34**3)
+ 3*(86341+77260*pyc**2+58270*pyc**4+38220*pyc**6
+ 19845*pyc**8)/(128*34**4)) .
EXECUTE .

COMPUTE s2xc0 = ((1 - pxc ** 2) ** 2 / 34) * (1+(14+11*pxc**2)/(2*34)
+ (98+130*pxc**2+75*pxc**4)/(2*34**2)+(2744+4645*pxc**2+4422*pxc**4
+ 2565*pxc**6)/(8*34**3)+19208+37165*pxc**2+44499*pxc**4+40299*pxc**6
+ 26685*pxc**8)/(8*34**4)) .
EXECUTE .

COMPUTE s2yc0 = ((1 - pyc ** 2) ** 2 / 34) * (1+(14+11*pyc**2)/(2*34)
+ (98+130*pyc**2+75*pyc**4)/(2*34**2)+(2744+4645*pyc**2+4422*pyc**4
+ 2565*pyc**6)/(8*34**3)+(19208+37165*pyc**2+44499*pyc**4+40299*pyc**6
+ 26685*pyc**8)/(8*34**4)) .

```

```

+ 2565*pyc**6)/(8*34**3)+(19208+37165*pyc**2+44499*pyc**4+40299*pyc**6
+ 26685*pyc**8)/(8*34**4)) .
EXECUTE .
COMPUTE s3xc0 = (-pxc*(1 - pxc ** 2)**3/(34**2))*(6+(69+88*pxc**2)/34
+ 3*(797+1691*pxc**2+1560*pxc**4)/(4*34**2)+3*(12325+33147*pxc**2
+ 48099*pxc**4+44109*pxc**6)/(8*34**3)) .
EXECUTE .

COMPUTE s3yc0 = (-pyc*(1 - pyc ** 2)**3/(34**2))*(6+(69+88*pyc**2)/34
+ 3*(797+1691*pyc**2+1560*pyc**4)/(4*34**2)+3*(12325+33147*pyc**2
+48099*pyc**4+44109*pyc**6)/(8*34**3)) .
EXECUTE .

COMPUTE s4xc0 =(3*(1 - pxc ** 2) ** 4 /(34**2))* (1+(12+35*pxc**2)/34
+ (436+2028*pxc**2+3025*pxc**4)/(4*34**2)+(3552+20009*pxc**2
+ 46462*pxc**4 +59751*pxc**6)/(4*34**3)) .
EXECUTE .

COMPUTE s4yc0 = (3*(1 - pyc ** 2) ** 4 /(34**2))* (1+(12+35*pyc**2)/34
+ (436+2028*pyc**2+3025*pyc**4)/(4*34**2)+(3552+20009*pyc**2
+ 46462*pyc**4+59751*pyc**6)/(4*34**3)) .
EXECUTE .

COMPUTE sxcyc0 = (pxy * (1 - pxc ** 2 - pyc ** 2)
- pxc * pyc * (1 - pxc ** 2 - pyc ** 2 - pxy ** 2) / 2)/34 .
EXECUTE .

COMPUTE s2xc = s2xc0 + bxc**2.
EXECUTE .

COMPUTE s2yc = s2yc0 + byc**2.
EXECUTE .

COMPUTE s3xc = s3xc0 - 3*s2xc0*bxc - bxc**3.
EXECUTE .

COMPUTE s3yc = s3yc0 - 3*s2yc0*byc - byc**3.
EXECUTE .

COMPUTE s4xc = s4xc0 - 4*s3xc0*bxc + 6*s2xc0*bxc**2 + bxc**4.
EXECUTE .

COMPUTE s4yc = s4yc0 - 4*s3yc0*byc + 6*s2yc0*byc**2 + byc**4.
EXECUTE .

COMPUTE sxcyc = sxcyc0 + bxc*byc.
EXECUTE .

COMPUTE e21 = bxc**2*byc + s2xc0*byc + 2*sxcyc0*bxc.
EXECUTE .

Compute e12 = byc**2*bxc + s2yc0*bxc + 2*sxcyc0*byc.
Execute.

Compute e31 = 3*s2xc0*sxcyc0 + bxc**3*byc + 3*s2xc0*bxc*byc
+ 3*sxcyc0*bxc**2.
Execute.

```

```
Compute e13 = 3*s2yc0*sxcyc0 + byc**3*bxc + 3*s2yc0*byc*bxc
+ 3*sxcyc0*byc**2 .
```

```
Execute.
```

```
Compute e22 = s2xc0*s2yc0 + 2*sxcyc0**2 + bxc**2*byc**2 + s2xc0*byc**2
+ s2yc0*bxc**2 + 4*sxcyc0*bxc*byc.
```

```
Execute.
```

```
COMPUTE m1 = pxc*pyc + pxc*byc + pyc*bxc + sxcyc.
```

```
EXECUTE .
```

```
COMPUTE m2 = pxc**2*pyc**2 + pxc**2*s2yc + pyc**2*s2xc + e22
+ 2*pxc**2*pyc*byc + 2*pxc*pyc**2*bxc + 2*pyc*e21 + 2*pxc*e12
+ 4*pxc*pyc*sxcyc.
```

```
EXECUTE .
```

```
COMPUTE m3 = pxc**3*pyc**3 + pxc**3*s3yc + pyc**3*s3xc
+ 3*pxc**3*pyc**2*byc + 3*pxc**3*pyc*s2yc + 3*pxc**2*pyc**3*bxc
+ 3*pxc*pyc**3*s2xc + 3*pxc**2*e13 + 3*pyc**2*e31
+ 9*pxc**2*pyc**2*sxcyc + 9*pxc**2*pyc*e12 + 9*pxc*pyc**2*e21
+ 9*pxc*pyc*e22.
```

```
EXECUTE .
```

```
COMPUTE m4 = pxc**4*pyc**4 + pxc**4*s4yc + pyc**4*s4xc
+ 4*pxc**4*pyc**3*byc + 4*pxc**3*pyc**4*bxc + 4*pxc**4*pyc*s3yc
+ 4*pxc*pyc**4*s3xc + 6*pxc**4*pyc**2*s2yc + 6*pxc**2*pyc**4*s2xc
+ 16*pxc**3*pyc**3*sxcyc + 16*pxc**3*pyc*e13 + 16*pxc*pyc**3*e31
+ 24*pxc**3*pyc**2*e12 + 24*pxc**2*pyc**3*e21 + 36*pxc**2*pyc**2*e22.
```

```
EXECUTE .
```

```
COMPUTE mu2 = m2 -m1**2.
```

```
EXECUTE .
```

```
COMPUTE mu3 = m3-3*m1*m2+2*m1**3.
```

```
EXECUTE .
```

```
COMPUTE mu4 = m4-4*m1*m3+6*m1**2*m2-3*m1**4.
```

```
EXECUTE .
```

```
SAVE OUTFILE='C:\My Documents\Academics\Apprenticeship\Smulation\PCM
and Approximate Moments (N = 28).sav'
```

```
/keep=pxc pyc pxy m1 mu2 mu3 mu4
```

```
/COMPRESSED .
```

## REFERENCES

- Abbott, A. (1998). The causal devolution. *Sociological Methods and Research*, 27(2), 148-181.
- Anderson, S., Auquier, A., Hauck, W. W., Oakes, D., Vandaele, W., & Weisberg, H. I. (1980). *Statistical Methods for Comparative Studies: Techniques for Bias Reduction*. New York: Wiley.
- Anderson, T. W. (1958). *Introduction to multivariate statistical analysis*. New York: John Wiley and Sons.
- Aroian, L. A. (1947). The probability function of the product of two normally distributed variables. *The Annals of Mathematical Statistics*, 18(2), 265-271.
- Aroian, L. A., Taneja, V. S., & Cornwell, L. W. (1978). Mathematical forms of the distribution of the product of two normal variables. *Communications in Statistics: Part A—Theory and Methods*, 7(2), 165-172.
- Bowman, K. O., & Shenton, L. R. (1979). Approximate percentage points for Pearson distributions. *Biometrika*, 66, 147-151.
- Boyer, J. E., Jr., Palachek, A. D., & Schucany, W. R. (1983). An empirical study of related correlation coefficients. *Journal of Educational Statistics*, 8(1), 75-86.
- Browne, M. W. (1968). A comparison of factor analytic techniques. *Psychometrika*, 33, 267-334.
- Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1991). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.
- Choi, S. C. (1977). Tests of equality of dependent correlation coefficients. *Biometrika*, 64(3), 645-647.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Erlbaum.
- Cohen, J., & Cohen, P. (1983). *Applied multivariate regression/correlation analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Erlbaum.
- Cook, T., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.

- Cornwell, L. W., Aroian, L. A., & Taneja, V. S. (1978). Numerical evaluation of the distribution of the product of two normal variables. *Journal of Statistical Computing and Simulation*, 7, 123-131.
- Craig, C. C. (1936). On the frequency function of  $xy$ . *The Annals of Mathematical Statistics*, 7(1), 1-15.
- Craig, C. C. (1942). On frequency distributions of the quotient and of the product of two statistical variables. *American Mathematical Monthly*, 49(1), 24-32.
- Duncan, O. D., Featherman, D. L., & Duncan, B. (1972). *Socioeconomic background and achievement*. New York: Seminar Press.
- Dunn, O. J., & Clark, V. (1969). Correlation coefficients measured on the same individuals. *Journal of the American Statistical Association*, 64, 366-377.
- Dunn, O. J., & Clark, V. (1971). Comparison of tests of the equality of dependent correlation coefficients. *Journal of the American Statistical Association*, 66, 904-908.
- Elderton, W. P., & Johnson, N. L. (1969). *Systems of frequency curves*. London, UK: Cambridge University Press.
- Featherman, D. L., & Hauser, R. M. (1976). Sexual inequalities and socioeconomic achievement in the U.S. 1962-1973. *American Sociological Review*, 41, 462-483.
- Fisher, R. A. (1924). The distribution of the partial correlation coefficient. *Metron*, 3, 329-332.
- Frank, K. A. (2000). The impact of a confounding variable on a regression coefficient. *Sociological Methods and Research*, 29(2), 147-194.
- Ghosh, B. K. (1966). Asymptotic expansions for the moments of the distribution of correlation coefficient. *Biometrika*, 35, 258-262.
- Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81, 945-970.
- Holland, P. W. (1988). Causal inference, path analysis, and recursive structural equations models (With discussion). In C.C. Clogg (Ed.), *Sociological methodology* (pp. 449-493). Washington, DC: American Sociological Association.
- Hotelling, H. (1936). Relations between two sets of variables. *Biometrika*, 28, 321-377.

- Hotelling, H. (1940). The selection of variables for use in prediction with some comments on the general problem of nuisance parameters. *The Annals of Mathematical Statistics*, 11(3), 271-283.
- Jacobs, J. E., Finken, L. L., Griffin, N. L., & Wright, J. D. (1998). The career plans of science-talented rural adolescent girls. *American Educational Research Journal*, 35, 681-704.
- Kendall, M., Sir, & Stuart, A. (1977). *The advanced theory of statistics, Vol. I, Distribution theory* (4th ed.). New York: Macmillan.
- Konishi, S. (1978). Asymptotic expansion for the distribution of functions of a correlation matrix. *Journal of Multivariate Analysis*, 9, 259-266.
- Konishi, S. (1979). An approximation to the distribution of the sample correlation coefficient. *Biometrika*, 65(3), 654-656.
- Kraemer, H. C. (1973). Improved approximation to the non-null distribution of the correlation coefficient. *Journal of the American Statistical Association*, 68, 1004-1008.
- Lee, O. (1999). Science knowledge, world views, and information sources in social and cultural contexts: making sense after a nature disaster. *American Educational Research Journal*, 36, 187-219.
- Mathai, A. M., & Saxena, R. K. (1969). Distribution of a product and the structural set up of densities. *The Annals of Mathematical Statistics*, 40(4), 1439-1448.
- Mathai, A. M., & Saxena, R. K. (1978). *The H-function with applications in statistics and other disciplines*. New York: John Wiley & Sons.
- May, K., & Hittner, J. B. (1997a). A note on statistics for comparing dependent correlations. *Psychological Reports*, 80, 475-480.
- May, K., & Hittner, J. B. (1997b). Tests for comparing dependent correlations revisited: A Monte Carlo study. *Journal of Experimental Education*, 65(3), 257-269.
- McKim, V. R., & Turner, S. P. (Eds.). (1997). *Causality in crisis?: Statistical methods and the search for causal knowledge in the social sciences*. Notre Dame, IN: University of Notre Dame Press.
- Meeker, W. Q., Cornwell, L. W., & Aroian, L. A. (1981). The product of two normally distributed random variables. In W. Kenney & R. Odeh (Eds.), *Selected tables in mathematical statistics* (Vol. VII, pp. 1-256). Providence, IR: American Mathematical Society.

- Meeker, W. Q., & Escobar, L. A. (1994). An algorithm to compute the CDF of the product of two normal random variables. *Communications in Statistics: Part B—Simulation and Computation*, 23(1), 271-280.
- Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, 111(1), 172-175.
- Miller, J. W., Ellsworth, R., & Howell, J. (1986). Public elementary schools which deviate from the traditional SES-achievement relationship. *Educational Research Quarterly*, 10, 31-50.
- National Center for Education Statistics (1996). *NELS:88-94 data files and electronic codebook system*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- Neill, J. J., & Dunn, O. J. (1975). Equality of dependent correlation coefficients. *Biometrics*, 31, 531-543.
- Okagaki, L., & Frensch, P. A. (1998). Parenting and children's school achievement: A multiethnic perspective. *American Educational Research Journal*, 35, 123-144.
- Olkin, I. (1967). Correlations Revisited. In J. C. Stanley (Ed.), *Improving experimental design and statistical analysis* (pp. 102-128). Chicago: Rand McNally & Company.
- Olkin, I., & Finn, J. (1990). Testing correlated correlations. *Psychological Bulletin*, 108, 330-333.
- Olkin, I., & Siotain, M. (1976). Asymptotic distribution of function of a correlation matrix. In S. Ikeda (Ed.), *Essays in probability and statistics* (pp. 235-251). Tokyo, Japan: Shinko Tsusho Co., Ltd.
- Pearl, J. (2000). *Causality: Models, reasoning, and inference*. New York: Cambridge University Press.
- Pearson, E. S., & Hartley, H. O. (1972). *Biometrika tables for statisticians, Vol. II*. New York: Cambridge University Press.
- Pearson, K. (1895). Contributions to the mathematical theory of evolution. II. Skew variations in homogeneous material. *Philosophical Transactions of the Royal Society of London, Series A*, 186, 343-414.
- Portes, P. R. (1999). Social and psychological factors in the academic achievement of children of immigrants: a cultural history puzzle. *American Educational Research Journal*, 36, 489-507.

- Rice, J. A. (1995). *Mathematical statistics and data analysis* (2nd Ed.). Belmont, CA: Duxbury Press.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688-701.
- Sobel, M. E. (1996). An introduction to causal inference. *Sociological Methods and Research*, 24(3), 353-379.
- Sobel, M. E. (1998). Causal inference in statistical models of the process of socioeconomic achievement: A case study. *Sociological Methods and Research*, 27(2), 318-348.
- Solomon, D., Battistich, V., & Hom, A. (1996). Teacher beliefs and practices in schools serving communities that differ in socioeconomic level. *Journal of Experimental Education*, 64, 327-347.
- Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search* [Lecture Notes in Statistics 81]. New York: Springer-Verlag.
- Springer, M. D. (1983). [Review of *Selected tables in mathematical statistics* (Vol. VII), *The product of two normally distributed random variables*]. *Technometrics*, 25(2), 211.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87(2), 245-251.
- Trusty, J., Peck, H. I., & Mathews, J. (1994). Achievement, socioeconomic status and self-concepts of fourth-grade students. *Child Study Journal*, 24, 281-298.
- Wallgren, C. M. (1980). The distribution of the product of two correlated  $t$  variables. *Journal of the American Statistical Association*, 75, 996-1000.
- Williams, E. J. (1959). Significance of difference between two non-independent correlation coefficients. *Biometrics*, 15, 135-136.
- Wolfe, D. A. (1976). On testing equality of related correlation coefficients. *Biometrika*, 63, 214-215.
- Wolfram, S. (1999). *The Mathematica book* (4th Ed.). New York: Cambridge University Press.
- Woodward, J. (1997). Causal models, probabilities, and invariance. In V. R. McKim & S. P. Turner (Eds.), *Causality in crisis?: Statistical methods and the search for causal knowledge in the social sciences* (pp. 265-315). Notre Dame, IN: University of Notre Dame Press.