1

35518

National Library
of CanadaBibliothèque nationale
du CanadaCANADIAN THESES
ON MICROFICHETHÈSES CANADIENNES
SUR MICROFICHENAME OF AUTHOR/NOM DE L'AUTEUR MICHAEL LERNERTITLE OF THESIS/TITRE DE LA THÈSE "Berlin 1961 - A Replication of Tanter's Model"UNIVERSITY/UNIVERSITÉ Carleton UniversityDEGREE FOR WHICH THESIS WAS PRESENTED/
GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE M.A.YEAR THIS DEGREE CONFERRED/ANNÉE D'OBTENTION DE CE DEGRÉ 1977NAME OF SUPERVISOR/NOM DU DIRECTEUR DE THÈSE John H. Sigler

Permission is hereby granted to the NATIONAL LIBRARY OF
CANADA to microfilm this thesis and to lend or sell copies
of the film.

The author reserves other publication rights, and neither the
thesis nor extensive extracts from it may be printed or other-
wise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈ-
QUE NATIONALE DU CANADA de microfilmer cette thèse et
de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la
thèse ni de longs extraits de celle-ci ne doivent être imprimés
ou autrement reproduits sans l'autorisation écrite de l'auteur.

DATED/DATE September 30/77 SIGNED/SIGNÉ Michael LernerPERMANENT ADDRESS/RÉSIDENCE FIXE Apt. 20 - 110 Gloucester St,Ottawa, Ontario

National Library of Canada

Cataloguing Branch
Canadian Theses Division

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada

Direction du catalogage
Division des thèses canadiennes

NOTICE

AVIS

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

**THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED**

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

**LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE**

BERLIN 1961: A REPLICATION OF TANTER'S MODEL

by

Michael Lerner, B.A.

A thesis submitted to the Faculty of
Graduate Studies and Research in partial fulfillment
of the requirements for the degree of
Master of Arts
Department of Political Science

Carleton University

OTTAWA, Ontario

September 12, 1977

(c)

Michael Lerner 1977

The undersigned recommend to the Faculty
of Graduate Studies and Research acceptance of the thesis

"Berlin 1961: A Replication of Tanter's Model"

submitted by Michael Lerner, B.A. (Hon.)
in partial fulfilment of the requirements for
the degree of Master of Arts
in Political Science

.....
John H. Sigler, Supervisor

.....
Robert J. Jackson, Chairman

Carleton University

September 30, 1977

The first Professor I saw was in a very large Room, with Forty Pupils about him. After Salutation, observing me to look earnestly upon a Frame, which took up the greatest Part of both the Length and Breadth of the Room; he said, perhaps I might wonder to see him employed in a Project for improving speculative Knowledge by practical and mechanical Operations. . . . He then led me to the Frame, about the Sides whereof all his pupils stood in Ranks. It was Twenty Foot square, placed in the Middle of the Room. The Superficies was composed of several Bits of Wood, about the bigness of a Dye, but some larger than others. They were all linked together by slender Wires. These Bits of Wood were covered on every Square with Papers pasted on them; and on these Papers were written all the Words of their Language in their several Moods, Tenses, and Declensions, but without any Order. The Professor then desired me to observe, for he was going to set his Engine at work. The Pupils at his command took each of them hold of an Iron Handle, whereof there were Forty fixed round the Edges of the Frame; and giving them a sudden Turn, the whole Disposition of the Words was entirely changed. He then commanded Six and Thirty of the Lads to read the several Lines softly as they appeared upon the Frame; and where they found three or four Words together that might make Part of a Sentence, they dictated to the four remaining Boys who were Scribes. . . .

Six Hours a-Day the young Students were employed in this Labour; and the Professor shewed me several Volumes in large Folio already collected, of broken Sentences, which he intended to piece together; and out of those rich Materials to give the World a compleat Body of all Arts and Sciences; which however might be still improved, and much expedited, if the Publick would raise a Fund for making and employing five Hundred such Frames in Lagado, and oblige the Managers to contribute in common their several Collections.

Swift Gulliver's Travels 3. 5.

ABSTRACT

This thesis is a replication of Tanter's (1974) analysis of the Berlin 1961 crisis which appeared in his Modelling and Managing International Conflicts. The data used are from a pre-release version of the Comparative Analysis of Dyadic Interstate Conflict event data set, which, unlike many other event data sets, focuses upon specific international crises. Tanter's model is critically examined, and suggestions for improved specifications are made. In particular, the functional form of the independent variables over time and their relationships to one another are more explicitly specified. A more appropriate estimation technique, autocorrelation coefficient-corrected pseudo-generalized least squares, is suggested and used for the autoregressive model suggested by Tanter. Moreover, hypothesis-testing in the replication makes use of pooled standard errors to provide more rigorous testing of the differences between coefficients than did the original. Contrary to Tanter, the results presented here indicate organizational behavior dominated both the North Atlantic Treaty Organization (NATO) and the Warsaw Treaty Organization (WTO) during the intensification of the crisis. During the reduction, more interaction between the two alliances was found than had been in the previous study. Similar to Tanter, however, these findings lead to the conclusion that organizational behavior dominated WTO during the entire crisis. Finally, this

thesis also suggests procedures to be followed in replication
in international politics and outlines some of the difficulties
that may be encountered along with means of dealing with them.

ACKNOWLEDGEMENTS

I wish to thank my two supervisors Professor Hart and Professor Sigler. Professor Hart constantly encouraged and advised me beyond the bounds of ordinary supervision. Professor Sigler still undertook my supervision despite being on leave and made available the data used in this thesis. Naturally, all faults found here are my own.

TABLE OF CONTENTS

INTRODUCTION: PRACTICAL METHODOLOGY.....	1
Overview	1
Rationale	2
Organization	7
Summary	11
Notes	12
CHAPTER 1	
REPLICATION AND AMELIORATION OF THE CUMULATION PROBLEM: IMPROVING INFERENCE.....	14
Overview	14
Cumulation and replication	14
Current practice of replication	18
Replicatory procedure	21
Summary	25
Notes	26
CHAPTER 2	
ESTABLISHING REPLICABILITY: ON TANTER'S AUTOREGRESSIVE CRISIS MODEL.....	27
Overview	27
The model	28
Specification	32
Data-generation	37
Estimation	45
Summary	51
Notes	51

CHAPTER 3

CREATING AND COMPARING THE DATA BASES.....	53
Overview	53
Creating the data bases	54
On estimation	60
The effect of the transformation	62
Stability of measurement	63
Comparison of the data sets	69
Summary	72
Notes	73

CHAPTER 4

ON THE GENERALIZABILITY OF TANTER'S RESULTS.....	74
Overview	74
A plausible solution for autoregressive models	74
Regression results and substantive findings	78
Comparison of Tanter's and replicatory results	95
Summary	104
CONCLUSIONS AND FURTHER SUGGESTIONS.....	106
Overview	106
On the results of the replicatory models	107
An alternative model	108
On the use of time-series	112
Measurement and replication	113
Replication as criticism	114
Final summation	116

APPENDIX A

DATA USED IN THIS THESIS..... 118

BIBLIOGRAPHY..... 139

LIST OF TABLES

2-1—Tanter's expectations of relative phase dominance of parameters.....	35
2-2—The East-West categories.....	38
2-3—The WEIS categories.....	39
2-4—The CADIC categories.....	44
2-5—Tanter's tri-phase results,.....	48
2-6—Tanter's two phase results.....	50
3-1—Daily comparison of transformed and untransformed data.....	64
3-2—Three day comparison of transformed and untransformed data.....	65
3-3—Daily comparison of scores and frequencies.....	67
3-4—Three-day comparison of scores and frequencies.....	68
3-5—Daily comparison of Tanter's and replicatory data..	70
3-6—Three day comparison of Tanter's and replicatory data.....	71
4-1—Replicatory regression results.....	79
4-2—Dominance of parameters.....	91
4-3—Comparison of Tanter's and replicatory coefficients.....	98
A-1—Tanter's WTO data daily.....	119
A-2—Tanter's NATO data daily.....	122
A-3—Replicatory WTO data daily.....	125
A-4—Replicatory NATO data daily.....	128
A-5—Tanter's WTO data three-day.....	131

A-6—Tanter's NATO data three-day.....	133
A-7—Replicatory WTO data three-day.....	135
A-8—Replicatory NATO data three-day.....	137

LIST OF ILLUSTRATIONS

3.1—The East-West scales.....	58
3.2—The East-West scales.....	59
6.1—Graph of polynomial distributed lags for intensification.....	110
6.2—Graph of polynomial distributed lags for reduction.....	111

INTRODUCTION: PRACTICAL METHODOLOGY

Overview

Many authorities (Deutsch, 1968; McClelland, 1972: 18-23) claim that the study of international politics has drastically changed during the past generation with an increasing emphasis upon positivism, transforming a hitherto legal and historical discipline into a behavioral social science. During the transformation, data-generation advanced, sophistication in statistical inference developed, models were proposed, and results were generated. Engaged in these activities, students of international politics became methodologists by debating the amenability of their discipline to scientific procedures as seen in the exchange between Hedley Bull (1968) and Morton A. Kaplan (1968), by setting forth general guide lines for scientific research (see McGowan, 1975), and by finally searching for signs of scientific accumulation¹ amidst the results generated during the past two decades. Although theoretical methodological statements abound, students of international politics avoid examining their practical methodology—problems of measurement, statistical estimation, and inference. This thesis, in addition to addressing the theoretical methodological issue of scientific accumulation, closely examines the intertwined practical problems of event data measurement, time-series regression, and replication. The substantive framework for these topics is the replication of

Raymond Tanter's (1974) autoregressive model² for the 1961 Berlin crisis. Here, Albert Legault et al.'s Comparative Analysis of Dyadic Interstate Conflict event data (1973; 1974; hereafter referred to as CADIC event data) is substituted for Walter H. Corson's Conflict and Cooperation in East-West Crises event data (1974; hereafter referred to as East-West event data)³ used by Tanter for the 1961 Berlin crisis.

Rationale

The rationale for this thesis rests on three simple observations of current preoccupations in the study of international politics: the scientific accumulation issue; the tremendous growth of the event data movement; the increasing use of time-series regression. Each of these three simple observations are dealt with in turn and then related to one another to reveal the rationale for this thesis.

The scientific accumulation issue can be epitomized by the following query: "Are the topics of their individual research efforts [those of students of international politics] close enough that each can build on the work of the others or is each working on such a different topic that there is little or no building?". One of the concerns here is whether students of international politics have created a body of knowledge that is confidently used in their ongoing research. In contrast to the scientifically accumulative situation where "a second research effort on a topic can rely on findings of a first research effort in such a way that it does not have to start from scratch," the study of international politics is

marked by repeated research efforts that fail to take cognizance of others and the creation of new variables that have no specified relationship to old variables (Rosenau, 1973: 14-15, citing Boynton, letter [8 February 1973]). Students of international politics cannot easily accept past results because their stability across various scope condition is unknown and are bewildered by uninvestigated measurement systems. Here, replication—repetition of previous research in a deliberately altered setting (scope conditions, measurement system, and procedures), where the two sets of results are compared in the light of known variation between the two settings so that the assumption of generalizability (similar settings produce similar results) is tested—ameliorates scientific accumulation concerns by assessing stability of results through testing for generalizability and by investigating the properties of different measurement systems. Replication, however, is not easy in the study of international politics due to poor reporting procedures that do not always clearly indicate what is supposed to be imitated and the prevalence of non-experimental situations that do not permit precise study of variations in settings. By actually attempting a replication, this thesis shall give an example of barriers against this particular avenue of scientific accumulation that may be encountered in practice.

The second strand in the rationale for this thesis is the tremendous growth of the event data movement (see Rosenau, 1973: 54; Peterson, 1975: 266). Within the event data movement, there is consistent controversy over measurement. Are there compatible event data sets? Is there sufficient degree of similarity between event

0

4

typologies and coding rules to permit the recoding of data of one set into another (Leng, 1975: 1)? What are the consequences of different sources? Do different sources yield different event data (see Sigler, 1972; Burrowes, 1974)? What are implications of different scaling procedures (Haevner and Peterson, 1975)? Do nominal scales produce different results than interval/ratio scales? What is to be done with missing data in event data (Deber, 1975)? What are the consequences of treating missing data with interpolation and extended time periods? The first two controversies definitely arise in replication because the student must know if the replicatory data have similar enough structure to the original to warrant use and must also be prepared to compare the original and replicatory event data sets that may have different sources in order to help establish the variation between settings. Scaling procedures, a constant concern of event data analysts, are especially important when the replication involves assigning scale values to originally nominal data and potential distortion must be checked by comparing the scale and nominal values on the basis of an a priori measurement model. Finally, treatment of missing data, like scaling procedures, only enters as a consideration with specific replications.

Finally, the increasing popularity of time-series regression in international politics must be noted. This is seen in such studies as Paul Smoker (1969), Nazli Choucri (1974), Choucri and Robert C. North (1972; 1974), John O. Field et al. (1972); Tanter (1972; 1974), Tanter and William C. Potter (1973), Edward E. Azar et al. (1975), and James M. McCormick (1975). All of these studies display varying degrees of sophistication in estimation of the

time-series—Choucrist and North (1974) probably being the most sophisticated—with more recent studies showing more acumen.

Since students of international politics are relatively new to the estimation of time-series, however, many dubious practices appear.

Many of these dubious practices appear in statistical inference, where students of international politics violate important assumptions of the statistical model and fail to properly specify and test their null hypotheses. . This, of course, results in Type-I error—failing to accept a true null hypothesis and consequently falsely inferring that the alternative hypothesis is true—which may lead to invalid findings and false predictions. Increased awareness and avoidance of the pitfalls of Type-I error in future time-series analyses coincide with the thrust of replication in increasing the confidence that may be placed upon results. Not only do replication and time-series analysis coincide, but there also is a fortuitous link between time-series and event data. All of the studies that were cited above, which use time-series analysis with varying degrees of sophistication, employed event data in whole or sometimes in part (see Azar et al., 1975 for an example of the former and Choucrist and North, 1974 for an example of the latter).

With regard to time-series estimation and its relationship to the other elements of the rationale for this thesis, Tanter (1974) is a typical example. The autoregressive crisis model—an alliance's present conflictive behavior is an additive function of that alliance's previous conflictive behavior and its opponent's previous conflictive behavior—uses event data only. In proposing his model, Tanter does

not specify as fully as possible what the values of the independent variables ought to be. He treats the assumption of no autocorrelation (independence of error terms over time [see ch. 2])—which is exacerbated in autoregressive models—incompletely, and fails to test certain hypotheses properly. Attempting to rectify these difficulties contributes toward better inference. Such rectification does not waste resources that could be better spent on other models. Tanter (1974) is the most extensive of a set of three studies conducted by Tanter (1972) and Tanter and Potter (1973)⁴ that employed the same autoregressive model. Moreover, the Tanter studies inspired McCormick's (1975) analysis of Arab-Israeli conflicts that also employed the same autoregressive model. At a more general level, the autoregressive model is popular elsewhere in studies of international politics (see Choucri and North, 1974) and elsewhere in political science as well (see Davis et al., 1969). Choosing Tanter (1974), then, is based upon a decision to select a model that displays typical characteristics of research in international politics and that is also generally popular in political science.

So far it has been shown that replication is a feasible approach to the scientific accumulation issue. Replication also involves consideration of at least two well-known event data measurement controversies. At the same time, increasingly sophisticated use of time-series analysis reflects a concern for improved inference that is also shared by replication. Moreover, increased use of event data and time-series analysis, often in combination, marks current research in international politics. Here, Tanter's (1974) auto-

7

regressive crisis model reflects the increasing use of event and time-series analysis as well as the general popularity of autoregressive models. Replication, the event data movement, and time-series analysis are closely connected. The rationale for this thesis, then, rests upon a set of interrelated practical methodological problems that are also connected to a theoretical methodological issue.

Organization

The organization of this thesis displays its rationale in the five remaining sections of this thesis: four chapters and a conclusion. The first chapter evaluates the scientific accumulation issue and the role of replication in ameliorating it. First, the origins and nature of the scientific accumulation issue are depicted. The role of replication in ameliorating the scientific accumulation issue by improving inference is then investigated. Next, the small fund of replicatory experience in the study of international politics is examined. Finally, replicatory practices are suggested on the basis of the definition that was provided above and the examination of the replicatory experience.

The next chapter analyzes Tanter's (1974) autoregressive crisis model. First, the general importance of Tanter's model (already briefly indicated above) as part of an ongoing specification of international political processes and as a parallel to specifications elsewhere in political science is described. The actual specification of Tanter's model is analyzed next and is found wanting due to failure to indicate fully what values the independent variables

are supposed to have in various situations and not providing hypotheses for certain obvious situations. Data-generation is then discussed. Here, a relatively clear picture of Tanter's data generation procedure is developed but certain arbitrary aspects of the data generation procedure are questioned. In addition to data generation, the similarities between CADIC event data and East-West event data are established in order to determine if there is the possibility of generating data similar to East-West event data for replication. The last part of chapter 2 is devoted to Tanter's estimation of his model. Once again, Tanter is found wanting due to dubious treatment of possible autocorrelation, the failure to test certain hypotheses, and not employing statistical inference properly in testing other hypotheses. As a result of Tanter's estimation procedures, the greatest barrier of all to replication is erected with certain results required by the model being unavailable and other results unable to be confidently accepted.

Chapter 3 is concerned with the generation of the replicatory data set and the reconstruction of Tanter's data for further analysis in the chapter. First, the precise technique of reordering CADIC event data into a resemblance of East-West event data and the assignment of scale values to the reordered CADIC event data is discussed. Here, possible deviations between the reordered CADIC event data and East-West event data are noted so that possible difference in measurement systems is not left undiscovered. Next, reconstruction of Tanter's data, necessary because East-West event

data are unavailable, from his data appendix is undertaken and possible deviations from what may have been Tanter's data are noted so that there is no misapprehension regarding the similarity of East-West event data in Tanter (1974) and this replication. Following generation of the data, the effects of various transformations used in the course of data generation, which tend to eliminate missing event data, are investigated for each data set. Not only are the scale values actually used considered, but the frequencies of the events are also investigated in order to provide a further validation of the effects found with the scale values. Possible distortion of CADIC event data due to the assignment of scale values is examined next by seeing if scale values diverge from the known measurements provided by the frequencies. Finally, CADIC event data and East-West event data are compared in order to determine if the settings are different. Here, despite comparing the data on the appropriate combinations of frequencies or scale values with sample space, actor, transformations, and regression algorithm, no similarities are found which immediately argues against stability of results.

The final chapter actually compares results obtained from the analysis of the CADIC event data with the results obtained by Tanter. First, of course, the CADIC data are analyzed. Here, recognizing that Tanter did not properly estimate his data, four analyses—ranging from no correction for autocorrelation with no test of hypotheses to correction for autocorrelation where deemed necessary with test of hypotheses—appear. These four separate

estimates are necessary if the possible impact of procedure upon the replicability of Tanter's results is to be understood. Unfortunately, despite the best efforts to rescue some replicability with four different approaches, similar results are not forthcoming.

The conclusion has several different directions. First, the substantive findings are briefly discussed. Proceeding next to a more general level, the autoregressive crisis model is discussed. Here, an alternative specification is offered. At a higher level of theoretical interest, the practical problems of event data measurement and time-series analysis are explored. Here, the implications of the impact of recoding event data sets, the use of transformations to eradicate the problem of missing data, and the relationship of frequencies to scale values upon the use of event data in similar situations is analyzed. Finally, the most general theoretical level is reached—barriers to replication and the scientific accumulation issue. From the outset here, it is noted that Tanter (1974) provides a typical, not a special, example of the barriers to replication. Here, the problems of poor reporting and failure to employ properly statistical inference are noted and suggestions for their rectification are offered. The need for more replication is also noted. Two problems loom large here: the unpopularity of replications as scholarly publications, and the danger that extensive replication may tend to indicate that many results are unstable. Therefore, potential encouragement of replication through a standardization of replicatory procedure that could be easily adapted to the research note is offered. The

danger of massive disconfirmation, it is argued, may be a blessing in disguise by clearing away a lot of the profuse growth that has impeded the establishment of a more coherent body of knowledge of international politics.

The organization of this thesis follows its rationale by proceeding from discussion of replication to analyzing Tanter's (1974) autoregressive crisis model, which demands examination of how Tanter estimated his models, to comparing the settings, which involves investigation of event data measurement controversies, to the actual comparison of the replicatory and original results under a set of estimation procedures and hypothesis testing techniques. Despite the apparent diversity of the topics covered in the course of this thesis, the desire to replicate as reasonably accurately as possible is a constant consideration.

Summary

This introduction has given a brief description of the role of replication in ameliorating the scientific accumulation issue as an example of practical, instead of theoretical, methodology. Thereafter the close relationship between replication and measurement controversies in the event data movement was noted, and the manner in which replication and the increasingly sophisticated use of time-series analysis both improve inference was depicted. Based upon these and other considerations, it was decided to replicate Tanter's (1974) autoregressive model for the 1961 Berlin crisis in order to show what the barriers were to replication in the study of international politics. This would involve consideration of the

estimation of time-series and certain event data measurement controversies. Finally, from these steps, the research design and certain leading results of this thesis were outlined.

Notes

¹Preoccupation with searching for signs of scientific accumulation may be seen in the creation of two workshops upon the topic at Ojai Valley in 1973 and Santa Monica in 1974 (see Rosenau, 1973 for more details).

²An autoregressive model involves a lagged endogenous variable. In the simplest case:

$$\hat{Y}_t = a + Y_{t-1} + u;$$

where $\hat{Y}_t$ = the estimate of Y at time t,
 a = the constant,
 Y_{t-1} = Y lagged once (the lagged endogenous variable),
 and u = the error term.

Tanter's (1974) autoregressive model, distinct from the model involving perception, is seen in the following example:

$$\hat{Y}_t = a + Y_{t-n} + X_{t-n} + e;$$

where $\hat{Y}_t$ = the estimate of Warsaw Pact conflict behavior at time t,
 a = the constant,
 Y_{t-n} = Warsaw Pact conflict behavior with a lag up to three,
 X_{t-n} = North Atlantic Treaty Organization conflict behavior with a lag up to three,
 and e = the error term.

³Although East-West event data did not originate in 1974, I am confident that the 1974 version of the event typology is the same that Tanter used because the 1974 event and resource are typologies are exactly the same as those found in earlier papers, employed by Tanter, that appear in Philip M. Burgess and Raymond W. Lawton (1972: 33-34). Presumably, then, all other information regarding East-West should be the same as Tanter's.

⁴Readers of the three Tanter studies will notice that although the 1973 study refers to the 1974 as forthcoming, the 1974 study makes

no note of the 1973. According to Tanter (letter to Professor Sigler [March 1977]) the 1973 study was written after the 1974, but, due to publication schedules, the 1974 study was published after the 1973. Such potential confusion could be avoided by clear indication of the date of receipt of manuscript for all published works. !

CHAPTER 1

REPLICATION AND AMELIORATION OF THE CUMULATION

PROBLEM: IMPROVING INFERENCE

Overview

This chapter establishes the desirability of replication in international politics by showing how it ameliorates the scientific accumulation issue by providing a firm base for inference. Unfortunately, many students of international politics, despite acknowledging the desirability of replication, leave their attempted replications incomplete. The consequences of this situation are discussed following the presentation of the desirability of replication. Finally, after showing the potential role and current practice of replication in international politics, an explicit replicatory procedure—expanded from the brief characterization given in the introduction—is developed.

Cumulation and replication

The cumulation problem in international politics refers to a situation where a coherent body of knowledge, depending on past efforts, has not been attained. Here, students of international politics cannot respond positively to such questions as:

What concrete findings, if any, have been uncovered that have become the basis of widespread intersubjective consensuses in the field? What theoretical models and/or perspectives have been rendered obsolete and what new ones have been generated that serve as organizing paradigms for

the field? How reliable and generalizable is the knowledge resulting from our systematic inquiries? What new methodologies, with what potentialities and limitations, have been fostered by the research programs?...How widespread has been the diffusion and scholarly utilization of the research [Rosenau, 1973: 12]?¹

International politics appears to many students as a "fragmented mass of bits and pieces, of correlation coefficients, amounts of variances explained, significances, factors, etc., that cannot or have not been put together in a meaningful way to tell us much about international behavior" (Zinnes, 1974:1). Thus, one of the symptoms of the cumulation problem is that students of international politics cannot confidently make valid claims about international behavior.

There are numerous suggested cures for the cumulation problem, ranging from Zinnes' (1974:10-14; 1976:211-216) demand for more formal deductive theory to Vincent's (1974:9-11) call for use of explicit overreaching frameworks. Boynton (1974:1-6) proposes a shift in emphasis from the present iconoclastic style to emphasize close links with past research. Certainly, all of these suggestions merit attention. Students of international politics will benefit from deductive theories and frameworks. They are not likely to create a coherent body of knowledge by iconoclasm. Zinnes' demand for deductive theory, despite her strong presentation of the possibilities of logical development and evaluation, probably does not appeal to the vast majority of students of international politics, who engage in "ad hoc hypothesis testing" where "beyond a very general argument the researcher does not explain why the particular relationship is being examined, or why he believes such a pattern would occur."

(Zinnes, 1976:4-5). The possibility of an overarching framework in the wake of the collapse of general systems theory (McClelland, 1974:29) seems remote because "there can be no paradigm of international politics, but of conflict, consensus, coalition, decision-making, etc. for which a specific range of behaviour dealing with international politics might provide a delimiting set of cases" (Sigler, 1974:2). Many students of international politics find deductive theory too restrictive and are uneasy at the prospect of forays into general social science theory in the name of some overarching framework. Despite misgivings, students of international politics seem best acquainted with "middle-range theory" and determined to retain it.

In the absence, then, of formal theory that leads to deductions and of overarching frameworks that develop generalization across levels-of-analysis, students of international politics must somehow develop inferences for their own respective "islands of theory." Clearly, the pattern of inference comes from the behavioral research design. Jack E. Vincent (1974:1) gives a rather crude version of this view:

To me, scientific accumulation is facilitated when a verified proposition about reality is recorded and is easily assessed...For me, statistical test, "implying an operationalization of variables is a "good" method. When I encounter studies using this method, I view the results with confidence, since the method presents the likely odds of being right or wrong...scientific accumulation occurs when a statistical test or procedures are applied to the data selected.

Here, accepting the deductive mathematical framework of statistics and a set of procedures generally agreed upon within the community of scholars regarding use of statistics, students of

international politics make inferences about international political phenomena. In following this procedure, students of international politics scrupulously avoid asserting that insignificant findings are significant as occurs in Type-I error. Most students of international politics stop at this point and do not wonder if "the same results must occur in future tests? (Vincent, 1974:1). As Vincent (1974:1-3) goes on to suggest (before pleading for overarching frameworks), replication tests findings for Type-I error by seeing if they are still stable under different conditions. Replication enlarges the sample space used to demonstrate propositions and therefore permits greater confidence in inferences about international political phenomena.

Replication is a natural extension of statistical generalization. There are, however, additional payoffs for scientific procedure and the building of cumulative knowledge that come from replication. Measurement and more particularly operationalization issues would be improved. Would-be replicators may use or develop operationalizations differing from those in the original study and may relate differences between sets of results to different operationalizations. Although such investigations of different operationalizations is best accomplished when only operationalizations and no other conditions have been changed, informed judgments may be reached with operationalizations and other conditions varying from the original study. In addition to "technical" spinoffs, replication has implications for other parts of the cumulation problem. Increased replication forces students of international politics to assess more closely past research.

Rash repudiations of past findings, typical of iconoclasm, would become rarer and a "post orientation" in research that leads to cumulation (Boynton, 1974:3-6) would be reinforced. Assuming that a set of stable findings is indeed built upon an "island of theory," these findings may well be accepted as axioms for formal models, or they may be used as suggestive isomorphisms for an overarching framework in international politics; political science, or the social sciences. Replication extends to other areas than inference in ameliorating the problems of inadequate cumulative knowledge.

There are two basic roles for replication in correcting the problem. First, replication carries forward the "technical" issues of inference and measurement that profoundly affect the acceptance of findings. Second, replication does not adversely affect other cumulation-provoking strategies but may actually hasten their eventual employment. Many students of international politics, however, have not gained these benefits of replication due to current malpractice in following scientific replication procedures.

Current practice of replication

Most students of international politics are aware that replication exists and that it would somehow enhance their work. The most commonplace attitude is, 'Of course replication is important; it's part of scientific procedure, isn't it?' Patrick J. McGowan (1974:54-55) clearly displays this attitude:

It is generally accepted among social scientists that the most reliable means for validating facts is through replication. Thus, one way to compare is to ask similar questions....Another is to repeat someone else's questions

on a new situation or in the same state at a later point in time.

Here, a very broad and vague description of replication is evident.

In fact, whether the first procedure is replication is uncertain: does asking similar questions mean the same hypotheses as before or questions about the same dependent variable? Moreover, aside from briefly invoking comparison, no indication is given how replication operates in practice—that differences between the original and replicatory study must be directly considered in arguing for generalizability. This particular aspect of replication, is usually forgotten by most students of international politics and has a deleterious impact upon the acceptability of their findings. Three replicatory studies² can be briefly examined to show how the current low status of replication actually appears in practice.

The first replication study, Raymond Tanter's (1965) well-known imitation of Rudolph J. Rummel's (1963) study, is an exemplar of replication. Before actually seeing whether his results and Rummel's were similar, Tanter carefully noted that concepts and operationalizations were the same as Rummel's and showed that the two sets of data were highly related which indicated that there were no differences that could not be accounted for. Thus, when similar regression and factor analytic results were discovered, Tanter was able to claim on the basis of no accountable differences that conflict behavior within and between nations was stable. Since similar conditions produced similar results, other students of international politics could give great credence to Tanter's findings.

A far less credible replication was recently published by Rosenau and Ramsey (1975). The replication was based upon a previous examination of an eight-fold nation attribute conflict and cooperation typology conducted by Rosenau and Hoggard (1974) that employed World Event/Interaction Survey (hereafter referred to as WEIS) event data. In the replication, the typology and hypotheses remained unmodified; the Conflict and Peace Data Bank (hereafter referred to as COPDAB) and Comparative Research on the Events of Nations (hereafter referred to as CREON) event data sets (see Burgess and Lawton, 1972:28-31, 37-40 and 45-49 for respective descriptions of WEIS, COPDAB and CREON) were used. Although similar results to the 1974 study were obtained, Rosenau's and Ramsay's failure to look for differences between the data sets, actually quite probable given problems of compatability, and to not see if the differences were reflected in the results of the replication do not bode well for generalizability. Are dissimilar data leading to similar results or are similar data leading to similar results? Other students of international politics—well aware of the potential difference between data sets due to the compatability problem—cannot give ready credence to the results of this replication.

Whereas the Rosenau-Ramsey (1975) replication could not be given much credence, the replication of Brams international hierarchy studies of the structure of the international political system by Kegley, and Wittkopf (1976) becomes extremely questionable on the basis of a typical replicatory fault. Brams' studies employed data upon international meetings that he derived from the New York

Times Index. Kegley and Wittkopf (1976:268) admit that they are using WEIS. Aside from a different operationalization of international meetings³ Kegley and Wittkopf overlook the known differences between editions of the New York Times (Sigler, 1972:25-26) that probably extend to the New York Times Index as sources of events. By failing to compare the two data sets, Kegley and Wittkopf not only detract from generalizability but also render their entire effort questionable by ignoring one of the most celebrated event data measurement problems. An opportunity to investigate the actual empirical results of a known measurement problem—one of the spinoffs of replication—was left untouched.

Failing to argue for the generalizability on the lack of differences between the original and replicatory study is commonplace. Usually, as in the case of Rosehau-Ramsey (1975), the failure detracts from the acceptability of the replication. On other occasions, the failure casts aspersions upon the replication as a study in its own right, turning a potentially fruitful exercise into frustration. Replicatory procedure, given the constant appearance of the failure, must emphasize how the differences between the original and replication one to be accounted for.

Replicatory procedure

From the foregoing discussion, the benefits of arguing for generalizability on the basis of differences between the original and replicatory study may be seen. But how is the would-be replicator to appropriately assess differences between the original and replicatory study, and once the differences between the original and

replicatory studies are established, how ought the argument for generalizability proceed? These two questions are represented in theory by what Fletcher (1970) has called the "replicability quotient" and "generalizability quotient."

The first question asks how the variation between the original and replicatory studies comes into being. Here, given a replicatory design, the analyst must assess how different replicatory measures are from those of the original. This is a specific existence of Fletcher's (1970:52) "replicability quotient," where "the quotient is dependent upon the number and severity of changes that would be necessary wherever the study was replicated."

Obviously, the would-be replicator's capability to unravel the original analyst's logic-in-use (Kaplan, 1964:ch.1) is at stake here. Can concepts and procedures be understood and which of them may be applied in the replication? These two items establish critical distance between the replication and the original study. The extension of this critical distance and consideration of different scope conditions leads to asking if "The extent of contextual changes vary in severity and affect the degree to which the researcher strains the applicability, the generalizability, of the original thesis (Fletcher, 1970:52)"—the "generalizability quotient." If the changes made in the replication relative to the original are not too severe, then the researcher may claim that the replication is a valid extension of the original. Generalizability is the claim that the juxtaposition of the findings of the original and replication are logically connected, that they do not respectively

represent so greatly different situations that any similarities are purely fortuitous. The first step in the argument for generalizability then, is to establish potential differences between the replication and original study and to see if the potential for valid generalization exists.

After having established potential generalizability, the researcher then enters upon the actual empirical procedure of replication. As noted before, generalizability at this stage depends upon the actual measurable differences between the conditions of the replication and original study. Measurement of the differences may involve several items. Difference in scope conditions may be considered. Does the replication represent a different system than that of the original? If the systems are in fact different, then differences may appear in the findings. The comparison of the measures used in the replication to those in the original study is the crux of establishing the measured differences. Here, simply looking for simple positive variation between parts of the two data sets is inadequate. If, for instance, it is known that the systems are different and the two sets of measurements are positively related then similarity between the two studies is not established. Similarly, if the replication employs a more fundamental measurement model than the original and the findings show no similarity, the variation is not yet accounted for. The replicator must note carefully all accountable differences, including differences in scope conditions as well as measurement models.

After establishing the source of variation between the two

studies, the researcher must analyze and compare the replicatory data with the original. Decisions and frustrations are rife here. With reference to analysis, the research must decide if the same procedures, as in the original, even if inappropriate, are to be used, or if the original procedure and the improved version should be employed, or if re-analysis of some parts of the original data should be attempted. In the first case, the researcher must choose between appropriate procedure and following the at times erroneous path of the original. The second and third cases involve more work but may shed extra light upon the consequences of various procedures. When comparing results, after deciding how those in the replication are to be obtained, the researcher must trust the results reported in the original study. Frequently, however, vital information, such as standard errors, that would improve the acceptability of results is missing. Thus, the researcher may well be in the unenviable situation of having to accept and use results that are definitely prone to Type I error—a most frustrating situation.

In comparing variation and similarity between studies four possible situations may be expected. First, there are no known sources of variation and similar results are obtained. Here, the researcher may generalize. Second, there is no known source of variation but results are dissimilar. Here, the researcher may not generalize. Third, the sources of variation are known and results are similar. Here, the researcher does not make any generalization statement but examines the sensitivity of data (measurement error) and design (bad research design). Fourth, and most frequently,

the sources of variation are known and results are dissimilar. Here, the researcher may make generalizability statements as in the first and second cases respectively, depending upon the degree of belief in whether variation accounts adequately or inadequately for the similarity or dissimilarity of the results (Galtung, 1969:66-67). As may be expected, the decision-fraught replication procedure only yields two clear-cut final decisions.

Replicatory procedure, then, is obviously complex. First, the potential for generalizability must be established. Following the establishment of potential, realization is an arduous process. Establishing variation is not straightforward but requires definite interpretation of scope conditions and consideration of measurement models. Analyzing replicatory data and comparing results have several strategies with their own payoffs and frustrations. Finally, as just demonstrated, making generalization statements, does not necessarily proceed smoothly.

Summary

This chapter has shown what the ~~potential~~ role of replication is, how replication is usually mishandled, and what a replicatory procedure would entail. Replication may aid in building cumulative knowledge, but replications are frequently done poorly so that cumulative benefits are lost. As a response to the misapplication of replication, a replicatory procedure has been established wherein criticism, measurement estimation theory, and the logical grounds for generalization—all contributory to cumulation—are combined. The next chapter starts the execution of the replicatory procedure

outlined here by determining the potential generalizability of Tanter's (1972) autoregressive crisis model.

Notes

¹This is from the text of a letter sent by Rosenau to participants at Workshop on the Successes and Failures of Scientific International Relations Research, dated 12 January 1973.

²These studies are not part of a longer study of the lack of replication current in international politics. I assume that international politics, like many other social sciences, always needs more replication.

³The different operationalizations would, of course, also contribute to Kegley's and Witkopf's woes.

CHAPTER 2

ESTABLISHING REPLICABILITY: ON TANTER'S AUTOREGRESSIVE CRISIS MODEL

Overview

The preceding chapter establishes the replicability of Tanter's autoregressive crisis model. Since the system analyzed in the replication, the 1961 Berlin crisis, is the same as one of those analyzed by Tanter (the other being the 1948 Berlin crisis), there is no need to debate the applicability of the model to the system. Immediate attention may therefore be given to the model in general, to Tanter's actual portrayal of the model, and to potential deviations from Tanter's technique involved in replication. By noting where and how severe these deviations are, the degree of replicability is established. The first step in this procedure is to demonstrate what the model is as Tanter describes it and to show the similarity of the model to others in international politics and elsewhere in political science. The general worthiness of the model for further investigation for its own sake and for the sake of the general class of models of which it is a part is thereby evident. Next, Tanter's specification of the model is analyzed: what hypotheses are proposed and how well do they match the development of the model? Following Tanter's specification, data-generation is dealt with. Here, not only must the actual way in which variables are

created be scrutinized, but the face-value similarity of CADIC data set to the East-West data set must also be examined in order to assess whether any grounds exist for claiming that the two event data sets are incompatible. Fourth, Tanter's analysis and results are reviewed: are estimation procedures reasonable? Are results subjected to tests of statistical significance so that their reliability is known? Lacunae in the execution of research design that have to be taken into account in the remainder of this thesis may well appear here. Finally, from the foregoing, an overall summation of the interpretation of Tanter's efforts must be given. Is there enough information and guidance with which to carry out the replication with some degree of confidence so that it is not a charade?

The model

Tanter's autoregressive crisis model attempts to integrate two approaches to international politics; the international interaction approach and the foreign policy approach. The first approach follows Charles A. McClelland's focus on the outputs of national systems. The second approach looks at McClelland's interacting nation-states as an organization where "if a nation performs an action of a certain type today, its organizational components must yesterday have been performing (or have established routines for performing) an action only marginally different from that action" (Allison, 1971:87). McClelland's approach views international politics as a system of reactions of actors to others, while Allison's approach—greatly indebted to organizational theorists such as

Herbert A. Simon (1957)—sees international politics generated by a simple set of decision rules that is evident in the maintenance of past behavior. Here, two usually separate approaches (Hermann; 1972:10) are to be combined, thus aiding scientific cumulation, into a model that claims that international political behavior is a function of interactions with an opponent and an actor's own past behavior toward an opponent. The actual form of the model must take into account the extent to which an actor reacts to its opponent and to its own past behavior. Tanter's formulation is straightforward:

$$\hat{Y}_t = a_y + Y_{t-n} + X_{t-n} + u_y;$$

where; a_y is the constant

$\hat{Y}_t$ is the actor's present estimated behavior;

Y_{t-n} is the actor's behavior in the

past up to a lag of three periods

(the organizational parameter;)

X_{t-n} is the opponent's behavior in the

past up to a lag of three periods

(the interaction parameter).

u_y is the error term.

The two parameters change their relative strength as the crisis goes through its phases. Two other models, inspired by the Stanford 1914 studies, were also analyzed by Tanter with respect to the 1948 and 1961 Berlin crisis, separately and combined, for the North Atlantic Treaty Organization (hereafter referred to as NATO) and the Warsaw Treaty Organization (hereafter

referred to as WTO) but shall not be discussed here.¹ The autoregressive model given above shall only be attempted for the 1961 Berlin crisis. Indeed, the autoregressive model is interesting enough to be investigated in its own right.

Tanter's autoregressive crisis model is not unique, appearing frequently international politics and elsewhere in political science. The most striking example of an autoregressive model in political science is given by Otto A. Davis, M. A. H. Dempster, and Aaron Wildavsky (1969) based upon Wildavsky's The Politics of the Budgetary Process (1964). This is one of the bases for Allison's formulation that Tanter has incorporated. The authors analyzed non-defense expenditure requests in the United States from 1947 to 1963. Proposing that "the agency request (through the Budget Bureau) [percentage of funds available requested] is a fixed mean percentage of the agency's request for the previous year plus a random variable (stochastic disturbance)" (1969:469), the authors specify the simplest of autoregressive models:

$$X = a_x + X_{t-1} + u_x;$$

where a_x is the constant

X_t is present estimated request,

X_{t-1} is previous year's request,

u_x is the error term.

We are not told how well the equation was specified but are informed (Davis et al., 1969:474) that the explained variance for all equations is high.

In international politics, Tanter's autoregressive model

appears of course in Tanter (1972) and Tanter and William C. Potter (1973). James M. McCormick (1975) imitated Tanter's work for the 1956 and 1957 Arab-Israeli conflicts, generally confirming Tanter's results.² Autoregressive models are not limited to Tanter and his imitators. Nazli Choucri and Robert North, (1975:169) in what is still the most sophisticated use of time-series in international politics, claim that military expenditures by European powers from 1870 to 1914 are a function of:

$$\begin{aligned} \text{Military Expenditures} = & a_3 + b_{10} (\text{military expenditures}_{t-1}) \\ & + b_{11} (\text{military expenditures of} \\ & \quad \text{non allies}) \\ & + b_{12} (\text{intensity of intersections}) \\ & + b_{13} (\text{colonial areas}) \\ & + b_{14} (\text{population-times-national-} \\ & \quad \text{income}_t) \\ & + u_3; \end{aligned}$$

where a_3 is the constant, intensity of intersections is the highest amount of violence for coloneal conflicts by the actor, and u_3 is the error term.

Here, in complete contrast to Davis et. al. (1969), defense expenditures are analyzed. Similar to Tanter (1974), there are more than two regressors one of which is a lagged endogenous variable and one of the others equivalent opponent's behavior to the lagged endogenous variable. Results for this equation yield explained variances ranging from .58 to .95 with the lagged endogenous variable positive and significant at least at the .05 level-of-risk, while military expenditures of non allies have mixed signs but is significant in a 2-tail test at least at the .025 level of risk for the 1870 to

1914 period.

Tanter's autoregressive crisis model is a synthesis of two important strands of investigation in international politics. The presence of the lagged endogenous variable in a regression equation is commonplace; it appears in the literature in public policy-organization analysis to which Allison and Tanter are indebted. Thus, Tanter's autoregressive crisis model, is part of a class of models that have intrigued political scientists. Replication of a particular instance of this model simply reconfirms belief in the importance of the model.

Specification

During the various phases of an international crisis, the regressors in Tanter's autoregressive model will attain different strengths relative to each other (parameter dominance).

The first and most general hypothesis is simply a statement of the model: "Current alliance intensity [of conflictual behavior] is a result of prior intensity of an adversary (event/interaction) as well as an alliance's own prior intensity." (Tanter, 1974:138). No further specification is given. Apparently, the model is correctly specified if in $Y_t = a_y + Y_{t-n} + X_{t-n} + u_y$, the two regressors do not equal zero. This is unfortunate because it is, of course, quite possible for both regressors to be less than zero and significant in which case the analyst is in the uncomfortable situation of explaining how present alliance behavior is instigated by previous behaviors that do not in any way match present behaviors and indicate

some continuity in crisis escalation and deescalation. By failing to elaborate upon the sign of the regressors for the entire crisis, Tanter casts doubt upon his theory. .

The second hypothesis refers to Tanter's historical expectations, based upon primarily Western and non-revisionist, sources;³

where "In contrast to the acknowledged inconsistency and indecisive action in Berlin, Soviet behavior generally is portrayed as clear, purposeful, consistent, and thoroughly planned" (Tanter, 1974:49).

If the USSR is the sole dominant member of WTO which is far more routinized than NATO, then, "For WTO, estimates of organizational processes should be larger than estimates of event/interaction, irrespective of conflict in question and particular phase" (Tanter, 1974:143). This hypothesis cannot be tested with the auto-

regressive model because we are asked to take alliances into account.

Hence, the two ordinary parameters plus the organizational parameter dummied by alliance must be regressed upon the pooled alliance data

in order to see if alliance is affecting the organizational parameter's size for the autoregressive model. No equivalent hypothesis regarding

interaction, indicative of the reactivity associated with NATO, is presented.⁴

Thus, not all potential hypotheses regarding historical expectation are presented, nor may the sole hypothesis presented be tested through the autoregressive model.

The next set of hypotheses are developed with respect to phase. According to Tanter (1974:28-31), drawing upon learning theory, game theory, organization theory, and theory of the firm,

the autoregressive—organizational procedure—parameter, will be dominant during normal times and the interaction parameter will prevail during a crisis. In order to test this proposition, Tanter creates a three-phase historical typology and a two-phase empirical typology. The three-phase typology, based on Tanter's historical sources, starts at 4 June 1961 and goes to 24 July 1961 for precrisis, 25 July 1961 to 12 September 1961 for the crisis, and 13 September 1961 to 17 October 1961 for postcrisis. Presumably precrisis and postcrisis are the normal phases (Tanter, 1974:85;87). The empirical typology was formed by dividing the total time period into an intensification phase that goes up to the period of highest conflictual behavior in the system (14 August) and the normal reduction phase which occurs from the peak onward (Tanter, 1974:88;89). This typology is completely empirical and it may not be precisely reproduced with another event data set. Tanter's expectations for the relative dominance of the parameters is summarized in Table 2-1.

This table represents the following hypothesis:

In the pre-and postcrisis phases, estimates of organizational processes should be larger than estimates of event/interaction irrespective of particular alliance and conflict in question.

In the crisis phase, estimates of event/interaction should be larger than estimates of organizational processes, irrespective of the particular alliance and conflict in question.

In the intensification phase, estimates of event/interaction should be larger than estimates of organizational processes, irrespective of the particular alliance and conflict in question.

In the reduction phase, estimates of organizational processes should be larger than estimates of event/interaction, irrespective of alliance and conflict in question. [Tanter, 1974:73].

TABLE 2-1—Tanter's expectations of relative phase dominance of parameters

	Interaction	Organizational Process
Historical Three Phase		
Precrisis	low	high
Crisis	high	low
Postcrisis	low	high
Empirical Two Phase		
Intensification	high	low
Reduction	low	high

Source: Tanter (1974:99).

All of these hypotheses follow from the logic of Tanter's expectations and are easily testable by seeing if the parameters that appear to have the appropriate relationship are significantly different by utilizing the difference between their respective t-statistics. But, as noted with the general and historical hypothesis, problems of specification and derivation of useful hypotheses arise. As in the case of the general hypotheses, Tanter seems content to accept his results if the regressors do not equal zero. If, however, a particular parameter is truly dominant, then it ought to be greater than zero. Claiming that a dominant parameter is less than zero is tantamount to claiming that the primary cause of the behavior is opposite to the one specified. Potentially missed hypotheses are also present, and under Tanter's conditions would render the replication faulty. No attempt is made to consider the way in which the base level of conflict, represented by the constant, changes. With the appropriate analysis-of-covariance design, it may be determined if the base level of conflict is lower during normal phases. More glaring than failing to consider changes in base level of conflict, is ignoring discontinuity of the regressors across phases; despite the relative dominance of regressors within a phase, if it is shown that the same regressors are equal across phases which are supposed to be different, then statements about regressor dominance are not supported. Once again, analysis-of-covariance would have to be employed. Like the preceding two sets of hypotheses, the third presents definite problems.

Overall, then, Tanter does not seem as willing to utilize

the full resources of the general linear model as some of his reviewers. (Jensen, 1974; Kegley, 1975) seem to believe. Failures to specify the sign of regressor will lead to interpretive difficulties in affirming that processor cause are present. In two instances, overlooking two hypotheses requiring analysis-of-covariance may be a 'mere oversight', having no farther ramifications for the theory. But, with respect to phases, analysis-of-covariance must be used to support the hypotheses. Although the procedures whereby Tanter's intent may be executed are quite clear, execution of these procedures alone and under Tanter's conditions would render the replication faulty. In sum, with severe technical reservations, Tanter's hypotheses and other specifications are acceptable and repeatable.

Data-generation

There are two phases to determining the replicability of Tanter's data-generation. Are data-generation procedures clear and understandable enough to be imitated? If these procedures are, understood, does CADIC provide a basis for similar data?

The event data set employed by Tanter is the Corson East-West set. As shown in Table 2-2, the East-West event data typology contains conflictual and cooperative categories, divided into verbal and physical acts categories, in a manner very similar to the WEIS event typology (Table 2-3) that was revised by Corson. (Burgess and Lawton, 1972:32) Tanter in his autoregressive model employs only conflictual data ignoring cooperative behavior and the macro-category "disapproval." "Disapproval" is excluded since it represents a

TABLE 3
CORSON EVENTS TYPOLOGY: A TYPOLOGY OF ACTION BETWEEN NATIONS

	WORDS (Verbal Actions)	DEEDS (Non-Verbal Actions)
CONFLICT (Negative Hostile Actions)	<p>Actor's Verbal: Evaluation or Perception of Past or Current Action by Target</p> <p>1 "DISAPPROVAL"</p> <p>1140 DENY (deny accusation attributed action policy)</p> <p>1170 ACCUSE (charge, criticize, blame, disapprove, denounce, denigrate)</p> <p>1020 COMMENT NEGATIVE (comment on situation, explain policy undesired by target)</p>	<p>Actor's Non-Verbal: Oppose Directed at Target</p> <p>2 "CONFLICTIVE DEEDS"</p> <p>7220 FORCE (threaten or violent use of military resources, equipment to achieve objectives. Military engagement, non-military destruction, non-injury destruction)</p> <p>7180 DEMONSTRATE (armed force mobilization, movement exercise, display. Non-military demonstration, meeting, boycott or walk-out)</p> <p>7230 INCREASE MILITARY CAPABILITY* (increase military spending, troop levels, develop weapons, authorize military action, reserve call-ups)</p> <p>7240 AID OPPONENT* (give military aid to opponents of target)</p> <p>7260 SUBVERT* (give aid to insurgents in target, make a non-verbal attempt to influence internal politics of target undesired by target)</p> <p>7210 SEIZE (seize position or possessions, detain or arrest personnel)</p> <p>7220 F. WELL (seize organization, group or personnel)</p> <p>7180 REDUCE RELATIONSHIP (cancel or postpone planned event, reduce or halt routine international activity, recall official, halt negotiations, break diplomatic relations)</p> <p>7300 NEGATIVE DEED* (carry out non-verbal non-violent conflict action not covered by other categories)</p>
	<p>Actor's Verbal: Desire or Preference for Future Action by Target</p> <p>3 "DEMANDS"</p> <p>3150 DEMAND (order, command, insist, demand compliance)</p> <p>3130 PROTEST (express formal or informal complaint)</p> <p>3100 PROPOSE NEGATIVE (offer proposal, suggest, urge action, policy undesired by target)</p> <p>3090 REQUEST NEGATIVE (appeal, ask for information and action in context undesirable to target)</p>	<p>Actor's Verbal: Intent Regarding Possible Future Action Toward Target</p> <p>5 "THREATS"</p> <p>5170 THREATEN (intention to take future action undesired by target, with or without specific negative sanctions, a contingent if then statement)</p> <p>5160 WARN (warn future situation or action by target undesired by actor, warn against action by target undesired by actor)</p> <p>5110 REJECT (turn down proposal, protest, demand, threat, refuse, oppose, harden position)</p>
COOPERATION (Positive Friendly Actions)	<p>2 "APPROVAL"</p> <p>2020 COMMENT POSITIVE (comment on situation, explain policy desired by target)</p> <p>2040 APPROVE (praise, hail, applaud, endorse, support)</p>	<p>8 "COOPERATIVE DEEDS"</p> <p>8010 YIELD (surrender, submit, retreat, evacuate)</p> <p>8060 GRANT (end negative sanctions, conflict action, threat of conflict action, express regret, apologize, give invitation, asylum, privilege, diplomatic recognition, release return persons or property)</p> <p>8290 DECREASE MILITARY CAPABILITY* (reduce military spending, troop levels, reduce readiness for military action)</p> <p>8030 CONSULT (make or receive official visit, participate in meeting involving foreign travel for at least one participant)</p> <p>8270 CARRY OUT AGREEMENT* (imply, meet a previous agreement, e.g., a diplomatic, economic, cultural or military agreement)</p> <p>8070 REEMBRU (give economic, technical or military aid)</p> <p>8290 INCREASE RELATIONSHIP* (increase routine international activity without explicit agreement, increase economic, technical, or military aid)</p>
	<p>WORDS (Verbal Actions)</p> <p>4 "PROPOSALS"</p> <p>4090 REQUEST POSITIVE (appeal, ask for information and action in context desirable to target)</p> <p>4100 PROPOSE POSITIVE (offer proposal, suggest, urge action, policy desired by target)</p> <p>4250 NEGOTIATE* (participate in substantive talks, negotiations on specific issue or interest area)</p>	<p>6 "OFFERS"</p> <p>6020 INTEND POSITIVE (explain future policy desired by target)</p> <p>6100 OFFER (offer to take future action desired by target)</p> <p>6050 PROMISE (assure of future action desired by target)</p> <p>6080 ACCEL (accept proposal, agree to meet or negotiate, agree to future action, sign or reach substantive agreement)</p>

SOURCE: Corson (1971: 25)
*Categories not included in the WEIS code

TABLE 2-2—The East-West categories.

Source: Burgess and Lawton (1972: 33).

[30]

TABLE 2
WORLD EVENT INTERACTION SURVEY EVENTS CATEGORIES

1. YIELD	11. REJECT
011 Surrender, yield to order, submit to arrest, etc.	111 Turn down proposal; reject protest demand, threat, etc.
012 Yield position; retreat; evacuate	112 Refuse; oppose; refuse to allow
013 Admit wrongdoing; retract statement	
2. COMMENT	12. ACCUSE
021 Explicit decline to comment	121 Charge; criticize; blame; disapprove
022 Comment on situation-pessimistic	122 Denounce; denigrate; abuse
023 Comment on situation-neutral	
024 Comment on situation-optimistic	13. PROTEST
025 Explain policy or future position	131 Make complaint (not formal)
	132 Make formal complaint or protest
3. CONSULT	14. DENY
031 Meet with; at neutral site; or send note	141 Deny an accusation
032 Visit; go to	142 Deny an attributed policy, action, role, or position
033 Receive visit; host	
4. APPROVE	15. DEMAND
041 Praise, hail, applaud, condolences	150 Issue order or command, insist; demand compliance, etc.
042 Endorse others policy or position; give verbal support	
5. PROMISE	16. WARN
051 Promise own policy support	160 Give warning
052 Promise material support	
053 Promise other future support action	17. THREATEN
054 Assure; reassure	171 Threat without specific negative sanctions
	172 Threat with specific nonmilitary negative sanctions
6. GRANT	173 Threat with force specified
061 Express regret; apologize	174 Ultimatum; threat with negative sanctions and time limit specified
062 Give state invitation	
063 Grant asylum	18. DEMONSTRATE
064 Grant privilege, diplomatic recognition; de facto relations, etc.	181 Nonmilitary demonstration; walk-out on
065 Suspend negative sanctions; truce	182 Armed force mobilization, exercise and/or display
066 Release and/or return persons or property	
7. REWARD	19. REDUCE RELATIONSHIP (as Neg. Sanction)
071 Extend economic aid (for gift and/or loan)	191 Cancel or postpone planned event
072 Extend military assistance	192 Reduce routine international activity; recall officials, etc.
073 Give other assistance	194 Halt negotiations
8. AGREE	195 Break diplomatic relations
081 Make substantive agreement	
082 Agree to future action or procedure; agree to meet, to negotiate	20. EXPEL
9. REQUEST	201 Order personnel out of country
091 Ask for information	202 Expel organization or group
092 Ask for policy assistance	
093 Ask for material assistance	21. SEIZE
094 Request action; call for	211 Seize position or possessions
095 Entreat; plead; appeal to; help me	212 Detain or arrest person(s)
10. PROPOSE	22. FORCE
101 Offer proposal	221 Noninjury destructive act
102 Urge or suggest action or policy	222 Nonmilitary injury-destruction
	223 Military engagement

SOURCE: McClelland and Young (1969: 29)

TABLE 2-3—The WEIS categories.

Source: Burgess and Lawton (1972: 30).

evaluation, e.g. (X says Y is hostile). The evaluation data are used in the other autoregressive model with perception added.

(Tanter, 1974:80). Events when considered for use were those that were directed by members of an alliance toward members of the opposing alliance, except for increases in the micro-category "increase military capability" that were included regardless of target (Tanter, 1974:81). The summed raw daily intensity of these events (ratio scaling in East-West diverges from WEIS nominal scaling) is then divided by the frequency of these events in order to form the basic unit of measurement, "mean action intensity" (Tanter, 1974:81-82)—otherwise known as "conflictive action intensity scores," "action intensity," and "intensity." (Tanter, 1974:43).

So far, the reconstruction of Tanter's logic is quite clear. Two problems, however, appeared: determining who the actors and targets are, and understanding the implications of alliance as the unit-of-analysis. Although Corson does code alliances as actor and target (Burgess and Lawton, 1972:56), such actors and targets are not coded often enough to be the sole criteria for event selection. Obviously, members of the alliances involved in the conflict must also be included. Moreover, if alliance are implicitly eligible, so too are the blocs ("Free World", "Socialist World") that they represent. The sole remaining criterion is whether the activity engaged in was committed with regard to the 1961 Berlin crisis. With Corson, there is a general monitoring of events world-wide from 1945 to 1965 with setting codes provided in order to filter out

events pertaining to a particular situation. Tanter (1974:81) employed the Berlin, East Germany, and West Germany settings for the 1948 crisis—presumably the same criterion prevailed for 1961. A similar arrangement would have to be present if the replicatory event data set also monitors events world wide and does not deal with a specific crisis. Based on this reasoning, it was determined that Tanter used alliance member's behavior as a surrogate for the alliance's behavior and that alliances and blocs are not excluded from consideration.

The problem of using individual alliance members as a surrogate for the alliance may now be dealt with. The individualist fallacy—inferring collective behavior from individuals, ignoring that a whole is more than the sum of its parts—faces us. Even if alliances and blocs do appear, individual nation states are still likely to be the most prevalent actors and targets. In the absence of a measurement of alliance cohesion in the model that could give some indication of how well individual nations partake in overall alliance behavior, a cautious strategy is appropriate. First, as Tanter (1974:81) attempts, an auxiliary theory of alliance cohesion based upon recent international history, is needed. Here, the two alliances in 1961 are seen as quite cohesive and hierarchical. Thus, if the plurality of the events engaged in by an alliance are perpetrated by alliance leaders, we may be certain that the events are those of the alliance. Finally, presentation of results must emphasize that they are based upon a surrogate measurement and must therefore be accepted with qualifications.

The daily mean intensity scores are not used by Tanter in his autoregressive model. Instead, these scores are subjected to two successive transformations that raise farther questions about data generation. Tanter is dissatisfied with the pattern presented by his scores. Smoothing—useful for covering up days with no events—is employed. Tanter justifies this smoothing by invoking the idea of memory (1944:99) The applicability of memory as a concept for alliances is at best uninvestigated. The function used to smooth is an exponential decay function:

$$f(X_t) = \sum M^{(n-t)} X_{t-n};$$

where X_t is current mean alliance intensity,

where M_t is the forgetting rate set at .5,

where n is thirteen days [Tanter, 1974:100]

We are not told why .5 and thirteen are used; a replicator would have to employ them without question. Moreover, the effect of the smoothing vis-a-vis the untransformed scores is uninvestigated; Tanter may well be employing a transformation that will yield different results for alliances in the various phases.

The next transformation is aggregating the scores every three days (Tanter, 1974:90) This logically follows smoothing since days were referred to in this procedure. This is easily done. No reason is given for the three day

period; again, the replication must simply follow orders.

Although Tanter may be correct in suggesting that the three day aggregation period may disguise certain measurement errors, there is no investigation of how aggregation affects the scores.

With elucidation of Tanter's critical variable creation process and the pitfalls attendant upon it, sufficient apparent similarity of CADIC's event typology to East-West's must now be shown. Looking at Table 2-4, the similarity between CADIC, WEIS, and East-West is readily apparent. Not all of the categories are the same but like East-West, there is a distinction between physical acts, threats, and other verbal acts exist. The precise combinations of keywords in CADIC are not the same as the micro-categories in East-West. This is especially evident with CADIC's "Condition Negatively" and East-West's "Disapproval and Demand." If a way to separate these events in CADIC's "Condition Negatively" that are equivalent to East-West's "Disapproval" and the CADIC "Punish" and "Interfere" collapsed into one category similar to East-West's "Conflictive Deeds," then a typology approximating East-West's is established. This approximation is contingent upon additional information available in CADIC with which to transform the CADIC data. Sufficient similarity between event data typologies

Specific Identifiers for Tactical Actions

14

As an aid to the coder, a description of the categories is provided here to help the coder to choose the proper category.

Behaviour					
Physical			Verbal		
(01) Punish	Destroy Attack Force Retaliate Seize	Deny Expel Resume (negative) Clash			
(02) Interfere	Interfere Obstruct Undermine Disrupt Suspend Impose Withhold	Freeze (i.e. funds, troop movements) Reduce (a routine international activity) Break diplomatic relations Block	(04) Threaten:		
(03) Mobilize	Mobilize Redeploy Resources Impose (domestic) Reorganize Recruit	Alert Adopt Demonstrate Increase capability	(05) Condition Negatively	Deny Condemn Accuse Reject Warn	Demand (negative) Criticize Protest Dissent Allege (negative)
(07) Consult	Discuss Visit Meet Negotiate Communicate		(06) Comment	Comment State Affirm Acknowledge Announce Report	
(11) Demobilize	Demobilize Terminate domestic alert Return to former condition Transfer Terminate coalition		(08) Appeal	Request for military assistance Request for information Request to join or form an Appeal / alliance Urge Request for third party mediation effort	
(12) Concede	Cease Retreat Evacuate Abandon	Reduce (positive) Withdraw Yield Give up	(09) Condition positively	Praise Approve Support Offer	Propose (positive) Assure Agree (Procedural issue)
(13) Reward	Reward Grant Ally Exchange	Recognize Return Increase Re-establish	(10) Promise:		

TABLE 2-4--The CADIC categories.

Source: Legault et. al. (1973:14).

already exists, however, for there to be hope of generating an approximation.

What has been shown here is that there is no clear set of instructions involving data-generation. The operational rule for selecting actors and targets had to be reconstructed and problems of a surrogate measure of alliance behavior had to be confronted. Next, a tangled skein of arbitrary constraints and time periods involved in transformation had to be investigated. Here, it was found that potential differences in measurement were left untouched. Finally, some similarities between CADIC and East-West were discovered. Overall, and especially upon the questions of the logical ordering of the transformations and the yet-to-be shown CADIC approximation of East-West, clarity and replicability exists for data-generation.

Estimation

Considering estimation involves examining how Tanter analyzed his data and whether the analyses and results present any difficulties of inference.

Tanter's general hypothesis is never actually examined. At no point in the text was the specification tested. If the general hypothesis remains untested but is still essential, then the general hypothesis is actually an axiom. All Tanter does is see whether the separate regressions for the two-phase typology yield organizational parameter estimates that are longer for WTO than NATO. Moreover, he notes that this is found in five of eight

instances (Tanter, 1974:141). Apparently, Tanter is not looking at the relative differences between the regressors of the two alliances in such a way that statistically takes the alliances to account by the covariance design given above. In only this way, can the statistical presence of alliance be accounted for in determining the size of organizational parameter for the autoregressive model.

Only the phase hypotheses are genuinely attempted. Here, however, Tanter does not always estimate the equations in an appropriate manner and certain aspects of statistical inference are also lacking. With respect to estimation, Tanter's primary fault is failing to report and take into account the problem of autocorrelation. Autocorrelation occurs when the error terms of a regression (or any other form of the general linear model)⁵ are highly correlated from one case to the next with a previous value. Usually, only the relationship between error_t and error_{t-1} and first-order autocorrelation (the only type considered in this thesis), is examined. The consequences of autocorrelation are that the sampling variances are liable to be understated which would give an artificially inflated set of F- and t-statistics (Johnston, 1972:246-249, Kmenta, 1971: 273-278) which has serious implications for statistical inference (Type-I error). At no point does Tanter list the amount of autocorrelation found; only for the 1961 reduction phase—alliance unspecified—is autocorrelation (.41) mentioned (Tanter, 1972:165). In order to avoid this problem, Tanter detrends the data and then chooses up to a lag

of structure of $t-3$ the regression that minimizes autocorrelation, whereas the usual procedures are to use first differences of the data if autocorrelation is close to $||$ or use an alternative regression procedure to ordinary least squares that will minimize the effect of autocorrelation (Kmenta, 1971:282-292). Indeed, Tanter's lack of concern about autocorrelation is shown when he writes (1974:168): "To estimate the organization process parameters, the alliance's conflictive behavior at $t-1$ is used because the first-order autocorrelations are the highest among autocorrelations for the model."

After deciding how to take autocorrelation into account, Tanter proceeds with his analysis. For the three phase typology, standardized coefficients were employed which means that Tanter's results cannot be compared to results found in this thesis since the actual differences of data are obfuscated. The results are summarized in Table 2-5. Note that only two parameters were statistically significant. The overall significance of the equation is nil, while the $\bar{R}^2$ reveal that the model seems to work best during crisis and for NATO. Neither t -statistics nor standard errors are given by Tanter so that the relative dominance of the parameters may not be evaluated. If, however, the reader is willing to assume unknown amounts of Type-I error, it is apparent that: for WTO the expectations hold for crisis only, and for NATO the expectations hold for precrisis only.

Tanter also presents results for the two phase typology

TABLE 2-5—Tanter's tri-phase results

Phase	Alliance	Parameters	R^2	R^{2a}	F^a
Precrisis	WTO	$-.16 \text{ WTO}_{t-1} + .12 \text{ NATO}_{t-1}$	.0370	-.1111	.2497
	NATO	$-.32 \text{ NATO}_{t-1} + .07 \text{ WTO}_{t-1}$	.1170	-.0188	.8613
Crisis	WTO	$-.1 \text{ WTO}_{t-1} + .31 \text{ NATO}_{t-1}$ ^b	.0820	-.0592	.5806
	NATO	$.23 \text{ NATO}_{t-1} + .13 \text{ WTO}_{t-1}$	.084	-.0569	.5961
Precrisis	WTO	$.02 \text{ WTO}_{t-1} + .04 \text{ NATO}_{t-1}$	.0200	-.1978	.0918
	NATO	$.16 \text{ NATO}_{t-1} + .32 \text{ WTO}_{t-1}$	.1300	-.0634	.6724

a Deduced from Tanter's information.

b Significance = .05 two-tailed.

Source: Tanter (1974:159-162)

in unstandardized form, making comparison with other results easier. The results Tanter obtained for the two phase typology are in Table 2-6. Intensification yields the best fitting model but no alliance seems best described by the model. The significance of parameters is not given, leaving the reader with less information than with the three phase typology results. Once again, since neither t -statistics nor standard errors are given, the relative dominance of parameters has to be taken on faith: for WTO all expectations are met, while for NATO no expectations are met.

Obviously, there are difficulties in accepting Tanter's estimations and results. The general hypothesis is raised to the level of an axiom when it was meant to be testable. The historical hypothesis is not tested properly. The three phase results are presented in a manner that does not aid comparison and full use of statistical inference is not made. With the two-phase results, the use of statistical inference actually declines below the level of the three phase results. Moreover, the method of handling autocorrelation leaves much to be desired. Presumably, imitators would appropriately correct for autocorrelation, use a lag structure of down to $t-3$, and choose the particular equation with the best explained variance. Overall, then, Tanter's estimation procedures and results are not easily accepted. Better estimation procedures and tests of hypotheses may be easily attained, and his results, with few exceptions, must be accepted on faith. Tanter's procedures cannot be imitated because they are inappropriate and his results are tenuous. Here, replicability is extremely limited.

TABLE 2-6—Tanter's two phase results

Phase	Alliance	Parameters	R^2	R^{2a}	F^a
Intensification	WTO	$-.5-.12 \text{ WTO}_{t-1} + .38 \text{ NATO}_{t-3}$	.5200	.4636	9.2083^b
	NATO	$.17-.18 \text{ NATO}_{t-1} -.43 \text{ WTO}_{t-2}$	.3100	.2333	4.0435^b
Reduction	WTO	$3.14+.26 \text{ WTO}_{t-1} + .05 \text{ NATO}_{t-1}$	.1700	.0778	1.8434
	NATO	$3.8+.22 \text{ NATO}_{t-1} + .54 \text{ NATO}_{t-1}$	.2900	.2111	3.6761^b

a Deduced from Tanter's information.

b Significance = .05 two-tailed.

Source: Tanter (1974:173,174)

Summary

As this chapter has progressed from one aspect of Tanter's autoregressive model to the next, confidence in replicability has decreased. The model may be misspecified. This, however, may be excused if we are willing to accept the tentativeness of Tanter's efforts. Data-generation must contend with Tanter's arbitrary use of certain transformations, aggregation periods, and the yet-to-be demonstrated compatibility of the East-West data set with the CADIC one. Estimation has numerous lacunae from improperly tested hypotheses to badly estimated equations. Despite these difficulties, Tanter's overall strategy is understandable. But, due to the lacunae we can never be sure if the interpretation given to Tanter is correct. Since Tanter's logic cannot be completely reconstructed or, in some cases, accepted, precise replication of procedures and variables is beyond reach. Only a rough approximation of Tanter, effectively altering aspects of the setting in unknown ways, is possible. At best, then, minimal replicability exists.

Notes

¹Hypotheses that compare the two alliances could not be attempted because CADIC does not have the 1948 crisis. Tanter also using alliance and phase as factors attempts in an ANOVA to determine if definition of the situation does offset behavior. This is not imitated here due to space limitations. Finally, Tanter, influenced still further by the 1914 studies, also uses an autoregressive model where perception interacts with the lagged endogenous variable. This, too, was not attempted due to space limitations.

²McCormick forms his variables differently than Tanter and looks at both nation-states and groups of nation states. Despite these differences, strong similarities are found to Tanter—marking further cumulation for this model since it has spread and been partially confirmed.

³This would seem to be the case with regard to Tanter's extensive current accounts and memory, originating in the West.

⁴Tanter's other historical hypotheses involve NATO but require the statistical presence of the 1948 and 1961 conflicts.

⁵Autocorrelation could affect ANOVA in the same way as it does regression. They could have potentially serious effects upon Tanter's attempt to examine the definition of the situation hypothesis since his data are time-serial.

CHAPTER 3

CREATING AND COMPARING THE DATA BASES

Overview

This chapter discusses the creation of the data bases and compares them so that the similarity in settings between Tanter and this replication may be understood. First, the general nature of the two data sets being used is discussed so that their similarity of purpose, which affect extracting and coding, is known. Second, the manner of generating the two data bases is discussed. These two steps raise the problem of potential measurement error. If the two data sets are dissimilar in purpose and the creation of the two data bases involves dubious procedures, then considerable measurement error which may yield dissimilar data bases but similar results is a distant possibility. This possibility of dissimilar data yielding similar results may be further compounded by bad research design in the selection of replicatory estimation procedures and techniques of comparing the two sets of findings. The first two topics, then, are vital to interpretation of replicatory results.

The next four topics in this chapter actually compare the data. First, estimation is discussed. In order to see if missing data are handled in the same way, the impact of the memory transformation is examined. There, potential distortion in the scaling of the CADIC event data is evaluated by seeing if its relationship

of frequency to scale values is similar to that in the East-West event data. Finally, the two data sets are formally compared by regressing the East-West event data upon the CADIC event data.

General nature of the data sets

By looking first at the general nature of the data sets, their purposes are revealed which may in time tell us about the implicit rules of and coding that will affect their similarity. The East-West event data set was an extended examination of interactions between the East and West from 1945 to 1965. The scope of the actors covered is large; no specific subset of actors is focussed upon a priori. Major sources are New York Times Deadline Data, and World Almanac (Burgess and Lawton, 1972:32-35). East-West event data are broadly based and includes the Berlin 1961 conflict as a matter of course.

The CADIC event data are not as broadly based as are the East-West event data. Only Berlin 1961 is looked at, and only one source, the New York Times, was used. Moreover, the CADIC event data concentrate solely upon alliance interactions, whereas the East-West event data concentrate upon both the local and alliance levels.¹ The differences in the purposes cannot be minimized. Given the differences of sources, breadth of coverage, and level-of-actors analyzed, extracting and coding cannot be the same. Potential differences between the two data bases are great.

Creating the data bases

In creating the data bases (see Appendix A), both untransformed and transformed frequencies and scores for the daily and

three-day aggregation periods were developed. The first set of data created of this type was East-West event data. Unfortunately, Corson has withdrawn East-West event data from public use. Tanter (1974:247-252) does provide a raw data appendix that is slightly different from the data he used. In addition to this possible source of error, events belonging to the "Disapproval" category (see table 2-2), which is the perception variable not used in the model being investigated here, had to be excluded. Each description was deemed to be one event, scale values were multiplied by -1 in order to make them consistent with the scaling of CADIC event data. The peak intensity, dividing between the intensification and reduction phases, was taken directly from Tanter as 14 August 1961 ($T=71$ for daily aggregation; $t=23$ for three-day aggregation). This, of course, may not match what was actually obtained from the event data coded from Tanter's appendix. Measurement error problems, then, exist with regard to the similarity of Tanter's actual data and its reconstruction, with regard to the extraction of events, and finally with regard to the selection of phases.

The CADIC event data presents more difficulties than the East-West event data. Since it was decided to recode the extent CADIC event data by using a selection process instead of recoding the data by reextracting and coding the raw event descriptions, usually used where others have attempted creation of data bases that imitate other event data sets (Harvner and Peterson, 1975; Leng, 1975), adequate information was needed to filter the CADIC event data. The CADIC event data set, as table 2-4 reveals, does not have as discrete a category structure as the East-West event data.

(see table 2-2). Fortunately, adequate information for filtering is provided by the double coding ("with what by whom was the actual event perpetrated?"). Thus, it is possible to omit certain original CADIC event data that would not exist in the East-West event data typology.

First, consider the East-West "disapproval" category. Here, the CADIC category "condition negatively" would appear most similar. Since this category represents perception, according to Tanter, and the model with perception is not being considered here, this category is automatically excluded from the analysis. There is potential loss of information, however, since the CADIC "condition negatively" contains the keywords 'demand (negative)' and 'protest' that are subcategories of the East-West "demands" category, and also contains 'reject' which is part of the East-West "threats" category. But, since the "disapproval" category in East-West requires that the injured party clearly state that it has been offended, any CADIC "condition negatively" event that is double coded is equivalent to the East-West "disapproval" category. Thus, any CADIC "condition negatively" events that were double-coded were excluded. Since the East-West subcategories, 'protest' and 'demand (negative)' would be included in the filtered CADIC "condition negatively" category, the resulting data were held to be equivalent to the East-West "demands" category. Only 'reject' is lost from the filtering. The CADIC category of "threats" and the East-West "threats" both rely upon a conditional "if...then" statement, and may be held to be equivalent. Only an equivalency for the East-West subcategory 'reject' cannot be established. The East-West "conflictive deeds" category

is simple to reconstruct because it contains the information contained in the CADIC categories "punish," "interfere," and "mobilize." After examining the data that are created, it may be noted that the replicatory data peak at the same time as the East-West data—a strong argument for the propriety of the filtering procedures. There is, then, overall, compatibility between the two event data sets. Only reconstruction of "demands" seems somewhat dubious.

Scaling the replicatory data was done very simply. Using the scaling sheets provided by Carson (1975:11,12; see figs. 3.1 and 3.2), the scale values of subcategories for all resource areas except "conflict about cooperation," "verbal conflict: trade basis," and "economic conflict" which do not apply to the Berlin 1961 crisis were assigned by inspection, and the median values for the categories they formed used as the score for each replicatory event in the equivalent category: 87 for "demands"; 147 for "threats"; 247 for "conflictive deeds." The median was preferred to the mean which would be subject to more distortion. Obviously, subcategories, being non-existent in the replicatory data, could not be given their own scale values for use in the analysis. This procedure yields a different level-of-measurement than Tanter used. This, however, does not appear to have very serious consequences for the replicatory regression results as reported in chapter four (see table 4-1) where only the intercepts are different than Tanter's but the other parameters are of the same order. Overall, then, the scaling procedure is reasonable.

On estimation

Before presenting results of the regressions of the variables, estimation of the regression must be discussed. This is divided into two parts: first, disclosing what the assumptions are about the disturbances involved in the estimates to be presented, second, stating what procedures will be followed in order to avoid the sole admitted disturbance—autocorrelation.

In both this and the following chapters, the behavior of alliances are being analyzed. By dealing with alliances such as NATO and WTO, we are examining what ought to be sets of behaviors emitted by members states of the alliances over time. Since no two states in either alliance has the same capacity to respond during the Berlin 1961 crisis, heteroskedacity may be expected. Since time-series is involved, autocorrelation is also a distinct possibility. Moreover, since alliance members communicate among themselves, mutual influence, mutual correlation, may also be expected. Situations of this type, as shown by Kmenta (1971: 512-517), may be estimated, but this is not attempted here nor was it attempted by Tanter. By using the transformed mean intensity of all alliance members, the more specific presentation of disturbances is avoided. This avoidance is defensible under the following conditions. With regard to heteroskedacity it must be assumed that the USA and USSR dominate these alliances throughout the period being examined. If this is so, then subordinate alliance members cannot dominate any part of the period being studied or appear vexed with alliance leaders throughout the crisis. Thus, the equations are homogenous

because they are dominated by the alliance leaders. With regard to mutual correlation, the assumption of homoskedasticity leads to the conclusion that there are no subordinate alliance members present influencing one another and that whatever mutual influence that exist flow solely from the alliance leader to subordinates.

Use of transformed mean intensity presupposes assumptions that yield a classic bipolar alliance system: two alliances with strict downward communication from alliance leaders which initiate the only interactions. Under such assumptions, the only disturbance to be dealt with is autocorrelation.

Autocorrelation is a situation where error terms are correlated with one another from one observation to another. Although autocorrelation does not affect the coefficients, the variances (standard errors) of the coefficients are usually underestimated (which leads to Type-I error) (Hibbs, 1944:257), producing falsely inflated t-ratios. Many possible mechanisms may produce autocorrelation. Only the simplest of these mechanisms, first order autocorrelation—error in the present period is correlated only with the immediately preceding error—is considered. For the results presented in this chapter only, the Durbin-Watson statistic at .05 confidence is used (see Kmenta, 1971:295-297 for a description). As a matter of caution, ordinary least squares estimates (OLS) are deemed autocorrelated if their Durbin-Watson statistic is in the uncertainty region between no autocorrelation and autocorrelation. Should autocorrelation be found, pseudo-generalized least squares (pseudo-GLS)—so called because it involves an estimate only of the

autocorrelated disturbance. Here, variables in the equation to be estimated are transformed by subtracting the lagged value of the variable multiplied by the estimate of autocorrelation from the variables' present value (Hibbs, 1974:282). This procedure will yield variances and standard errors that are not falsely deflated. This procedure coupled with the assumptions of no heteroskedacity and no mutual correlation, yields a simple estimation procedure that will be followed in this chapter and with some modifications in the following chapter.

The effect of the transformation

The first issue to be dealt with is the effect of the transformation: how is it related to the untransformed series? Does the transformation have a stable systematic relationship across the data sets? The issue here is whether the transformation, supposedly representing memory but actually indicating missing data, has the same effect on both the Tanter and replicatory data sets. Should the effect of the transformation be different, then neither memory nor the elimination of missing data, is the same for both data set. Use of the decay function serves as an interpolation device is the same as Deber's (1975:65) linear interpolation which "assumes a smooth trend...and minimal fluctuation...a quiet bridging of present and future hostilities must have taken place..." Use of this mechanism presupposes a stable process; if many jumps occur, then the mechanism is not warranted. Thus, by examining the effect of the transformation, we shall see if both Tanter's and replicatory

data display the same smoothing of an already stable process.

If smoothing should occur, then a slope less than or equal to zero between the transformed and untransformed data is the null hypothesis. The alternative is that a positive slope will exist. Examination of table 3-1, however, reveals that for both alliances, whether using frequencies or scores, null or negative slope occur across all phases for both daily data sets. Both data sets show that a random relationship exists; there is no systematic relationship between the transformed and untransformed. To a large extent, this may be due to measurement error involved in using daily time aggregation. When the three day time aggregation is considered (see table 3-2), the pattern is reversed. The null hypothesis is accepted one fourth of the time at best. For all alliances, phases, and data sets, there is a significant smoothing of the untransformed data. Thus, the use of memory as a device to eliminate missing data is successful only where it is supposed to be used: the three day data used in regression. This applies to both data sets; therefore, the smoothing process is equally successful, stable, across all data sets.

Stability of measurement

The next issue to be dealt with is stability of measurement. Do the two data sets differ in the relationship between the more primitive measuring instrument, frequencies, and the more sophisticated instrument, the scale values? Based on past experience in the event data movement, positive slopes are expected between the

TABLE 3-1—Daily comparison of transformed and untransformed data.

Intensification

Tanter

Replication

TWFD-TWFDT -.1466
 TWSD-TWSDT .2444
 TNFD-TNFDT .3248
 TNSD-TNSDT -.3439

RWFD-RWFDT .0485
 RWSD-RWSDT -.7925
 RNFD-RNFDT -.2180
 RNSD-RNSDT -.1842

% H_0 accepted 100

100

Reduction

Tanter

Replication

TWFD-TWFDT -.0586
 TWSD-TWSDT -.4207*
 TNFD-TNFDT .3807
 TNSD-TNSDT .6293*

RWFD-RWFDT .0531
 RWSD-RWSDT -.1058
 RNFD-RNFDT .1238
 RNSD-RNSDT .4115

% H_0 accepted 100

100

Note: Underlined coefficients estimated by pseudo-GLS.

Here: TWFD=Tanter's WTO frequency daily; TWFDT=Tanter's WTO frequency transformed daily; RWFD=replicatory WTO frequency daily; RWFDT=replicatory WTO frequency transformed daily; TWSD=Tanter's WTO score daily; TWSDT=Tanter's WTO score transformed daily; RWSD=replicatory WTO score daily; RWSDT=replicatory WTO score transformed daily; TNFD=Tanter's NATO frequency daily; TNFDT=Tanter's NATO frequency transformed daily; RNFD=replicatory NATO frequency daily; RNFDT=replicatory NATO frequency transformed daily; TNSD=Tanter's NATO score daily; TNSDT=Tanter's NATO score transformed daily; RNSD=replicatory NATO score daily; RNSDT=replicatory NATO score transformed daily.

*significance = .01

TABLE 3-2—Three day comparison of transformed and untransformed data

Intensification

Tanter

Replication

TWFT-TWFTT	<u>.5031*</u>	RWFT-RWFTT	<u>.7827*</u>
TWST-TWSTT	<u>.3115</u>	RWST-RWSTT	<u>.5747*</u>
TNFT-TNFTT	<u>.7659*</u>	RNFT-RNFTT	<u>.5605*</u>
TNST-TNSTT	<u>.6705*</u>	RNST-RNSTT	<u>.8916*</u>

% H_0 accepted

25

0

Reduction

Tanter

Replication

TWFT-TWFTT	<u>.5221*</u>	RWFT-RWFTT	<u>.6309*</u>
TWST-TWSTT	<u>.8960*</u>	RWST-RWSTT	<u>.7420*</u>
TNFT-TNFTT	<u>.9011*</u>	RNFT-RNFTT	<u>.4534</u>
TNST-TNSTT	<u>.9100*</u>	RNST-RNSTT	<u>.7212*</u>

% H_0 accepted

0

25

Note: Underlined coefficients estimated by pseudo-GLS.

Here: TWFT=Tanter's WTO frequency three-day; TWFTT=Tanter's WTO frequency transformed three-day; RWFT=replicatory WTO frequency three-day; RWFTT=replicatory WTO frequency transformed three-day; TWST=Tanter's WTO score three-day; TWSTT=Tanter's WTO score transformed three-day; RWST=replicatory WTO score three-day; RWSTT=replicatory WTO score transformed three-day; TNFT=Tanter's NATO frequency three-day; TNFTT=Tanter's NATO frequency transformed three-day; RNFT=replicatory NATO frequency three-day; RNFTT=replicatory NATO frequency transformed three-day; TNST=Tanter's NATO score three-day; TNSTT=Tanter's NATO score transformed three-day; RNST=replicatory NATO score three-day; RNSTT=replicatory NATO score transformed three-day.

*significance = .05

two measuring systems. Handelman (1975:180) by way of example, that Azar's COPDAB yields strong and significant correlations between the scale values and frequencies of Arab-Israeli conflict from 1965 to 1970. Since a simple one-to-one relationship exists between replicatory frequencies (categories) and scale values, significantly positive slopes are expected here. What remains to be seen, however, is whether Tanter's data yield the same pattern. If it does not, and the replicatory data display the expected positive significant slope, then the replicatory data are distorted compared to Tanter's. If this should happen, then the reliability of the scaling procedures used for the replicatory data is dubious, and potential measurement error that could effect the success of the replication is discovered.

If we expect significant positive slopes, then the null hypothesis is that the slopes are less than or equal to zero.

Examination of daily data (see table 3-3) shows that this is never the case in all instances, strong positive relationships between frequencies and scores exist for all data set, phases, and alliances. The three day data reveal a slightly weaker pattern (see table 3-4). The replicatory transformed NATO data during intensification has a significant positive slope, whereas the equivalent Tanter data does not. Only here, then, does potential measurement error arise.

Both aggregation periods, however, generally support the reliability of the replicatory assignment of scale values—out of thirty-two chances, only once is there disagreement between the two data sets.

TABLE 3-3—Daily comparison of scores and frequencies

Intensification

Tanter

Replication

TWSD-TWFD	23.3907*	RWSD-RWFD	75.1206*
TNSD-TNFD	18.8072*	RNSD-RNFD	<u>57.2220*</u>
TWSDT-TWFDT	30.6238*	RWSDT-RWFDT	<u>65.3152*</u>
TNSDT-TNFDT	19.7130*	RNSDT-RNFDT	51.9386*

% H_0 accepted

0

0

Reduction

Tanter

Replication

TWSD-TWFD	<u>26.2148*</u>	RWSD-RWFD	<u>46.8056*</u>
TNSD-TNFD	<u>37.3206*</u>	RNSD-RNFD	<u>44.0355*</u>
TWSDT-TWFDT	<u>41.4326*</u>	RWSDT-RWFDT	<u>43.8756*</u>
TNSDT-TNFDT	<u>41.9571*</u>	RNSDT-RNFDT	<u>42.0828*</u>

% H_0 accepted

0

0

Note: Underlined coefficients estimated by pseudo-GLS.

Here: TWSD=Tanter's WTO score daily; TWFD=Tanter's WTO frequency daily; RWSD=replicatory WTO score daily; RWFD=replicatory WTO frequency daily; TNSD=Tanter's NATO score daily; TNFD=Tanter's NATO frequency daily; RNSD=replicatory NATO score daily; RNFD=replicatory NATO frequency daily; TWSDT=Tanter's WTO score transformed daily; TWFDT=Tanter's WTO frequency transformed daily; RWSDT=replicatory WTO score transformed daily; RWFDT=replicatory WTO frequency transformed daily; TNSDT=Tanter's NATO score transformed daily; TNFDT=Tanter's NATO frequency transformed daily; RNSDT=replicatory NATO score transformed daily; RNFDT=replicatory NATO frequency transformed daily.

*significance = .01 one-tailed

TABLE 3-4—Three-day comparison of scores and frequencies

Intensification

Tanter

Replication

TWST-TWFT	26.2889*	RWST-RWFT	63.5376*
TNST-TNFT	<u>19.2375*</u>	RNST-RNFT	50.7036*
TWSTT-TWFTT	26.7997*	RWSTT-RWFTT	63.2678*
TNSTT-TNFTT	20.8487	RNSTT-RNFTT	<u>51.4730*</u>

% H₀ accepted

25

0

Reduction

Tanter

Replication

TWST-TWFT	<u>45.6412*</u>	RWST-RWFT	59.6364*
TNST-TNFT	<u>64.5271*</u>	RNST-RNFT	<u>23.5880*</u>
TWSTT-TWFTT	<u>44.5341*</u>	RWSTT-RWFTT	<u>77.4773*</u>
TNSTT-TNFTT	<u>63.8049*</u>	RNSTT-RNFTT	<u>49.1408*</u>

% H₀ accepted

0

0

Note: Underlined coefficients estimated by pseudo-GLS.

Here: TWST=Tanter's WTO score three-day; TWFT=Tanter's WTO frequency three-day; RWST=replicatory WTO score three-day; RWFT=replicatory WTO frequency three-day; TNST=Tanter's NATO score three-day; TNFT=Tanter's NATO frequency three-day; RNST=replicatory NATO score three-day; RNFT=replicatory NATO frequency three-day; TWSTT=Tanter's WTO score transformed three-day; TWFTT=Tanter's WTO frequency transformed three-day; RWSTT=replicatory WTO score transformed three-day; RWFTT=replicatory WTO frequency transformed three-day; TNSTT=Tanter's NATO score transformed three-day; TNFTT=Tanter's NATO frequency transformed three-day; RNSTT=replicatory NATO score transformed three-day; RNFTT=replicatory NATO frequency transformed three-day.

*significance = .05 one-tailed

Comparison of the two data sets

Finally, the two data sets may be compared. Here, both daily and three day data shall be compared. Should any differences exist here, then it is obvious that the aggregation period is having an effect. Moreover, transformed and untransformed scores and frequencies are also included. Should differences exist between transformed and untransformed data, then the transformations are having an effect. Most important of all is the inclusion of frequencies. Here, frequencies, subject to less distortion due to assignment of different scale values, may provide another view of the similarity between the two data sets than scale values. If, in fact, frequencies yield a significant positive slope while scale values do not, then the similarity of the two data sets for that phase, alliance, and transformation may be argued. But, if scales are similar and frequencies dissimilar, this argument is not forthcoming.

If positive significant slopes are expected then the null hypothesis is that the slopes are less than or equal to zero. As seen in table 3-5, the null hypothesis is accepted the vast majority of times. Only during intensification do two significantly positive slopes appear. These slopes are, however, positive for untransformed daily scores of both alliances, and cannot be construed to imply similarity and potential generalizability of Tanter's findings. The three day results (see table 3-6) uniformly accept the null hypothesis all the time. Similarity of data sets and generalizability of Tanter's findings cannot be established. Difficulties

TABLE 3-5—Daily comparison of Tanter's and replicatory data

Variables	Intensification	Reduction
TWFD-RWFD	.1817	-.0168
TWSD-RWSD	.1653*	-.0017
TNFD-RNFD	-.0147	.0495
TNSD-RNSD	<u>-.1292*</u>	.0254
TWFDT-RWFDT	<u>.2194</u>	-.0057
TWSDT-RWSDT	-.0100	-.1896
TNFDT-RNFDT	<u>-.0321</u>	.0392
TNSDT-RNSDT	<u>-.0384</u>	<u>.0057</u>
% H ₀ accepted	75	100

Note: Underlined coefficients estimated by pseudo-GLS.

Here: TWFD=Tanter's WTO frequency daily; RWFD=replicatory WTO frequency daily; TWSD=Tanter's WTO score daily; RWSD=replicatory WTO score daily; TNFD=Tanter's NATO frequency daily; RNFD=replicatory NATO frequency daily; TNSD=Tanter's NATO score daily; RNSD=replicatory NATO score daily; TWFDT=Tanter's WTO frequency transformed daily; RWFDT=replicatory WTO frequency transformed daily; TWSDT=Tanter's WTO score transformed daily; RWSDT=replicatory WTO score transformed daily; TNFDT=Tanter's NATO frequency transformed daily; RNFDT=replicatory NATO frequency transformed daily; TNSDT=Tanter's NATO score transformed daily; RNSDT=replicatory NATO score transformed daily.

TABLE 3-6—Three day comparison of Tanter and replicatory data

Variables	Intensification	Reduction
TWFT-RWFT	.0552	-.0714
TWST-RWST	.0957	-.2187
TNFT-RNFT	-.0784	<u>.0508</u>
TNST-RNST	-.0332	<u>-.1991</u>
TWFTT-RWFTT	<u>.4107</u>	-.0958
TWSTT-RWSTT	<u>.1350</u>	-.2762
TNFTT-RNFTT	<u>-.1637</u>	-.0599
TNSTT-RNSTT	<u>-.0674</u>	<u>-.2583</u>
% H ₀ accepted	100	100

Note: Underlined coefficients estimated by pseudo-GHS.

Here: TWFT-Tanter's WTO frequency three-day; RWFT-replicatory WTO frequency three-day; TWST-Tanter's WTO score three-day; RWST-replicatory WTO score three-day; TNFT-Tanter's NATO frequency three-day; RNFT-replicatory NATO frequency three-day; TNST-Tanter's NATO score three-day; RNST-replicatory NATO score three-day; TWFTT-Tanter's WTO frequency transformed three-day; RWFTT-replicatory WTO frequency transformed three-day; TWSTT-Tanter's WTO score transformed three-day; RWSTT-replicatory WTO score transformed three-day; TNFTT-Tanter's NATO frequency transformed three-day; RNFTT-replicatory NATO frequency transformed three-day; TNSTT-Tanter's NATO score transformed three-day; RNSTT-replicatory NATO score transformed three-day.

with replication may be due to differences in the origins of the two data sets, problems in creating the two data sets, inadequate reporting of Tanter's results, and possible difficulties in estimation in this thesis. All of these possibilities are definite considerations in evaluating generalizability.

Summary

This chapter began by discussing the procedures by which the two data bases were created. In both instances, suboptimal procedures were employed. It was necessary to recreate Tanter's data by extracting the events from a list that was not identical to the data that he used. With the CADIC event data, recoding and assignment of scale values presented difficulties. Once the data were obtained, they were analyzed with respect to their similarities in three ways. First the two data sets were examined for similarities in the relationship between untransformed and transformed series (for memory and missing data). Second, the relationships between frequencies and scale values were examined as a means of validating the assignment of scale values to the replicatory data. Finally, the two data sets were compared in order to establish generalizability. Similar, strong positive slopes related the transformations to the original data for the three day data. Frequencies and scales were closely related, validating the assignment of scale values to the replicatory data. No similarities were found however, between the two data sets. This independence, probably a function of measurement error

introduced in the attempt to recreate Tanter's data, limits the generalizability of the findings of this replication.

Notes

¹According to a personal communication from Professor Sigler (25 August 1977).

CHAPTER 4

ON THE GENERALIZABILITY OF TANTER'S RESULTS

Overview

This chapter assesses the generalizability of Tanter's results. Since the previous chapter has shown that the two data sets are dissimilar, assessing the generalizability of Tanter's results involves determining if research design as seen in estimation procedures and selection of models will rescue any findings. Attempting to rescue Tanter's results in this manner deliberately avoids immediate rejection and seeks to retain some of these findings as a basis for further research. The first step here is to develop a more plausible procedure for estimating autoregressive models which shall be used in conjunction with ordinary least squares' results in the remainder of the chapter. Next, the two sets of replicatory findings are presented and substantive interpretations are given with due consideration of the ordinary least squares' and the more plausible procedure's results. Finally, assessment of the generalizability of Tanter's results is undertaken in the manner described above.

A plausible solution for autoregressive models

During the second chapter, it was suggested that Tanter's autoregressive model exacerbated his questionable treatment of autocorrelation. According to Douglas A. Hibbs (1974:294), the

most commonplace problems of an autoregressive model suffering from autocorrelation estimated via ordinary least squares are

upwardly biased coefficients for the lagged y_{t-1} [lagged endogenous variable] and downwardly biased (toward zero) coefficients for the exogenous x_{t-1} [exogenous variable]. Since the lagged dependent variables tend to dominate the outcomes of such equations (partially) at the expense of exogenous influences, this leads to erroneous inferences about the importance of the dynamic [lagged endogenous] terms....[Moreover], OLS residuals cannot...be used to determine the structure or magnitude of the disturbance time dependence [sic].

Students of international politics who naively employ autoregressive models are caught in a dilemma that may be partially relieved if a more plausible procedure can be developed.

The basic problem is to determine how autocorrelated an autoregressive equation is and then use this information for proper reestimation. Unfortunately, the Durbin-Watson statistic is inappropriate (see Nerlove and Wallis, 1966) and the small sample spaces used in this thesis ($n < 30$) do not permit use of Durbin's autocorrelation statistic (1970) for autoregressive models. Therefore, a cumbersome procedure, shown by Hibbs (1974:290-296), shall be used and will now be briefly outlined. Consider an autoregressive model equivalent to that being estimated in this thesis

$$Y_t = a + Y_{t-1} + X_t + u$$

where Y_t = dependent variable

a = intercept

Y_{t-1} = lagged endogenous variable at $t-1$

X_{t-m} = exogenous variable at $t-m$ (n not necessarily equal to m)

u = error term

The actual systematic presence of first order autocorrelation (ρ) in the above model is shown by

$$\rho u_{t-1} = a + \rho Y_{t-1} + \rho Y_{t-2} - \rho X_{t-1} + v$$

in contrast to

$$u_t = a + \rho u_{t-1} + e$$

which is a satisfactory representation of first order autocorrelation in nonautoregressive equations. The autoregressive equation must be revised on the basis of the actual/systematic presence of autocorrelation by substituting the correct autocorrelation generation mechanism for the incorrect autocorrelation mechanism. With appropriate manipulation, this yields

$$Y_t = a + (b + \rho)Y_{t-1} + X_t - \rho Y_{t-2} - \rho X_{t-1} + v,$$

The only "true" coefficients here are $b + \rho$ for Y_{t-1} and X_t is an identity, while Y_{t-2} and ρX_{t-1} are variables that were incorrectly omitted prior to integration of the autocorrelation mechanisms into the equation. Realizing here that $\rho = -(-\rho X_{t-1}/X_t)$, Y_{t-1} therefore is $Y_{t-1} - \rho$. A more plausible estimate of an autoregressive model may be generated with the above four variable equation by taking advantage of the equalities and the identity of the exogenous variable. Hibbs (1974:296) observes that this procedure will not yield as sharp a set of estimates as using the ρ generated above in a pseudo-GLS estimation of the original model. The more plausible procedure followed in this thesis for estimating autoregressive models is to generate the four variable model for the proper estimate of ρ and then use ρ for pseudo-GLS.

The procedure outlined above is not free of possible difficulties.

First, the estimation of p from the four variable equation may yield values greater than $|||$. If this is the case, then it is quite probable that an autocorrelation mechanism greater than first order is present. Presumably, p_1 and p_2 could be estimated with the inclusion of additional variables. Since this thesis assumes the simplest possible case, first-order autocorrelation, only those instances where $p \leq |||$ are considered. In those instances where $p > |||$ obtains uncorrected OLS estimates are used. The more often this occurs, the more suspect are the assumption of first order autocorrelation and, above all else, so are the parameter estimates arising from the default OLS algorithm. Hibbs (1974:296), after deriving the above procedure, notes that this procedure may not give as good results as OLS if the order of autocorrelation is misspecified. Once again, the simple assumption of first-order autocorrelation may lead to incorrect results. Finally, the generation of estimates of the autocorrelation process used here is not the best that is possible. Nonlinear regression upon the excluded variables model yields excellent small sample results; but, computational expense and lack of a priori information with which to set the range of values of variables close this route. Two instrumental variable approaches, using the exogenous variables as instruments, are also available (Hibbs, 1974:295-296, 298). Unfortunately, as shown below (see table 4-1), the exogenous variables and endogenous variables are not highly correlated during the intensification phase which means that the exogenous variables cannot be used as instrumental variables. (Kmenta, 1971:309). By default, then, the most plausible

procedure is the one outlined above.

Regression results and substantive findings

After discussing estimation, replicatory regression results and substantive findings may now be attempted. Here, the most general results of the lag structure apply. Nine equations for two alliances for two phases were estimated by OLS and pseudo-GLS where possible. From these results, the best fitting models are chosen and evaluated. Should the best fitting model have been estimated by OLS, the results may well be suspect and the best fitting pseudo-GLS results are given. Moreover, the dominance of one parameter over another, (see ch. 2, p. 33) is evaluated from the OLS and pseudo-GLS evidence. Changes, if any, in substantive results from the use of pseudo-GLS are carefully noted along with any differences between selection of best fitting models.

The OLS and pseudo-GLS results are presented in table 4.1. Here, note that estimates of p for OLS are presented in only the cases where it was less than .1; otherwise no value of p is given. Both the OLS and pseudo-GLS equations present a measure of multicollinearity between the two independent variables. In the case of two independent variables, the Pearson product moment between them is a sufficient indication of multicollinearity. Should this correlation be significant, .05 two-tailed, then the standard errors are inflated, rendering parameter estimates imprecise (Kmenta, 1971:389).

The most immediately noticeable aspect of results presented in table 4-1 is that many of the independent variable parameter for all phases, estimated by either OLS or pseudo-GLS, are statistically

TABLE 4-1—Replicatory regression results

Intensification

OLS

$$\hat{NATO} = 20.1364 + .5071 NATO_{t-1} + .1017 WTO_{t-1} + 44.2239$$

$$14.0256 \quad .1904* \quad .1745$$

$$R^2 = .2946 \quad \bar{R}^2 = .2203 \quad F(2,19) = 3.9677$$

$$p = \text{---} \quad \text{multicol.} = .1224 \quad n = 22$$

$$\hat{WTO} = 34.1749 + .2913 WTO_{t-1} - .0548 NATO_{t-1} + 55.7240$$

$$17.6729 \quad .2199 \quad .2399$$

$$R^2 = .0848 \quad \bar{R}^2 = -.0115 \quad F(2,19) = .8802$$

$$p = .9264 \quad \text{multicol.} = .1224 \quad n = 22$$

$$\hat{NATO} = 19.6257 + .4511 NATO_{t-1} + .2375 WTO_{t-2} + 43.6223$$

$$13.9683 \quad .1940* \quad .1902$$

$$R^2 = .3198 \quad \bar{R}^2 = .2442 \quad F(2,18) = 4.2314*$$

$$p = -.9069 \quad \text{multicol.} = .1995 \quad n = 21$$

$$\hat{WTO} = 28.1024 + .2629 WTO_{t-1} + .1334 NATO_{t-2} + 56.3348$$

$$19.2712 \quad .2235 \quad .2454$$

$$R^2 = .0859 \quad \bar{R}^2 = -.0157 \quad F(2,18) = .8457$$

$$p = \text{---} \quad \text{multicol.} = .0093 \quad n = 21$$

$$\hat{NATO} = 41.9567 + .0392 NATO_{t-2} + .1314 WTO_{t-1} + 49.7123$$

$$17.8798* \quad .2222 \quad .2180$$

$$R^2 = .0235 \quad \bar{R}^2 = -.0850 \quad F(2,18) = .2166$$

$$p = \text{---} \quad \text{multicol.} = .0093 \quad n = 21$$

$$\hat{WTO} = 40.9446 + .2768 WTO_{t-2} - .1118 NATO_{t-1} + 56.9113$$

$$18.2235* \quad .2482 \quad .2531$$

$$R^2 = .0671 \quad \bar{R}^2 = -.0366 \quad F(2,18) = .6473$$

$$p = .7158 \quad \text{multicol.} = .1995 \quad n = 21$$

TABLE 4-1--(continued)

$$\hat{NATO} = 39.2283 - .0354 NATO_{t-2} + .3335 WTO_{t-2} + 49.7123$$

$$15.8391 * .2222 \quad .2180$$

$$R^2 = .1166 \quad \bar{R}^2 = .0184 \quad F(2,18) = 1.1879$$

$$p = \text{---} \quad \text{multicol.} = .2235 \quad n = 21$$

$$\hat{WTO} = 33.1932 + .2370 WTO_{t-2} + .0821 NATO_{t-2} + 57.0548$$

$$18.1785 \quad .2502 \quad .2550$$

$$R^2 = .0624 \quad \bar{R}^2 = -.0418 \quad F(2,18) = .5990$$

$$p = \text{---} \quad \text{multicol.} = .2235 \quad n = 21$$

$$\hat{NATO} = 16.6353 + .4539 NATO_{t-1} + .2651 WTO_{t-3} + 42.8061$$

$$13.8072 \quad .2037 * \quad .2045$$

$$R^2 = .3720 \quad \bar{R}^2 = .2981 \quad F(2,17) = 5.0350 *$$

$$p = -.2725 \quad \text{multicol.} = .3498 \quad n = 20$$

$$\hat{WTO} = 51.6306 + .2854 WTO_{t-1} - .3334 NATO_{t-3} + 54.9288$$

$$18.9951 * .2238 \quad .2471$$

$$R^2 = .1497 \quad \bar{R}^2 = .0497 \quad F(2,17) = 1.4965$$

$$p = -.1322 \quad \text{multicol.} = .1519 \quad n = 20$$

$$\hat{NATO} = 55.3757 - .3274 NATO_{t-3} + .1999 WTO_{t-1} + 50.1931$$

$$17.3574 * .2258 \quad .2054$$

$$R^2 = .1366 \quad \bar{R}^2 = .0350 \quad F(2,17) = 1.3448$$

$$p = \text{---} \quad \text{multicol.} = .1519 \quad n = 20$$

$$\hat{WTO} = 51.8052 + .0644 WTO_{t-3} - .1259 NATO_{t-1} + 59.2098$$

$$19.0983 * .2742 \quad .2817$$

$$R^2 = .0120 \quad \bar{R}^2 = -.1042 \quad F(2,17) = .1032$$

$$p = .9790 \quad \text{multicol.} = .3498 \quad n = 20$$

$$\hat{NATO} = 25.6422 - .0392 NATO_{t-2} + .4291 WTO_{t-3} + 48.6151$$

$$15.9413 * .2235 \quad .2176$$

$$R^2 = .1900 \quad \bar{R}^2 = .1065 \quad F(2,17) = 2.2746$$

$$p = .8411 \quad \text{multicol.} = .2452 \quad n = 20$$

TABLE 4-1--(continued)

$$\hat{WTO} = 53.5154 + .2597 WTO_{t-2} - .3120 NATO_{t-3} + 55.6399$$

$$19.0926^* .2419 \quad .2486$$

$$R^2 = .1276 \quad \bar{R}^2 = .0250 \quad F(2,17) = 1.2432$$

$$p = \text{---} \quad \text{multicol.} = .0989 \quad n = 20$$

$$\hat{NATO} = 50.2150 - .3329 NATO_{t-3} + .3843 WTO_{t-2} + 46.9151$$

$$16.0987^* .2096 \quad .2039$$

$$R^2 = .2457 \quad \bar{R}^2 = .1570 \quad F(2,17) = 2.7687$$

$$p = \text{---} \quad \text{multicol.} = .0989 \quad n = 20$$

$$\hat{WTO} = 42.8970 - .0019 WTO_{t-3} + .0981 NATO_{t-2} + 59.3315$$

$$19.4553^* .2656 \quad .2727$$

$$R^2 = .0080 \quad \bar{R}^2 = -.1087 \quad F(2,17) = .0685$$

$$p = \text{---} \quad \text{multicol.} = .2452 \quad n = 20$$

$$\hat{NATO} = 51.9738 - .4406 NATO_{t-3} + .5548 WTO_{t-3} + 42.7967$$

$$13.9515^* .1975^* \quad .1928^*$$

$$R^2 = .3723 \quad \bar{R}^2 = .2985 \quad F(2,17) = 5.0415^*$$

$$p = -.4577 \quad \text{multicol.} = .2678 \quad n = 20$$

$$\hat{WTO} = 59.7867 + .1036 WTO_{t-3} - .3140 NATO_{t-3} + 57.2249$$

$$18.6551^* .2578 \quad .2641$$

$$R^2 = .0772 \quad \bar{R}^2 = -.0314 \quad F(2,17) = .7111$$

$$p = -.4278 \quad \text{multicol.} = .2678 \quad n = 20$$

pseudo-GLS

$$\hat{NATO} = 25.1152 + .5014 NATO_{t-1} + .2682 WTO_{t-1} + 81.0228$$

$$27.3882 .1923^* \quad .1866$$

$$R^2 = .3812 \quad \bar{R}^2 = .3084 \quad F(2,16) = 5.2363^*$$

$$\text{multicol.} = .1224 \quad (n = 21)$$

TABLE 4-1--(continued)

$$\widehat{WTO} = 2.7966 + .2718 WTO_{t-1} - .0834 NATO_{t-1} \\ 1.3935 \quad .2556 \quad .2478$$

$$R^2 = .0767 \quad \bar{R}^2 = -.0259 \quad F(2,18) = .7476$$

$$\text{multicol.} = .1224 \quad n = 21$$

$$\widehat{NATO} = 23.4199 + .4938 NATO_{t-1} + .2224 WTO_{t-2} + 54.3735 \\ 17.6740 \quad .2070^* \quad .2022$$

$$R^2 = .3920 \quad \bar{R}^2 = .3160 \quad F(2,17) = 5.1579^*$$

$$\text{multicol.} = .3498 \quad n = 20$$

$$\widehat{WTO} = 12.9659 + .2546 WTO_{t-2} + .1482 NATO_{t-1} + 16.3743 \\ 5.5346^* \quad .2530 \quad .2608$$

$$R^2 = .0645 \quad \bar{R}^2 = -.0456 \quad F(2,17) = .5861$$

$$\text{multicol.} = .1995 \quad n = 20$$

$$\widehat{WTO} = 47.9995 + .3312 WTO_{t-1} - .2656 NATO_{t-3} + 60.7344 \\ 22.3837^* \quad .2212 \quad .2465$$

$$R^2 = .1610 \quad \bar{R}^2 = .0561 \quad F(2,16) = 1.5352$$

$$\text{multicol.} = .1519 \quad n = 19$$

$$\widehat{NATO} = 6.1637 - .0650 NATO_{t-2} + .4169 WTO_{t-3} + 7.8865 \\ 2.7374^* \quad .2328 \quad .2233$$

$$R^2 = .1797 \quad \bar{R}^2 = .0767 \quad F(2,16) = 1.8621$$

$$\text{multicol.} = .2452 \quad n = 19$$

$$\widehat{WTO} = 1.0189 + .1438 WTO_{t-3} - .2000 NATO_{t-1} + 23.3356 \\ .3926^* \quad .2722 \quad .2786$$

$$R^2 = .0358 \quad \bar{R}^2 = -.0847 \quad F(2,16) = .2970$$

$$\text{multicol.} = .3498 \quad n = 19$$

$$\widehat{NATO} = 83.1217 - .4818 NATO_{t-3} + .5158 WTO_{t-3} + 62.1871 \\ 21.4443^* \quad .2007^* \quad .1930^*$$

$$R^2 = .3939 \quad \bar{R}^2 = .3266 \quad F(2,16) = 5.8490^*$$

$$\text{multicol.} = .2678 \quad n = 19$$

2 2
OF/DE

TABLE 4-1--(continued)

$$\text{WTO} = 75.2041 + .1264 \text{ WTO}_{t-3} - .2559 \text{ NATO}_{t-3} + 81.1331$$

$$27.9775^* .2492 \quad .2674$$

$$R^2 = .0586 \quad \bar{R}^2 = .0669 \quad F(2,16) = .2970$$

$$\text{multicol.} = .2678 \quad n = 19$$

Reduction

OLS

$$\text{NATO} = 245.9812 - .2691 \text{ NATO}_{t-1} + .8967 \text{ WTO}_{t-1} + 99.3367$$

$$70.1226^* .2791 \quad .2803^*$$

$$R^2 = .6463 \quad \bar{R}^2 = .6091 \quad F(2,19) = 17.3589^*$$

$$p = -.1488 \quad \text{multicol.} = .9173^* \quad n = 22$$

$$\text{WTO} = 244.8057 + 1.1305 \text{ WTO}_{t-1} - .5193 \text{ NATO}_{t-1} + 104.3610$$

$$73.6990^* .2944^* \quad .2932^*$$

$$R^2 = .6420 \quad \bar{R}^2 = .6043 \quad F(2,19) = 17.0363^*$$

$$p = .4713 \quad \text{multicol.} = .9173^* \quad n = 22$$

$$\text{NATO} = 296.3940 + .0836 \text{ NATO}_{t-1} + .4459 \text{ WTO}_{t-2} + 130.1360$$

$$90.6829 .2993 \quad .2714$$

$$R^2 = .4433 \quad \bar{R}^2 = .3845 \quad F(2,19) = 7.5648^*$$

$$p = \text{---} \quad \text{multicol.} = .8737^* \quad n = 22$$

$$\text{WTO} = 213.7155 + .5790 \text{ WTO}_{t-1} + .0792 \text{ NATO}_{t-2} + 111.9950$$

$$77.1533^* .1999^* \quad .1681$$

$$R^2 = .5877 \quad \bar{R}^2 = .5433 \quad F(2,19) = 13.5415^*$$

$$p = \text{---} \quad \text{multicol.} = .7769^* \quad n = 22$$

$$\text{NATO} = 231.9015 - .0448 \text{ NATO}_{t-2} + .6902 \text{ WTO}_{t-1} + 101.5080$$

$$69.9287^* .1523 \quad .1811^*$$

$$R^2 = .6307 \quad \bar{R}^2 = .5918 \quad F(2,19) = 16.2243^*$$

$$p = .2056 \quad \text{multicol.} = .7769^* \quad n = 22$$

TABLE 4-1--(continued)

$$\hat{WTO} = 215.5311 + .0178 WTO_{t-2} + .6351 NATO_{t-1} + 112.6280$$

$$77.4898 * .2254 \quad .2497 *$$

$$R^2 = .5830 \quad \bar{R}^2 = .5391 \quad F(2,19) = 13.2818 *$$

$$p = .7158 \quad \text{multicol.} = .8737 * \quad n = 22$$

$$\hat{NATO} = 341.2756 - .0599 NATO_{t-2} + .5382 WTO_{t-2}$$

$$77.0449 * .3261 \quad .3501$$

$$R^2 = .4206 \quad \bar{R}^2 = .3596 \quad F(2,19) = 6.8963 *$$

$$p = -.0757 \quad \text{multicol.} = .9296 * \quad n = 22$$

$$\hat{WTO} = 308.5508 + .4094 WTO_{t-2} + .1029 NATO_{t-2} + 130.0780$$

$$78.8249 * .3582 \quad .3336$$

$$R^2 = .4438 \quad \bar{R}^2 = .3853 \quad F(2,19) = 7.5802 *$$

$$p = \text{---} \quad \text{multicol.} = .9296 * \quad n = 22$$

$$\hat{NATO} = 282.6597 + .3670 NATO_{t-1} + .2014 WTO_{t-3} + 118.4820$$

$$82.5099 * .1978 \quad .1617$$

$$R^2 = .4968 \quad \bar{R}^2 = .4445 \quad F(2,19) = 9.3699 *$$

$$p = .3293 \quad \text{multicol.} = .7423 * \quad n = 22$$

$$\hat{WTO} = 224.5492 + .4079 WTO_{t-1} + .2472 NATO_{t-3} + 103.0450$$

$$71.0456 * .1718 * \quad .1284$$

$$R^2 = .6509 \quad \bar{R}^2 = .6142 \quad F(2,19) = 17.7128 *$$

$$p = .4302 \quad \text{multicol.} = .7387 * \quad n = 22$$

$$\hat{NATO} = 232.6777 + .0507 NATO_{t-3} + .5987 WTO_{t-1} + 101.310$$

$$69.8495 \quad .1263 \quad .1689 *$$

$$R^2 = .6321 \quad \bar{R}^2 = .6127 \quad F(2,19) = 16.3222 *$$

$$p = .3163 \quad \text{multicol.} = .7387 * \quad n = 22$$

$$\hat{WTO} = 291.0364 + .3170 WTO_{t-3} + .2254 NATO_{t-1} + 128.4590$$

$$89.4576 * .1753 \quad .2145$$

$$R^2 = .4575 \quad \bar{R}^2 = .4004 \quad F(2,19) = 8.0115 *$$

$$p = -.1480 \quad \text{multicol.} = .7423 * \quad n = 22$$

TABLE 4-1—(continued)

$$\hat{NATO} = 375.0579 + .0572 NATO_{t-2} + .3741 WTO_{t-3} + 128.6290$$

$$73.6540^* .2825 \quad .2734$$

$$R^2 = .6321 \quad \bar{R}^2 = .6127 \quad F(2,19) = 16.3222^*$$

$$p = .2105 \quad multicol. = .9027^* \quad n = 22$$

$$\hat{WTO} = 316.7014 + .1261 WTO_{t-2} + .3860 NATO_{t-3} + 116.2490$$

$$70.5249^* .2104 \quad .1742^*$$

$$R^2 = .5557 \quad \bar{R}^2 = .5089 \quad F(2,19) = 11.8819^*$$

$$p = -.0322 \quad multicol. = .8281^* \quad n = 22$$

$$\hat{NATO} = 344.4910 + .1695 NATO_{t-3} + .3089 WTO_{t-2} + 124.5800$$

$$75.5991^* .1867 \quad .2255$$

$$R^2 = .4437 \quad \bar{R}^2 = .3851 \quad F(2,19) = 7.5771^*$$

$$p = \text{---} \quad multicol. = .8281^* \quad n = 22$$

$$\hat{WTO} = 334.1377 + .2924 WTO_{t-3} + .1848 NATO_{t-2} + 130.7230$$

$$74.8531 .2779 \quad .2871$$

$$R^2 = .4382 \quad \bar{R}^2 = .3791 \quad F(2,19) = 7.4099^*$$

$$p = + \quad multicol. = .9027^* \quad n = 22$$

$$\hat{NATO} = 382.0283 + .1255 NATO_{t-3} + .2951 WTO_{t-3} + 128.2660$$

$$68.9585^* .3255 \quad .3544$$

$$R^2 = .4103 \quad \bar{R}^2 = .3482 \quad F(2,19) = 6.6099^*$$

$$p = .8016 \quad multicol. = .8921^* \quad n = 22$$

$$\hat{WTO} = 361.1050 - .2883 WTO_{t-3} + .7224 NATO_{t-3} + 130.7230$$

$$61.7584^* .3174 \quad .2915^*$$

$$R^2 = .5662 \quad \bar{R}^2 = .5187 \quad F(2,19) = 12.3994^*$$

$$p = -.7528 \quad multicol. = .8921^* \quad n = 22$$

pseudo-GLS

$$\hat{NATO} = 282.5835 - .2691 NATO_{t-1} + .8967 WTO_{t-1} + 114.118$$

$$80.5569^* .2791 \quad .2803^*$$

$$R^2 = .6463 \quad \bar{R}^2 = .6091 \quad F(2,18) = 17.3589^*$$

$$multicol. = .9173^* \quad n = 21$$

TABLE 4-1—(continued)

$$\hat{WTO} = 129.4287 + 1.1305 WTO_{t-1} - .5193 NATO_{t-1} + 55.1755$$

$$39.9488* .2944 \quad .2932$$

$$R^2 = .6420 \quad \bar{R}^2 = .5938 \quad F(2,18) = 13.3166*$$

$$\text{multicol.} = .9173* \quad n = 21$$

$$\hat{NATO} = 184.2228 - .0448 NATO_{t-2} + .6902 WTO_{t-1} + 80.6377$$

$$55.5513* .1524 \quad .1811*$$

$$R^2 = .6306 \quad \bar{R}^2 = .5923 \quad F(2,18) = 16.4625*$$

$$\text{multicol.} = .7769* \quad n = 21$$

$$\hat{WTO} = 198.8591 + .4459 WTO_{t-2} + .0836 NATO_{t-1} + 87.3083$$

$$60.8390* .2714 \quad .2993$$

$$R^2 = .4433 \quad \bar{R}^2 = .3847 \quad F(2,18) = 7.5648*$$

$$\text{multicol.} = .8737* \quad n = 21$$

$$\hat{NATO} = 367.1101 - .0599 NATO_{t-2} + .5382 WTO_{t-2} + 136.7650$$

$$82.8772* .3261 \quad .3501$$

$$R^2 = .4206 \quad \bar{R}^2 = .3482 \quad F(2,18) = 6.8963*$$

$$\text{multicol.} = .9296* \quad n = 21$$

$$\hat{NATO} = 189.5798 + .3670 NATO_{t-1} + .2014 WTO_{t-3} + 79.4660$$

$$55.3394* .1978 \quad .1617$$

$$R^2 = .4968 \quad \bar{R}^2 = .4438 \quad F(2,18) = 9.3792*$$

$$\text{multicol.} = .7423* \quad n = 21$$

$$\hat{WTO} = 127.9477 + .4079 WTO_{t-1} + .2472 NATO_{t-3} + 58.7149$$

$$40.4819* .1718* \quad .1284$$

$$R^2 = .6509 \quad \bar{R}^2 = .6151 \quad F(2,18) = 17.1284*$$

$$\text{multicol.} = .7387* \quad n = 21$$

$$\hat{NATO} = 159.0819 + .0507 NATO_{t-3} + .5987 WTO_{t-1} + 69.2654$$

$$47.7561 .1263 \quad .1689*$$

$$R^2 = .6321 \quad \bar{R}^2 = .5934 \quad F(2,18) = 16.3222*$$

$$\text{multicol.} = .7387* \quad n = 21$$

TABLE 4-1—(continued)

$$\hat{WTO} = 334.1096 + .3170 WTO_{t-3} + .2254 NATO_{t-1} + 147.4710 \\ 102.6973 * .1753 \quad .2145$$

$$R^2 = .4575 \quad \bar{R}^2 = .3954 \quad F(2,18) = 8.0115*$$

$$\text{multicol.} = .7423 * \quad n = 21$$

$$\hat{NATO} = 296.1084 + .0572 NATO_{t-2} + .3741 WTO_{t-3} + 101.5530 \\ 58.1498 * .2825 \quad .2734$$

$$R^2 = .4069 \quad \bar{R}^2 = .3444 \quad F(2,18) = 6.5175*$$

$$\text{multicol.} = .9027 * \quad n = 21$$

$$\hat{WTO} = 326.8997 + .1261 WTO_{t-2} + .3860 NATO_{t-3} + 119.9920 \\ 72.7957 * .2104 \quad .1742 *$$

$$R^2 = .5557 \quad \bar{R}^2 = .5089 \quad F(2,18) = 11.8819*$$

$$\text{multicol.} = .8281 * \quad n = 21$$

$$\hat{NATO} = 75.7945 + .1256 NATO_{t-3} + .2951 WTO_{t-3} + 25.4480 \\ 13.6813 * .3255 \quad .3544$$

$$R^2 = .4102 \quad \bar{R}^2 = .3481 \quad F(2,18) = 6.6072*$$

$$\text{multicol.} = .8921 * \quad n = 21$$

$$\hat{WTO} = 632.9445 - .3883 WTO_{t-3} + .7224 NATO_{t-3} + 201.3510 \\ 108.2502 * .3174 \quad .2915 *$$

$$R^2 = .5662 \quad \bar{R}^2 = .5205 \quad F(2,18) = 12.3905*$$

$$\text{multicol.} = .8921 * \quad n = 21$$

* Significance = .05.

insignificant. Thus, whatever substantive results may be offered cannot be said to be based upon good estimates of the 1961 Berlin crisis model. With this caveat in mind, presentation of results may be given.

The simplest criterion of substantive meaning is goodness-of-fit. Which phase and alliance best fits the model? Which particular lag structure of a lagged endogenous variable (Y_{t-n}) and exogenous variable (X_{t-n}) best fits a phase and alliance? The first question is rather obvious—we are finding out if anything is being explained. The second question, however, raises more points. Is there a shift in lag structure as we proceed from intensification to reduction? As discussed in chapter two (p. 36) if, in fact, we expect that the interaction parameter is the most important during intensification, then we should expect that the best fitting equation here would have the interaction parameter occurring before the organizational parameter. Obviously, similar logic applies to reduction with respect to the organizational parameter. This temporal propinquity applies to theoretically dominant parameters. This same principle may also be applied to the style of the alliances. By examining the type of lag structure that best explains WTO and NATO, we may see if WTO is more bureaucratic by having an organizational parameter that is closer to the present in its best fitting equations than NATO's.

Examining the OLS $\bar{R}^2$ s with respect to the first question reveals that: the model explains NATO far better during intensification than WTO; reduction sees WTO explained slightly better than NATO;

NATO overall fits best to the model. Explained variance ranges from the extremely low (negative $\bar{R}^2$ s) to moderately good ($\bar{R}^2 \approx .60$). Substitution of the pseudo-GLS results does not change the findings. Essentially, then, the model best fits NATO moderately well, with WTO being explained in reduction.

Examining the OLS $\bar{R}^2$ s with respect to the relationship between phase, alliance, and temporal propinquity does not depend upon one equation because many equations seem to provide very similar good explanations. Therefore, the equations that are among the best fitting are considered. For NATO, during intensification, the OLS $\bar{R}^2$ s reveal that the best fitting equations have lag structures (Y_{t-n}, X_{t-n}) of 1, 3 and 3, 3. Since the best lag structure is a tie, it may be ignored, showing that interaction does not have temporal propinquity. Substitution of the pseudo-GLS $\bar{R}^2$ s adds lag structure 1,1 which leaves the substantive result the same. For WTO, during intensification, the OLS $\bar{R}^2$ s reveal only two equations with positive (but low) fits with lag structures of 1,3 and 3,3. This structure is precisely the same as NATO's and is not altered by the pseudo-GLS $\bar{R}^2$ s. For NATO, during reduction, the OLS $\bar{R}^2$ s reveal four equations with lag structures of 1,1; 2,1; 3,1; and 2,3. These lags indicate that the organizational parameter does not have temporal propinquity. Substitution of pseudo-GLS $\bar{R}^2$ s eliminating lag structure 2,3 which strengthens the preceding result. For WTO, during reduction, the OLS $\bar{R}^2$ s reveal two equations with lag structure 1,1 and 1,3. Here, the organizational parameter does have temporal propinquity. This result remains unchanged after substituting the

pseudo-GLS $\tilde{R}^2$ s. Using the preceding results to see if there is a style to the various alliances shows that NATO switches from organization temporal propinquity during intensification to interaction temporal propinquity during reduction, while WTO has consistent organizational temporal propinquity.. What has been seen here is that neither NATO nor intensification conforms to expectation in temporal propinquity, only WTO does so. What is intriguing, however, is that NATO and WTO reveal the same lag structure during intensification which may mean that their escalatory processes were so similar that a tit for tat scenario seems suggestive. The consistent organizational temporal propinquity of WTO, seems to parallel Tanter's (1974:72) contention that organization should be larger than interaction.

There is one final substantive consideration, the dominance of parameters across the two phases. Here, as shown in table 4-2, two null hypotheses are entertained: for intensification that $Y_{t-n} \geq X_{t-n}$; for reduction that $X_{t-n} \geq Y_{t-n}$. Examination of table 4-2 reveals that for both OLS and pseudo-GLS during intensification the alternative hypothesis cannot be accepted. Here, however, NATO comes closest of the alliances to making the alternative acceptable with a 55.55 per cent rate of failing to reject the null hypothesis. Higher rates of accepting the null hypothesis appear during reduction. Overall there would appear to be little evidence for the hypothesized parameter dominance; only NATO during intensification appear to show some small glimmer of parameter dominance.

Substantively, then, this replication has shown that only

TABLE 4-2—Dominance of parameters

Intensification					
NATO		OLS		pseudo-GIS	
LAG	STRUCTURE	H_0 accepted by inspection		H_0 accepted by inspection	
X_{t-n}	Y_{t-n}	by testing		by testing	
1	1	Yes		Yes	
1	2	Yes			
2	1	No	Yes; t=4.189		
2	2	No	Yes; t=1.6761		
1	3	Yes		Yes	
3	3	No	No; t=2.4457*		No; t=2.1127*
2	3	No	No; 5=2.1238*	No	
3	2	No	No; t=3.4681*		No; t=5.0665*
3	3	No	No; t=5.0994*	No	
H_0 accepted		33.33	55.55	33.33	55.55

TABLE 4-2--(continued)

WTO	LAG	STRUCTURE	OLS		pseudo-GLS	
	X_{t-n}	Y_{t-n}	H_0 accepted by inspection	H_0 accepted by testing	H_0 accepted by inspection	H_0 accepted by testing
	1	1	Yes		Yes	
	1	2	Yes			
	2	1	Yes		Yes	
	2	2	Yes			
	1	3	Yes		Yes	
	3	3	Yes		Yes	
	2	3	Yes			
	3	2	No	Yes; $t = .3715$		
	3	3	Yes		Yes	
% H_0 accepted			88.88	100	88.88	100

TABLE 4-2--(continued)

Reduction		LAG	STRUCTURE	Y_{t-n}	OLS		pseudo-GLS	
NATO	X_{t-n}				H_0 accepted by inspection	H_0 accepted by testing	H_0 accepted by inspection	H_0 accepted by testing
1	1			1	Yes		Yes	
1	1			2	Yes			
2	2			1	Yes		Yes	
2	2			2	Yes		Yes	
3	3			3	No	Yes; t = .9210	No	Yes; t = .8372
3	3			1	Yes		Yes	
3	3			3	Yes		Yes	
3	3			2	Yes			
3	3			3	Yes		Yes	
% H_0 accepted					88.88	100	88.88	100

TABLE 4-2--(continued)

WTO	LAG	STRUCTURE	OLS	H ₀ accepted by inspection	H ₀ accepted by testing	pseudo-GLS	H ₀ accepted by inspection	H ₀ accepted by testing
X_{t-n}	Y_{t-n}							
1	1	1	No	No	No; t=5.6154*	No	No	No; t=5.5235*
1	2	2	No	No	No; t=2.7613*	No	No	Yes; t=1.2694
2	1	1	Yes	Yes	Yes; t=.8861	No	No	Yes; t=1.0706
2	2	2	No	No	Yes; t=1.0706	No	No	Yes; t=.4700
1	3	3	No	No	Yes; t=.4700	Yes	Yes	
3	1	1	Yes	Yes	Yes; t=.3809	Yes	Yes	
2	3	3	No	No				
3	2	2	Yes	Yes				
3	3	3						
% H ₀ accepted			33.33	77.77	22.22	77.77		

Note: Percentage of H₀ accepted for OLS by testing and all pseudo-GLS represent the proportion obtained by substituting these results for previous results.

* Significance = .05

NATO seems to conform to the model, and that temporal propinquity and parameter dominance are contrary to expectation. Moreover, WTO does seem to react first to its own past behavior, indicating the presence of a strong bureaucratic tendency. The only puzzle that is raised is the NATO-WTO tit-for-tat scenario during an intensification that did not see temporal propinquity or parameter dominance of interaction. If we are willing to abide with the acceptance of the null hypotheses, perhaps the most appropriate scenario is not classic escalation but a controlled bureaucratic attempt to extend influence with little interest in what the opponent was doing. This lack of interest is suggested by the acceptance of the null hypothesis for parameter dominance during intensification. The Berlin 1961 crisis may well be an example of an international incident that had become routinized. The reduction phase lack of appropriate temporal propinquity and parameter dominance would seem to show that it is possible to deescalate an international crisis in much the same manner as they are supposedly escalated. The mechanisms underlying the 1961 Berlin crisis are mixed with a bureaucratic process controlling escalation and a more interactive (GRIT?) process appearing in reduction. The Berlin 1961 crisis was not the consistent East-West confrontation that Tanter (1974:178) depicts.

Comparison of Tanter's and replicatory results

Comparison of Tanter's and replicatory results proceeds in much the same manner as the preceding discussion with three exceptions.

First, comparison of temporal propinquity of parameters is not attempted since Tanter did not consider this possibility. The overall similarity of coefficients obtained by Tanter and the replication is considered. Third, specific equations obtained by Tanter and compared to their replicatory OLS and pseudo-GLS estimates which may be considered in the same light since they are similar. The comparison begins with Tanter compared to the set of replicatory equations and ends with consideration of the specific equations.

The first question that may be asked here is the same as that which started the preceding discussion: how well does the model fit each phase and each alliance? If similarities are found, then Tanter and this replication may have the same pattern of crude explanation. At best, the dissimilarity of the data throughout and difficulties with Tanter's results (provided in ch. 2) make a case for measurement error or bad research design. Examining for the best fit reveals that WTO fits best during intensification, and NATO fits best during reduction. Considered in the light of the replicatory OLS and pseudo-GLS results, which are the same, only the first observation is similar. Here, a case for measurement error or research design may be made. The rest of the observations, however, do not conform at all to the replication and no generalizability can be said to exist. Tanter's pattern of crude explanation does not exist in this replication.

The next question that may be asked is whether the same sort of structure exists in Tanter and the replication. This may be

examined by seeing if there is any difference between replicatory and Tanter's coefficients. Presumably, the two sets of coefficients should have the same sign, so a one-tailed test is employed. Here, since Tanter did not provide standard errors, tests of the null hypothesis were carried out with respect to the replicatory standard errors, not the pooled replicatory and Tanter standard errors.

The tests, then, presented in table 4-3, are not as robust as they ought to be. Inspection of the results reveal the following: for NATO during intensification, whether by OLS or pseudo-GLS, no similarity of parameters seems forthcoming; results for WTO are more mixed. During intensification, WTO, whether by OLS or pseudo-GLS, rejects the null hypothesis for the lagged endogenous parameter most of the time but does not do so for the exogenous parameters. During reduction, OLS shows that the null hypothesis for either parameter may be accepted on 44.44 per cent of the time, while pseudo-GLS lowers this to 33.33 per cent. Overall, then, replicatory WTO seems to show much the same sort of structure as did Tanter. Generalizability does not, of course, exist, but, the confounding effects of measurement error and bad research design may be present. The results of the structure of the data are more encouraging than those of crude explanation.

Still concerned with the parameters, the most specific question that may be asked is whether the expected pattern of parameter dominance prevails. Once again, Tanter's lack of standard errors make judgment of the two sets of results difficult; we can never have confidence that the differences between Tanter's parameters would

TABLE 4-3—Comparison of Tanter's and replicatory coefficients

Intensification	OLS		pseudo-GLS	
	H_0 accepted	X_{t-n}	H_0 accepted	X_{t-n}
NATO	Y_{t-n}	X_{t-n}	Y_{t-n}	X_{t-n}
	1 Yes; t=3.6087*	1 Yes; t=3.6201*	1 Yes; t=3.5434*	1 Yes; t=4.2776*
	1 Yes; t=2.8699*	2 Yes; t=4.0352*	1	2
	2 No; t=.9865	1 Yes; t=3.0339*	2	1
	2 No; t=.6508	2 Yes; t=3.9702*	2	2
	1 Yes; t=3.1119*	3 Yes; t=3.8880*	1 Yes; t=3.2551*	3 Yes; t=3.7211*
	3 Yes; t=.6528	1 Yes; t=3.5536*	3	1
	2 No; t=.6299	3 Yes; t=4.4076*	2 No; t=.4940	3 Yes; t=4.2405*
	3 Yes; t=.7295	2 Yes; t=4.4777*	3	2
	3 Yes; t=3.5434*	3 Yes; t=5.6266*	3 Yes; t=.8516	3 Yes; t=4.4187*
	% H_0 accepted	66:66	66:66	100
		100		

TABLE 4-3--(continued)

WTO

		OLS		pseudo-GLS	
		H_0 accepted		H_0 accepted	
Y_{t-n}	X_{t-n}	Y_{t-n}	X_{t-n}	Y_{t-n}	X_{t-n}
1 Yes; t=1.8203*	1 Yes; t=1.8124*	1 No; t=1.5329	1 Yes; t=-1.8701*		
1 No; t=-.6394	2 Yes; t=1.0049	1	2		
2 Yes; t=1.5987	1 Yes; t=-1.9431*	2 No; t=1.4806	1 Yes; t=-.8888		
2 No; t=1.4269	2 Yes; t=-1.1682	2	2		
1 Yes; t=1.8114*	3 Yes; t=-2.884*	1 Yes; t=2.0398	3 Yes; t=-2.6191*		
3 No; t=-.6795	1 Yes; t=-1.7959*	3 No; t=-.9691	1 Yes; t=-2.0818*		
2 No; t=1.5697	3 Yes; t=-2.7836*	2	3		
3 No; t=-.4447	2 No; t=-.0211	3	2		
3 No; t=-.8673	3 Yes; t=-2.6278*	3 No; t=-.9887	3 Yes; t=-2.3781*		
% H_0 accepted. 33.33	88.88	33.33	88.88		

TABLE 4-3—(continued)

Reduction

NATO

OLS

pseudo-GLS

H₀ acceptedH₀ acceptedY_{t-n}X_{t-n}Y_{t-n}X_{t-n}

1 Yes; t=-1.7524*
 1 Yes; t=-.4557
 2 Yes; t=-1.7387*
 2 Yes; t=-.8461
 1 No; t=-.7431
 3 Yes; t=-1.3405
 2 Yes; t=-.5763
 3 Yes; t=-.2705
 3 Yes; t=-.2919

1 No; t=1.2726
 2 Yes; t=-.3099
 1 No; t=.8294
 2 No; t=-.0000
 3 Yes; t=-2.0191*
 1 No; t=.3475
 3 Yes; t=-.6068
 2 Yes; t=-1.0248
 3 Yes; t=-.69103

1 Yes; t=-1.7524*
 1 Yes; t=-1.7375*
 2 Yes; t=-.8583
 1 No; t=-.7432
 3 Yes; t=-1.3405
 2 Yes; t=-.5763
 3 Yes; t=-.2974

1 No; t=1.2725
 2 No; t=.8923
 2 No; t=-.0000
 3 Yes; t=-2.0940*
 1 No; t=.3475
 3 Yes; t=-.6068
 2 Yes; t=-.6910

% H₀ accepted 88.88

55.55

88.88

55.55

TABLE 4 3—(continued)

WTO

OLS		pseudo-GLS	
H ₀ accepted		H ₀ accepted	
Y _{t-n}	X _{t-n}	Y _{t-n}	X _{t-n}
1 Yes; t=3.1548*	1 Yes; t=-2.1630*	1 Yes; t=2.9908*	1 Yes; t=-2.1872
1 No; t=1.0992	2 No; t=1.4587	1	2
2 Yes; t=-1.0302	1 Yes; t=2.3432*	2 No; t=-.7218	1 No; t=-.1123
2 No; t=-.4450	2 No; t=-.1586	2	2
1 No; t=-.9191	3 No; t=1.5358	1 No; t=-.9191	3 No; t=1.5358
3 No; t=-.3822	1 No; t=-.8143	3 No; t=-.3822	1 No; t=-.8177
2 Yes; t=-.5889	3 Yes; t=1.9288*	2 Yes; t=-.5889	3 Yes; t=1.9288*
3 No; t=-.1526	2 No; t=-.4695	3	2
3 Yes; t=-1.6960	3 Yes; t=2.3067*	3 Yes; t=-1.6960	3 Yes; t=2.3067*
% H ₀ accepted 44.44	44.44	33.33	33.33

Note: Percentage pseudo-GLS H₀ accepted components proportion obtained by substituting pseudo-GLS results for OLS results.

*. Significance = .05 one-tailed.

be statistically significant at any standard criterion level.

Examination of Tanter's equations reveal that, during intensification the lagged endogenous parameter is dominant in WTO but not so for NATO; during reduction the exogenous parameter is dominant in WTO but not so for NATO.

These mixed results are matched by the replication for NATO in all phases. During intensification, WTO results are dissimilar. During reduction, however, the replicatory test by inspection yields the same result as Tanter's but is erased by formal testing. At best, similarities exist three-fourths of the time; at worst, one-half of the time. Thus, whether under generous or stringent conditions, sufficient similarities are found to question once again the replication, measurement procedure and research design. This general pattern appears in the majority of replicatory questions assessed by comparing general replicatory results to Tanter's.

Consideration may now be given to specific equations. (see table 4-1). Once again, the pattern of crude explanation is the starting point. For either OLS or pseudo-GLS, during intensification, replicatory NATO $\bar{R}^2$ for lag structure 1,2 is in excess of Tanter's; the replicatory WTO $\bar{R}^2$ for lag structure 1,3 is very close to Tanter's; the replicatory WTO $\bar{R}^2$ is greater than NATO's. During reduction, however, when both Tanter's alliances have a lag structure of 1,1, the appropriate replicatory $\bar{R}^2$ s are well in excess of Tanter's. Moreover, there is little difference between replicatory $\bar{R}^2$ s of either alliance. Across phases, the replicatory WTO $\bar{R}^2$ is slightly greater than NATO's. Those results are different than the general

results and are also more similar to Tanter's results (except for reduction. Thus, measurement error or bad research design in the replication may be suspected.

With regard to the structure of the models, we find that NATO during intensification is dissimilar as is WTO for the same period. During reduction, WTO is dissimilar, while only NATO's exogenous variable during this period shows similarity. This is contrary to what was found in the general case with there being no possible appeal to measurement error or bad research design.

Parameter dominance for specific equations shows replicatory WTO during reduction rejecting the null hypotheses and NATO during intensification accepting the null hypothesis. Other phase and alliance combinations are dissimilar. Similar results are forthcoming one half of the time. Therefore, measurement error or bad research design appears in the replication. Indeed, this possibility arises more often than it did when the general set of replicatory models were considered.

We are faced with a similar set of replicatory interpretations. Under the most general conditions, problems in replication is a definite possibility. When specific equations are compared, potential error of measurement and research design also appear as a genuine consideration. Choosing between similar results depends upon how confidently one set of results may be accepted over another. Despite the one-to-one correspondence of the specific equations, they may be an unreliable technique of assessing replicability because they represent a very strict interpretation of the validity of the

data generation procedures and indeed the supposed similarity of purpose in forming the two sets of event data. As noted in chapter three, such a strict interpretation is unwarranted: recoding CADIC event data into an approximation of East-West event data was not entirely satisfactory; nor were the two event data sets collected with the same purpose. Admitting immediately to measurement error precludes consideration of specific equations. Presumably, the OLS procedures used in this replication were similar to Tanter's and so this aspect of the research design cannot be considered to be different (pseudo-GLS yielded the same general results). Replication, then, ought to be seen in terms of the general results. Thus, bearing in mind the problems of measurement, measurement error and bad research design are highly probable—dissimilar data representing the same situation yielded similar results.

Summary

This chapter began by briefly demonstrating a plausible procedure for estimating autoregressive models that was applicable to the majority of the estimated equations. Results of the replicatory models were then presented. Here, the null hypotheses were accepted. The model that Tanter specified did not appear. Instead a model of a bureaucratically directed intensification and an interactive reduction seemed suggested, contrary to all expectation. This general dissimilarity between the replication and Tanter did not continue with comparison of the general sets of replicatory results and Tanter. Consideration of specific equations led to the same conclusion.

The different substantive interpretation and similar enough statistical results is primarily a product of strictness of interpretation. With substantive results, the null hypothesis accepted one-half of the time was sufficient to change the alternative hypothesis; whereas, with replication, the same level of accepting the null hypotheses still permitted sufficient similarity to be seen. Moreover, Tanter's results, showing temporal propinquity of the organizational parameter during intensification, no temporal propinquity during reduction, and expected parameter dominance occurring half of the time—could lead to the same substantive conclusion offered for the replicatory results. Thus, overall similarity may exist. Although problems of statistical inference exist with respect to Tanter's results, we must interpret the results of the replication to mean that there is considerable measurement error or bad research design (despite attempts to reasonably imitate Tanter) that do not permit successful replication. Successful replication may not be possible in the contemporary study of international politics.

CONCLUSIONS AND FURTHER SUGGESTIONS

Overview

This thesis has raised more puzzles than it solved. Therefore, this conclusion brings together the various questions and points that have been raised, and makes suggestions based upon them. The most substantive area is dealt with first: the findings of the replicatory model. Next, based upon the critique of Tanter's model in chapter two and the interpretation of the replicatory results in chapter four, an alternative model is proposed. After the two preceding 'substantive' issues have been dealt with, 'theoretical' issues come to the fore. The first of these issues is the case of time-series regression in international politics. Here, minimal standards of inference and reporting are suggested. Next, as part of a review of the replicatory findings, questions of event data measurement are raised. As noted in chapter four, these issues are vital to interpreting the replicability of Tanter's results. Here, simple statistical inference about the similarity of data sets is insufficient; the event data analyst must impose a priori subjective probabilities upon what the two data sets are supposed to represent. In this way, variation between sources may be more easily attributable and replication ameliorated. Finally, replication as a technique leading to scientific accumulation is discussed. Based on the experience of this thesis, it is suggested that replication does not lead to the clear acceptance

or rejection of findings, but leads instead to a focus upon the technical barrier to such clear cut acceptance or rejection. This is a valuable critical function that points to research areas that must be further integrated into the study of international politics. This is best accomplished through deliberate creation of a replicatory research note, which is discussed at the end of this conclusion.

On the results of the replicatory models

The results of the replicatory models are contrary to expectations. By accepting the null hypotheses throughout the interpretation of the replicatory models, it would appear that escalation during the 1961 Berlin crisis was primarily a bureaucratic process, whereas the reduction phase saw more interaction. If we are willing to accept that greater bureaucratic impact implies 'business as usual', then we must conclude that the Berlin 1961 crisis was not unusual enough to warrant great amounts of interaction. Both NATO and WTO, and presumably the bloc leaders as well, were preset to autonomously manage an international crisis. Winding down an international crisis proves more difficult for each alliance alone; they must therefore negotiate by slowly decreasing the level of hostility accorded to one another in an interactive fashion. The 1961 Berlin crisis, then, displays the heritage of the extreme bipolarity of the preceding decade: it is easier to maintain a crisis than to break out of it. In this light, Berlin 1961 may be seen behaviorally as a classic case of Cold War conflict management.

An alternative model

The above interpretations of the replication's substantive conclusions is not based upon a model that lends itself well to testing the suggested alternative model. Based, however, upon the suggestions regarding the original model in chapters two and four, an alternative model may be suggested. The bases of the model are the following: The first assumption is that there is no adequate interpretation of the impact of past behavior by examining only one set of their values at a time. Dominant parameters cannot be equal to or less than zero. Finally, dominant parameters should have greater temporal propinquity than subordinate parameters. Placing these three assumptions together yields a model that would test sets of lagged endogenous and exogenous parameters for parameter dominance and temporal propinquity. One way of representing this model is through a polynomial distributed lag (Almon lag).

A polynomial distributed lag (see Kmenta, 1971; 492-494 for a relatively simple explanation) is a technique whereby a set of lagged parameters (X_{t-1} , X_{t-2} , X_{t-n}) may be weighted to a polynomial function. The simplest function here, of course, is the parabola. Considering, then, the parabola for each set of variables, the following null hypotheses with regard to intensification for the original model may be derived. Only Y_{t-n} is greater than or equal to its temporally corresponding X_{t-n} ($Y_{t-1} \geq X_{t-1}$, $Y_{t-2} \geq X_{t-2}$, ..., $Y_{t-n} \geq X_{t-n}$). All exogenous parameters are less than or equal to zero ($X_{t-1} \leq 0$, $X_{t-2} \leq 0$, ..., $X_{t-n} \leq 0$). The peak value

of Y_{t-n} precedes or occurs at the same time as X_{t-n} ($Y_{t-n, \max} \geq X_{t-n, \max}$).

If these null hypotheses are rejected, then a graph similar to figure 6-1 may be obtained. During reduction, the null hypotheses are reversed and if rejected may yield a graph similar to figure 6-2. This sort of model adds to the information that may be gained because it is more completely specified than Tanter's.

Practical difficulties may arise in estimating this model due to the problem of uneven lags between the two sets of parameters. Usually, with models of this sort, lagged variables are added until no new information is added (seen through little change in the successive $\bar{R}^2$ s). If this were to be carried out naively, an embarrassing situation where there are more subordinate parameters than dominant parameters, implying that the subordinate parameters carry more information and are hence more important, may arise. Several solutions are possible. First, add pairs of parameters only and reject any pair that does not add information. Even sets of parameters may thus be obtained. Second, admit to the possibility of having more subordinate than dominant parameters but test the null hypotheses that the additional parameters are individually less than the last of the dominant parameters. In this way, it may be possible to say that the additional parameters do not have as much impact upon behavior as the last dominant parameter. Finally, admit to the additional subordinate parameter and accept the possible consequences of claims that overall the explained behavior seems more responsive to a supposedly less important set of variables. Despite these potential difficulties, it is possible to formulate

Fig. 6.1—Graph of polynomial distributed lags for intensification

Fig. 6.2—Graph of polynomial distributed lags for reduction.

a more attractively specified model than Tanter's that is deserving of investigation.

On the use of time-series

This thesis, like Tanter and many other modern works in international politics, depends upon time-series analysis. Most political scientists are still probably unfamiliar with time-series and are apt to make mistakes and false inferences. Such realization should not paralyze political scientists but should urge them on to acquire more time-series analysis acumen and to leave enough information about what they attempted so that effective methodological revision may be forthcoming at a later date. Toward this second goal, the political scientist must be aware of fundamental inferential problems, most notably autocorrelation, and problems of specific models, such as the autoregressive model. This thesis suggested simple procedures to mitigate these problems. Hopefully, these procedures, which are not necessarily the most effective, are described well enough that future analysts may easily improve upon them. This is not a matter of science, but of good scholarship.

In a related vein, reportorial standards of results must be adequate. If an analyst omits standard errors, F-statistics, and thorough discussion of functional form, future analysts are at a disadvantage. The most serious instance of this in the replication was Tanter's omission of the standard errors which made retrieval of parameter dominance results and similarity of coefficients far less robust than they ought to have been. At the very least, then, for the sake of improving inference and replicability, students of

international relations ought to report what inferential problems may arise and how they are to be handled; give as complete a specification of the model as possible; report fully all necessary statistical information pertaining to the regression coefficients (the standard errors) and the equation (R^2 , $\bar{R}^2$, and F). Minimal reporting standards along these lines ought to be established.

Measurement and replication

Replication depends upon measurement. The more successful we are measuring phenomena, the easier and more conclusive replication. Recall from chapter four that measurement error rather than estimation procedures seemed the most likely source of the dissimilar data producing similar results. There can be no doubt that ordinary statistical inference leads to the conclusion that the two data sets do not covary together. For all alliances, phases, aggregations of time, measurement and transformations, there is no significant positive slope between the two data sets. The recoding and scaling of CADIC event data to resemble East-West event data may have caused some of the problem along with being forced to extract an approximation of what Tanter used. There are sound reasons to believe that the creation of the two data bases had sufficient errors of measurement that may have erased any similarity. Part of the error of measurement may be accounted for in this way.

Another potential source of measurement error is the nature of the two data sets. The East-West event data set was collected to assess the pattern of behavior of the blocs over an extended period of time using several sources. The CADIC event data, on the other

hand, is limited to the Berlin 1961 crisis, does not refer to behavior of the local antagonists (East and West Germany), and is drawn from one source (see ch. 3, p. 54). The context in which the coders were working in each event data project was different. The subjective probability of the two event data sets being similar is low. Despite the similarity of both event data sets' categories to WEIS, despite similar structure of measurement devices (both data sets showed significant positive covariation between scale values and frequencies under all conditions), event extraction and coding in both data sets were probably different enough to create the dissimilarities.

These points must, of course, be investigated further. Should further investigation show that the above speculations are true, then, in retrospect, a known source of variation in the setting may be said to exist. If this should be the case, then some of Tanter's results become generalizable.

Replication as criticism

The major question of this thesis has been the relationship between scientific accumulation and replication. Ideally, replication could determine if certain models are generalizable or not. Often, as was the case in this thesis, the results are mixed: inaccuracies of replication do not permit arguing completely for or against generalizability. If such a situation could be avoided, then it might be possible to validate what knowledge has been gained or, more likely, reduce the number of verified models to a parsimonious and easily manipulable few (an ideal situation). Failure of a

replication to attain either extreme is not complete, for the replication's critical value is still extremely important.

What critical value, then, may a replication leave? There is the valuable process of reconstructing the other analyst's logic. This tells the replicator what the original analyst intended to do and whether it was properly attempted. This covers the area referred to earlier (see ch. 1) as replicability and also extends into discussion of specification. Thus, the replicator shows to all the world whether the original study could inspire further coherent scientific research that could build up into a scientifically accumulative enterprise. Discovery of potentially more stringent specification is simply furthering the development and testing of the model—yet another aspect of advancing the scientific endeavour.

On a more mundane level, replication also has definite benefits for the development of methodological techniques. Consider briefly, the range of methodological techniques raised by this thesis: reporting standards; minimal inferential considerations, measurement systems. Replication acts as a link between a growing substantive area and its 'mere methodological' concerns.

Overall, then, replication's critical functions are valuable. First, replication criticizes and fosters an area of substantive research. Second, it gives purpose to 'mere methodology' and makes it an integral part of the research enterprise. Yet, despite these benefits, replication is not popular. To a large extent, replication's unpopularity is due to a bias against it in the study of international politics. To a lesser extent, replication is held back by difficulties

of reporting.

Replication need not, however, be difficult to report. Wilson et al. (1973:147) wisely suggest that the would-be replicator label the effort as a replication; discuss the research being analyzed; bring forward the differences in the replication; which in this thesis' case would include measurement of differences between data and the creation of data; discussion of replicatory results (arguing for or against generalizability on invoking problems in replicatory design), and perhaps an evaluation of the implications of the replication. According to the authors, this format may be contained within the space of a research note. Hopefully, more research notes will become replications and more students of international politics will be aware of them. Otherwise, the study of international politics will still be maintaining a great barrier to building up critical, self-correcting, ongoing research. Obviously, the current low status of replication in the study of international politics is a sad testimonial to the feelings of misplaced effort that have recently swept the discipline.

Final summation

This conclusion has sought to bring together all of the suggestions that have been inherent in this thesis. First, then, the results of the replicatory model were seen, suggesting a conflict management model of the Berlin 1961 crisis. Second, based upon the limitations of Tanter's model, a more appealing and informative polynomial distributed lag model was specified for the crisis.

Given the difficulties of inference, minimal standards for time series research and reporting (for the present stage of the field) were given. Event data measurement problems also received attention with the suggestion that the context of event data sets is also a vital consideration on assessing their similarity. Finally, the critical role of replication, as ongoing criticism, in scientific accumulation was discussed and a palatable replication format was suggested. Although a definite decision on Tanter's model is not forthcoming, this thesis has made a strong case for replication's critical role in scientific accumulation.

APPENDIX A

DATA USED IN THIS THESIS

The following tables contain all data used in this thesis. In order of appearance the tables are:

- A-1.—Tanter's WTO data daily;
- A-2.—Tanter's NATO data daily;
- A-3.—Replicatory WTO data daily;
- A-4.—Replicatory NATO data daily;
- A-5.—Tanter's WTO data three-day;
- A-6.—Tanter's NATO data three-day;
- A-7.—Replicatory WTO data three-day;
- A-8.—Replicatory NATO data three-day.

All abbreviations used in the tables are noted after the table in which they appear.

TABLE A-1.—Tanter's WTO data daily.

TIME	TWFD	TWFDT	TWSD	TWSDT
1	2	---	23.500	---
2	0	1	0	11.750
3	0	.500	0	5.875
4	0	.250	0	2.937
5	0	.125	0	1.468
6	0	.062	0	.734
7	1	.031	20	.367
8	0	.515	0	10.183
9	0	.257	0	5.091
10	0	.128	0	2.542
11	0	.064	0	1.271
12	2	.032	18	.635
13	0	1.016	0	9.318
14	0	.508	0	4.658
15	0	.253	0	2.327
16	0	.126	0	1.160
17	0	.063	0	.580
18	0	.031	0	.290
19	0	.015	0	.145
20	0	.007	0	.072
21	0	.003	0	.032
22	0	.001	0	.016
23	0	0	0	.008
24	0	0	0	.004
25	2	0	35	.001
26	0	1	0	17.500
27	0	.500	0	8.750
28	0	.250	0	4.375
29	0	.125	0	2.187
30	0	.062	0	1.093
31	0	.031	0	.547
32	0	.015	0	.273
33	0	.007	0	.136
34	0	.003	0	.063
35	2	.001	27.500	.031
36	0	1	0	13.765
37	0	.500	0	6.765
38	0	.250	0	3.441
39	0	.125	0	1.718
40	1	.062	4	.859
41	0	.531	0	2.429
42	0	.265	0	1.214
43	0	.132	0	.607
44	0	.066	0	.299
45	0	.033	0	.149

TABLE A-1.—continued

TIME	TWFD	TWFDT	TWSD	TWSDT
46	0	.016	0	.074
47	0	.008	0	.037
48	0	.004	0	.018
49	0	.001	0	.007
50	0	0	0	.003
51	0	0	0	.001
52	0	0	0	0
53	0	0	0	0
54	1	0	10	0
55	0	.500	0	5
56	0	.250	0	2.500
57	0	.125	0	1.250
58	1	.062	22	.625
59	0	.531	0	11.312
60	1	.265	7	5.656
61	1	.632	67	6.328
62	0	.816	0	36.663
63	1	.408	18	18.330
64	0	.704	0	18.165
65	2	.352	25	9.082
66	0	1.176	0	17.041
67	1	.587	92	8.517
68	1	.793	23	50.257
69	0	.896	0	36.627
70	0	.448	0	18.303
71	2	.224	130	9.151
72	1	1.111	30	69.571
73	0	1.055	0	49.786
74	1	.527	140	24.888
75	1	.763	22	82.440
76	1	.881	105	52.205
77	1	.940	140	78.599
78	0	.970	0	109.299
79	0	.485	0	54.647
80	1	.242	10	27.304
81	1	.620	35	18.641
82	0	.810	0	26.819
83	0	.405	0	13.389
84	0	.202	0	6.689
85	0	.100	0	3.321
86	0	.050	0	1.638
87	0	.025	0	.818
88	1	.012	20	.401
89	0	.506	0	10.198
90	0	.252	0	5.087
91	1	.126	12	2.535
92	0	.563	0	7.268
93	0	.281	0	3.634
94	1	.140	2	1.816

TABLE A-1.—continued

TIME	TWFD	TWFDT	TWSD	TWSDT
95	0	.570	0	1.906
96	0	.285	0	.953
97	1	.142	5	.473
98	1	.571	20	2.736
99	1	.785	60	11.368
100	1	.892	25	35.682
101	0	.946	0	30.341
102	0	.473	0	15.169
103	0	.236	0	7.584
104	0	.118	0	3.792
105	0	.059	0	1.895
106	2	.029	33	.946
107	0	1.014	0	16.969
108	1	.507	4	8.475
109	1	.753	20	6.234
110	1	.876	5	13.117
111	0	.938	0	9.058
112	0	.469	0	4.528
113	0	.234	0	2.260
114	0	.117	0	1.128
115	1	.058	4	.559
116	1	.529	5	2.279
117	0	.764	0	3.639
118	1	.382	6	1.816
119	0	.690	0	3.907
120	1	.345	8	1.951
121	1	.672	10	4.976
122	0	.836	0	7.487
123	1	.418	35	3.742
124	1	.708	30	19.370
125	1	.854	18	24.684
126	0	.927	0	21.342
127	1	.463	2	10.670
128	1	.731	50	6.335
129	0	.865	0	28.166
130	0	.432	0	14.081
131	0	.216	0	7.040
132	1	.107	75	3.514
133	0	.553	0	39.252
134	1	.276	5	19.623
135	0	.638	0	12.311
136	1	.318	4	6.154

Note: TWFD-Tanter's WTO frequency daily; TWFDT-Tanter's WTO frequency daily transformed; TSWD-Tanter's WTO score daily; TWSDT-Tanter's WTO score daily transformed.

TABLE A-2.—Tanter's NATO data daily

TIME	TNFD	TNFDT	TNSD	TNSDT
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	1	0	7	0
7	0	.500	0	3.500
8	1	.250	22	1.750
9	0	.625	0	11.875
10	0	.312	0	5.937
11	0	.156	0	2.968
12	0	.078	0	1.484
13	0	.039	0	.742
14	0	.019	0	.371
15	0	.009	0	.184
16	0	.004	0	.092
17	0	.002	0	.042
18	0	.001	0	.021
19	0	0	0	.010
20	0	0	0	.005
21	0	0	0	.002
22	0	0	0	0
23	0	0	0	0
24	1	0	12	0
25	1	.500	26	6
26	1	.750	18	16
27	0	.875	0	17
28	0	.437	0	8.500
29	0	.218	0	4.250
30	0	.109	0	2.125
31	1	.054	5	1.062
32	0	.527	0	3.031
33	0	.263	0	1.513
34	0	.131	0	.752
35	1	.065	9	.373
36	0	.532	0	4.686
37	1	.266	25	2.343
38	1	.633	30	13.671
39	1	.816	14	21.833
40	1	.908	6	17.915
41	0	.954	0	11.957
42	0	.477	0	5.978
43	1	.238	9	2.989
44	1	.619	23	5.993
45	1	.809	18	14.496

TABLE A-2.—continued

TIME	TNFD	TNFDI	TNSD	TNSDI
46	1	.904	15	16.244
47	0	.952	0	15.617
48	1	.475	16	7.806
49	0	.737	0	11.902
50	1	.368	15	5.951
51	0	.684	0	10.474
52	2	.341	66	5.233
53	1	1.170	25	35.612
54	0	1.085	0	30.303
55	1	.542	12	15.149
56	0	.771	0	13.574
57	0	.385	0	6.784
58	2	.192	16.500	3.391
59	1	1.096	55	9.942
60	1	1.047	8	32.469
61	0	1.023	0	20.224
62	2	.511	16.500	10.107
63	1	1.255	36	13.303
64	1	1.127	14	24.649
65	0	1.063	0	19.324
66	1	.531	20	9.658
67	0	.765	0	14.825
68	0	.382	0	7.404
69	0	.191	0	3.700
70	0	.095	0	1.850
71	1	.047	130	.922
72	2	.523	22.500	65.454
73	1	1.261	130	43.972
74	2	1.130	40	86.985
75	2	1.565	164	63.489
76	1	1.782	120	113.743
77	1	1.391	30	116.870
78	1	1.195	28	73.434
79	1	1.097	50	50.716
80	0	1.048	0	50.337
81	1	.524	120	25.165
82	1	.761	80	72.562
83	1	.880	65	76.274
84	2	.940	107.500	70.611
85	0	1.469	0	89.031
86	0	.734	0	44.508
87	0	.367	0	22.244
88	1	.183	18	11.111
89	0	.591	0	14.545
90	1	.295	10	7.248
91	0	.647	0	8.609
92	0	.323	0	4.292
93	0	.161	0	2.128
94	2	.080	16	1.063

TABLE A-2.—continued

TIME	TNFD	TNFDT	TNSD	TNSDT
95	0	1.040	0	8.525
96	1	.520	18	4.258
97	1	.759	65	11.122
98	1	.879	65	38.055
99	0	.939	0	51.526
100	0	.469	0	25.763
101	1	.234	120	12.881
102	0	.617	0	66.439
103	1	.308	30	33.217
104	1	.654	10	31.608
105	1	.826	4	20.800
106	0	.913	0	12.390
107	0	.456	0	6.185
108	0	.228	0	3.091
109	0	.114	0	1.545
110	1	.056	15	.753
111	0	.528	0	7.872
112	0	.263	0	3.928
113	0	.131	0	1.963
114	0	.065	0	.980
115	0	.032	0	.483
116	1	.016	4	.241
117	0	.508	0	2.119
118	0	.254	0	1.058
119	0	.126	0	.527
120	0	.063	0	.263
121	0	.031	0	.131
122	0	.015	0	.065
123	1	.007	35	.032
124	0	.503	0	17.515
125	0	.251	0	8.757
126	0	.125	0	4.378
127	0	.062	0	2.189
128	0	.031	0	1.094
129	1	.015	22	.547
130	0	.507	0	11.273
131	2	.253	26.500	5.636
132	0	1.126	0	16.062
133	0	.563	0	8.031
134	0	.281	0	4.015
135	1	.140	3	2.008
136	0	.570	0	2.503

Note: TNFD-Tanter's NATO frequency daily; TNFDT-Tanter's NATO frequency daily transformed; TNSD-Tanter's NATO score daily; TNSDT-Tanter's NATO score daily transformed.

TABLE A-3.—Replicatory WTO data daily

TIME	RWFD	RWFDT	RWSD	RMSDT
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	0	0	0	0
8	0	0	0	0
9	0	0	0	0
10	2	0	86.999	0
11	0	1	0	43.499
12	1	.500	86.999	21.749
13	1	.750	148.998	54.374
14	0	.875	0	101.686
15	0	.437	0	50.843
16	0	.218	0	25.422
17	0	.109	0	12.710
18	0	.054	0	6.355
19	1	.027	148.998	3.164
20	0	.513	0	76.080
21	0	.256	0	38.026
22	0	.128	0	18.991
23	0	.064	0	9.495
24	0	.031	0	4.743
25	0	.015	0	2.372
26	3	.007	107.666	1.180
27	0	1.503	0	54.414
28	0	.751	0	27.184
29	0	.375	0	13.592
30	0	.187	0	6.796
31	0	.093	0	3.398
32	0	.047	0	1.699
33	0	.023	0	.840
34	0	.011	0	.419
35	0	.005	0	.193
36	0	.002	0	.096
37	0	.001	0	.048
38	0	0	0	.024
39	0	0	0	.011
40	0	0	0	0
41	0	0	0	0
42	0	0	0	0
43	0	0	0	0
44	0	0	0	0
45	0	0	0	0

TABLE A-3.—continued

TIME	RWFD	RWFDT	RWSD	RWSDT
46	0	0	0	0
47	0	0	0	0
48	0	0	0	0
49	0	0	0	0
50	0	0	0	0
51	0	0	0	0
52	0	0	0	0
53	0	0	0	0
54	1	0	86.999	0
55	0	.500	0	43.499
56	0	.250	0	21.749
57	0	.125	0	10.874
58	0	.062	0	5.437
59	0	.031	0	2.718
60	0	.015	0	1.359
61	0	.007	0	.679
62	0	.003	0	.339
63	1	.001	148.998	.156
64	0	.500	0	74.577
65	0	.250	0	37.288
66	1	.125	148.998	18.644
67	0	.562	0	83.821
68	0	.281	0	41.905
69	0	.140	0	20.953
70	1	.070	86.999	10.476
71	2	.535	248.998	48.736
72	0	1.267	0	148.846
73	1	.633	248.997	74.422
74	0	.816	0	161.709
75	3	.408	215.665	80.832
76	0	1.704	0	148.249
77	0	.852	0	74.117
78	2	.426	248.998	37.057
79	1	1.212	248.997	143.014
80	0	1.106	0	195.959
81	0	.553	0	97.979
82	1	.276	248.997	48.951
83	1	.638	248.997	148.974
84	2	.818	248.998	198.948
85	2	1.409	248.998	223.961
86	1	1.704	248.997	236.476
87	1	1.351	248.997	242.685
88	1	1.175	248.997	245.804
89	1	1.087	248.997	247.389
90	1	1.043	248.997	248.193
91	1	1.021	248.997	248.556
92	1	1.010	248.997	248.723
93	2	1.004	198.999	248.808
94	3	1.502	140.999	223.866

TABLE A-3.—continued

TIME	RWFD	RWFDT	RWSD	RWSDT
95	1	2.250	248.997	182.394
96	1	1.625	248.997	215.643
97	1	1.312	248.997	232.268
98	1	1.155	248.997	240.580
99	4	1.077	127.499	244.736
100	1	2.538	248.997	186.065
101	1	1.769	248.997	217.480
102	2	1.384	248.998	233.194
103	1	1.691	248.997	241.060
104	1	1.345	248.997	244.976
105	2	1.172	248.998	246.934
106	3	1.586	248.999	247.914
107	2	2.292	198.999	248.408
108	2	2.145	248.998	223.675
109	1	2.072	248.997	236.284
110	1	1.536	248.997	242.588
111	4	1.2677	248.999	245.740
112	2	2.633	167.999	247.318
113	2	2.316	167.999	207.612
114	1	2.157	248.997	187.753
115	1	1.578	248.997	218.323
116	1	1.288	248.997	233.615
117	1	1.144	248.997	241.254
118	1	1.071	248.997	245.073
119	1	1.035	248.997	246.982
120	1	1.017	248.997	247.939
121	1	1.008	248.997	248.431
122	1	1.003	248.997	248.674
123	2	1.001	248.998	248.783
124	0	1.500	0	248.838
125	1	.749	248.997	124.366
126	1	.874	248.997	186.634
127	2	.937	248.998	217.768
128	0	1.468	0	233.331
129	2	.733	248.998	116.613
130	1	1.366	248.997	182.752
131	1	1.183	248.997	215.824
132	2	1.091	167.999	232.359
133	0	1.345	0	200.164
134	2	.772	248.998	100.028
135	0	1.386	0	174.462
136	0	.692	0	87.179

Notes: RWFD=replicatory WTO frequency daily; RWFDT=replicatory WTO frequency daily transformed; RWSD=replicatory score daily; RWSDT=replicatory score daily transformed.

TABLE A-4.—Replicatory NATO data daily

TIME	RNFD	RNFDT	RNSD	RNSDT
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	2	0	86.999	0
8	0	1	0	43.499
9	0	.500	0	21.749
10	0	.250	0	10.874
11	0	.125	0	5.437
12	0	.062	0	2.718
13	0	.031	0	1.359
14	0	.015	0	.679
15	2	.007	86.999	.339
16	0	1.003	0	43.656
17	0	.501	0	21.828
18	0	.250	0	10.914
19	0	.125	0	5.457
20	3	.062	194.999	2.728
21	0	1.531	0	98.859
22	0	.765	0	49.429
23	2	.382	167.999	24.714
24	0	1.191	0	96.343
25	0	.595	0	48.171
26	2	.297	86.999	24.086
27	0	1.148	0	55.542
28	0	.574	0	27.769
29	0	.286	0	13.851
30	0	.143	0	6.925
31	0	.071	0	3.461
32	0	.035	0	1.707
33	0	.017	0	.852
34	0	.008	0	.414
35	0	.004	0	.195
36	0	.002	0	.096
37	0	0	0	.039
38	0	0	0	.020
39	0	0	0	.009
40	0	0	0	0
41	0	0	0	0
42	0	0	0	0
43	2	0	86.999	0
44	0	1	0	43.499
45	0	.500	0	21.749

TABLE A-4.—continued

TIME	RNFD	RNFDT	RNSD	RNSDT
46	3	.250	107.666	10.874
47	0	1.625	0	59.270
48	0	.812	0	29.635
49	0	.406	0	14.818
50	0	.203	0	7.408
51	1	.101	86.999	3.703
52	0	.550	0	45.338
53	0	.275	0	22.668
54	0	.137	0	11.333
55	0	.068	0	5.651
56	1	.034	86.999	2.825
57	0	.516	0	44.907
58	0	.258	0	22.453
59	0	.129	0	11.226
60	0	.064	0	5.594
61	0	.032	0	2.797
62	0	.016	0	1.398
63	0	.008	0	.699
64	0	.004	0	.348
65	0	.001	0	.156
66	0	0	0	.078
67	0	0	0	.039
68	0	0	0	.020
69	2	0	86.999	.009
70	0	1	0	43.499
71	2	.500	248.998	21.749
72	0	1.250	0	135.374
73	1	.625	248.997	67.687
74	1	.812	248.997	158.342
75	2	.906	248.998	203.670
76	1	1.453	86.999	226.334
77	0	1.226	0	156.667
78	2	.612	248.998	78.319
79	1	1.306	248.997	163.659
80	0	1.152	0	206.290
81	0	.576	0	103.144
82	2	.288	167.999	51.532
83	2	1.143	198.999	109.723
84	3	1.571	194.999	154.323
85	4	2.285	167.999	174.635
86	1	3.142	248.997	171.315
87	1	2.071	248.997	210.104
88	1	1.535	248.997	229.499
89	1	1.267	248.997	239.234
90	1	1.133	248.997	244.111
91	2	1.066	248.998	246.508
92	1	1.592	248.997	247.718
93	1	1.265	248.997	248.314
94	1	1.132	248.997	248.630

TABLE A-4.—continued

TIME	RNFD	RNFDT	RNSD	RNSDT
95	1	1.066	248.997	248.775
96	1	1.032	248.997	248.838
97	1	1.016	248.997	248.868
98	3	1.007	140.999	248.884
99	1	2.003	248.997	194.893
100	1	1.501	248.997	221.893
101	1	1.250	248.997	235.393
102	2	1.125	248.998	242.143
103	1	1.562	248.997	245.518
104	1	1.280	248.997	247.205
105	3	1.14	194.999	248.049
106	3	2.069	248.999	221.472
107	1	2.534	248.997	235.199
108	2	1.767	248.998	242.046
109	1	1.883	248.997	245.470
110	1	1.441	248.997	247.181
111	1	1.220	248.997	248.037
112	1	1.109	248.997	248.471
113	1	1.054	248.997	248.682
114	1	1.026	248.997	248.795
115	2	1.012	167.999	248.844
116	1	1.506	248.997	208.369
117	1	1.252	248.997	228.630
118	1	1.126	248.997	238.762
119	1	1.062	248.997	243.830
120	1	1.031	248.997	246.361
121	1	1.015	248.997	247.626
122	1	1.007	248.997	248.260
123	2	1.003	248.998	248.577
124	0	1.501	0	248.747
125	1	.750	248.997	124.321
126	2	.875	167.999	186.607
127	2	1.437	248.998	177.250
128	0	1.718	0	213.072
129	2	.858	248.998	106.488
130	1	1.429	248.997	177.690
131	1	1.214	248.997	213.292
132	1	1.106	248.997	231.093
133	0	1.053	0	240.031
134	2	.526	248.998	119.962
135	0	1.262	0	184.441
136	0	.631	0	92.168

Note: RNFD=replicatory NATO frequency daily; RNFDT=replicatory NATO frequency daily transformed; RNSD=replicatory NATO score daily; RNSDT=replicatory NATO score daily transformed.

TABLE A-5.—Tanter's WTO data three day.

TIME	TWFT	TWFTT	TWST	TWSTT
1	0	1.749	0	20.562
2	1	.219	20	2.567
3	0	.903	0	17.817
4	2	1.111	18	11.224
5	0	.889	0	8.147
6	0	.111	0	1.015
7	0	.014	0	.120
8	2	.001	35	.013
9	0	1.749	0	30.624
10	0	.219	0	3.829
11	0	.028	0	.473
12	2	1.502	27.500	20.680
13	1	.438	4	6.018
14	0	.930	0	4.251
15	0	.116	0	.525
16	0	.014	0	.064
17	0	.001	0	.006
18	1	.500	10	5
19	1	.438	22	4.374
20	2	1.429	74	23.296
21	1	1.927	18	73.134
22	3	2.115	117	34.641
23	1	2.138	23	105.189
24	3	2.391	160	128.510
25	3	2.173	267	159.533
26	1	2.395	140	242.545
27	2	1.672	45	72.765
28	0	.708	0	23.399
29	1	.088	20	2.850
30	1	.885	12	17.820
31	1	.986	2	12.717
32	1	.997	5	3.332
33	3	2.249	105	49.706
34	0	1.654	0	53.094
35	2	.206	33	6.634
36	2	2.274	24	31.679
37	1	2.283	5	26.702
38	1	.409	4	3.947
39	2	1.674	11	7.737
40	2	1.709	18	10.824
41	2	1.962	65	30.600
42	2	2.243	20	56.695

TABLE A-5.—continued

TIME	TWFT	TWFTT	TWST	TWSTT
43	1	2.030	50	48.581
44	1	.878	75	49.806
45	2	1.232	9	38.098

Note: TWFT=Tanter's WTO frequency three-day; TWFTT=Tanter's WTO frequency three-day transformed; TWST=Tanter's WTO score three-day; TWSTT=Tanter's WTO score three-day transformed.

TABLE A-6.—Tanter's NATO data three day.

TIME	TNFT	TNFTT	TNST	TNSTT
1	0	0	0	0
2	1	.500	7	3.500
3	1	1.187	22	19.562
4	0	.273	0	5.194
5	0	.033	0	.647
6	0	.004	0	.075
7	0	0	0	.007
8	2	.500	28	6
9	1	2.062	18	41.500
10	1	.383	5	7.437
11	0	.923	0	5.297
12	2	.865	34	7.404
13	3	2.356	50	53.419
14	1	1.669	9	20.925
15	3	2.332	56	36.732
16	1	2.165	16	35.326
17	3	1.394	81	21.658
18	2	2.797	37	81.063
19	2	1.348	16.500	23.749
20	2	3.167	63	62.636
21	4	2.895	66.500	48.061
22	1	2.360	20	43.807
23	0	.670	0	12.953
24	4	1.832	282.500	110.349
25	5	4.477	324	264.219
26	3	3.684	108	241.021
27	2	2.334	200	148.064
28	3	3.290	172.500	235.916
29	1	1.283	18	77.864
30	1	1.562	10	30.403
31	2	.566	16	7.484
32	2	2.319	83	23.904
33	1	2.289	65	115.343
34	2	1.159	150	112.538
35	2	2.393	14	64.798
36	0	.838	0	10.822
37	1	.849	15	12.554
38	0	.230	0	3.426
39	1	.778	4	3.419
40	0	.222	0	.923
41	1	.528	35	17.614
42	0	.441	0	15.324

TABLE A-6.—continued

TIME	TNFT	TNFTT	TNST	TNSTT
43	1	.550	22	12.914
44	2	1.943	26.500	29.730
45	1	.933	3	8.527

Note: TNFT=Tanter's NATO frequency three-day; TNFTT=Tanter's NATO frequency three-day transformed; TNST=Tanter's NATO score three-day; TNSTT=Tanter's NATO score three-day transformed.

TABLE A-7.—Replicatory WTO data three day.

TIME	RWFT	RWFTT	RWST	RWSTT
1	0	0	0	0
2	0	0	0	0
3	2	0	87	0
4	2	2.249	236	119.624
5	0	1.532	0	177.951
6	1	.191	149	22.230
7	0	.908	0	133.097
8	0	.112	0	16.610
9	3	2.262	107.665	82.778
10	0	.660	0	23.786
11	0	.081	0	2.959
12	0	.009	0	.339
13	0	.001	0	.037
14	0	0	0	0
15	0	0	0	0
16	0	0	0	0
17	0	0	0	0
18	1	.500	87	43.500
19	0	.438	0	38.056
20	0	.055	0	4.757
21	1	.507	149	75.073
22	1	.938	149	139.874
23	0	.492	87	73.335
24	3	2.435	498	272.004
25	3	2.928	215.665	390.790
26	3	2.490	498	254.188
27	1	1.934	249	342.889
28	5	2.865	747	571.883
29	3	4.232	747	724.965
30	3	3.153	747	744.138
31	6	3.517	589	721.937
32	3	5.187	747	630.305
33	6	4.772	625.500	671.383
34	4	4.844	747	691.735
35	6	4.104	747	739.824
36	5	6.511	697	708.457
37	7	5.437	666	735.646
38	4	6.051	666	613.690
39	3	3.504	747	720.842
40	3	3.060	747	743.352
41	3	3.515	498	745.296
42	4	2.561	747	528.768

TABLE A-7.—continued

TIME	RWFT	RWFTT	RWST	RWSTT
43	3	3.568	498	532.696
44	3	3.820	417	648.348
45	2	2.851	249	361.670

Note: RWFT=replicatory WTO frequency three-day;
RWFTT=replicatory WTO frequency three-day transformed;
RWST=replicatory WTO score three-day; RWSTT=replicatory WTO
score three-day transformed.

TABLE A-8.—Replicatory NATO data three day.

TIME	RNFT	RNFTT	RNST	RNSTT
1	0	0	0	0
2	2	0	87	0
3	0	1.749	0	76.124
4	0	.219	0	9.515
5	2	1.027	87	44.673
6	0	.878	0	38.198
7	3	2.359	195	151.016
8	2	2.162	168	169.229
9	2	2.020	87	107.399
10	0	.502	0	24.238
11	0	.063	0	2.973
12	0	.007	0	.331
13	0	0	0	.030
14	2	0	87	0
15	3	1.749	107.665	76.124
16	0	2.844	0	103.723
17	1	.856	87	56.451
18	0	.481	0	39.653
19	1	.810	87	70.186
20	0	.225	0	19.616
21	0	.028	0	2.447
22	0	.002	0	.274
23	2	1	87	43.519
24	3	2.374	498	224.880
25	4	3.171	585	588.345
26	3	3.145	498	398.644
27	3	2.016	168	360.967
28	9	5.002	561	438.690
29	3	6.749	747	610.918
30	4	3.468	747	729.874
31	3	3.931	747	744.661
32	3	3.114	747	746.481
33	5	4.511	639	665.670
34	4	3.972	747	723.054
35	7	4.490	693	716.727
36	4	6.183	747	722.715
37	3	3.770	747	743.688
38	4	3.094	666	746.321
39	3	3.884	747	675.761
40	5	3.108	666	737.819
41	3	3.510	498	745.584
42	5	3.061	666	488.179

TABLE A-8.—continued

TIME	RNFT	RNFTT	RNST	RNSTT
43	3	4.006	498	497.250
44	2	3.373	498	684.417
45	2	2.418	249	396.572

Note: RNFT=replicatory NATO frequency three-day;
RNFTT=replicatory NATO frequency three-day transformed;
RNST=replicatory NATO score three-day; RNSTT=replicatory
NATO score three-day transformed.

BIBLIOGRAPHY

ALLISON, G. T. (1969) "Conceptual models and the Cuban missile crisis." *American Political Science Review* 63 (September): 689-718.

——— (1971) *The Essence of Decision*. Boston: Little, Brown.

ASHLEY, R. K. (1974) "Noticing pre-paradigmatic progress." Presented at the Conference on Cumulation in Scientific International Relations Research, Santa Monica, Calif., March 27-29.

AZAR, E. E., J. P. BENNETT, and T. J. SLOAN (1974) "Steps toward forecasting in international interactions." *Peace Science Society (International) Papers* 23: 27-67.

BLALOCK, H. M., Jr. (1968) "The measurement problem; a gap between the languages of theory and research, pp. 5-27 in H. M. Blalock, Jr. and A. B. Blalock (eds.) *Methodology in Social Research*. Chicago: Aldine.

——— (1969) *Theory Construction*. Englewood Cliffs, N.J.: Prentice-Hall.

BOEROW, D. B. (1972) "The relevance potential of different products," pp. 204-228 in R. Tanter and R. H. Ullman (eds.) *Theory and Policy in International Relations*. Princeton, N.J.: Princeton University Press.

BOYNTON, G. R. (1974) "Cumulateness in political science." Presented at the Conference on Cumulation in Scientific International Relations, Santa Monica, Calif., March 27-29.

BREMER, S. A. (1974) "Obstacles to the accumulation of knowledge in international relations research." Presented at the Conference on Cumulation in Scientific International Relations Research, Santa Monica, Calif., March 27-29.

BULL, H. (1969) "International theory: the case for a classical approach," pp. 20-38 in Klaus Knorr and James N. Rosenau (eds.) *contending approaches to international politics*. Princeton, N.J., Princeton University Press.

BURGESS, P. G. and R. W. LAWTON (1972) Indicators of International Behavior. Beverly Hills, Calif.: Sage.

BURROWES (1974) "Mirror, mirror, on the wall....: a comparison of event data sources," pp. 383-406 in James N. Rosenau (ed.) Comparing Foreign Policies. Beverly Hills, Calif.: Sage.

CARLETON UNIVERSITY COMPUTER SERVICES (1976) TSP (mimeo).

CHOUCRI, N. (1974) "Applications of economic analysis to forecasting in international relations." Peace Science Society (International) Papers 21: 15-39.

_____ and R. C. NORTH (1972) "Dynamics of international conflict: some policy implications of population, resources, and technology," pp. 80-122 in R. Tanter and R. H. Ullman (eds.) Theory and Policy in International Relations. Princeton, N.J.: Princeton University Press.

_____ and _____ (1974) Nations in Conflict. San Francisco: W. H. Freeman.

CORSON, W. H. (1974) "Bases of United States-Soviet conflict and cooperation, 1945-1965." Presented at the annual meeting of the International Studies Association, St. Louis, March 20-23.

CYERT, R. M. and J. G. MARCH (1963) A Behavioral Theory of the Firm, Englewood Cliffs, N.J.: Prentice-Hall.

DAVIS, O. A., M. A. H. DEMPSTER, and A. WILDAVSKY (1969) "A theory of the budgetary process," pp. 462-495 in Charles F. Cnudde and Deane E. Neubauer (eds.) Empirical Democratic Theory. Chicago: Markham.

DEBER, R. B. (1975) "Missing data in events analysis: the problem and suggested solution," pp. 63-79 in Edward E. Azar and J. Ben-Dak (eds.) Theory and Practice of Events Research. New York: Gordon and Breach.

DEUTSCH, K. W. (1968) "The coming crisis of cross-national and international research in the United States." American Council of Learned Societies Newsletter 19 (April): 1-7.

DURBIN, J. (1970) "Testing for several correlation in least-squares regression when some of the regressors are lagged dependent variables." Econometrica 38 (May): 410-421.

FIELD, J. O., N. CHOUCRI, and W. C. MITCHELL (1972) "The Sino-Indian border conflict: an exploratory analysis of action and perception," pp. 31-59 in John H. Sigler, John O. Field, and Murray L. Adelman (comps.) Applications of Event Data Analysis. Beverly Hills, Calif.: Sage.

- FLETCHER, C. (1970) "On replication: notes on the notions of replicability quotient and a generalizability quotient." *Sociology* 4 (January): 51-70.
- GALTUNG, J. (1967) *Theory and Methods of Social Research*. New York: Columbia University Press.
- HAVENER, T. and A. PETERSON (1975) "Measuring conflict cooperation in international relations: a methodological inquiry," pp. 57-61 in Edward E. Azar and Joseph Ben-Dak (eds.) *Theory and Practice of Events Research*. New York: Gordon and Breach.
- HANDELMAN, J. R. (1975) "Projecting violence in the middle east," pp. 171-201 in Michael K. O'Leary and William D. Coplin (eds.) *Quantitative Techniques on Foreign Policy Analyses and Forecasting*. New York: Praeger.
- HERMANN, C. F. (1972) "Some issues in the study of international crises," pp. 3-17 in Charles F. Hermann (ed.) *International Crises*. New York: Free Press.
- HIBBS, D. A., Jr. (1974) "Problems of statistical estimation and causal inference in time-series regression models." *Sociological Methodology*, 1973-1974: 252-308.
- HILTON, G. (1976) *Intermediate Politometrics*. New York: Columbia University Press.
- HOGGARD, G. D. (1974) "Differential source coverage in foreign policy analysis," pp. 353-381 in James N. Rosenau (ed.) *Comparing Foreign Policies*. Beverly Hills, Calif.: Sage.
- HOPMANN, P. T. (1974) "Cumulation in scientific international relations research." Presented at the Conference on Cumulation in Scientific International Relations Research. Santa Monica, Calif., March 27-29.
- JOHNSTON, J. (1972) *Econometric Methods*, 2nd ed. New York: McGraw-Hill.
- JONES, S. D., and J. D. SINGER (1972) *Beyond Conjecture in International Politics*. Itasca, Ill.: Peacock.
- KAPLAN, A. (1964) *The Conduct of Inquiry*. San Francisco: Chandler.
- KAPLAN, M. A. (1969) "The new great debate: traditionalism vs. science in international relations," pp. 39-61 in Klaus Knorr and James N. Rosenau (eds.) *Contending approaches to international politics*. Princeton N.J.: Princeton University Press.

KEGLEY, C. W., Jr. (1975) "Review of Modelling and Managing International Conflicts." *Journal of Politics* 37 (August): 824-826.

_____, S. A. SALMORE, and D. J. ROSEN (1974) "Convergences in the measurement of state behavior." *Sage International Yearbook of Foreign Policy Studies* 2: 309-339.

KMENTA, J. (1971) *Elements of Econometrics*. New York: Macmillan.

LEGAULT, A., J. STEIN, J. SIGLER, and B. STEINBERG (1973) "L'analyse comparative des conflits interetatiques dyadiques (CADIC)." *Etudes Internationales* 4 (December): 480-501.

_____, _____, _____, and _____ (1974) "CADIC—comparative analyses of dyadic interstate conflict—coding actor behavior." *Universite Laval* (mimeo).

LENG, R. J. (1974) "Scientific accumulation in international relations research." Presented at the Conference on Cumulation in Scientific International Relations, Santa Monica, Calif., March 27-29.

_____, (1975) "'The future of events data marriages: a question of compatability'." *International Interactions* 2 (July): 45-62.

MCCLELLAND, C. A. (1972) "On the fourth wave: past and future in the study of international systems," pp. 15-40 in James N. Rosenau, Vincent Davis, and Maurice A. East (eds.) *The Analysis of International Politics*. New York: Free Press.

MCCORMICK, J. M. (1975) "Evaluating models of crises behavior: some evidence from the Middle East." *International Studies Quarterly* 19 (March): 17-66.

MCGOWAN, P. J. (1974) "Meaningful comparisons in the study of foreign policy: a methodological discussion of objectives, techniques, and research design," pp. 52-87 in Charles W. Kegley, Jr., Gregory A. Raymond, Robert M. Rood and Richard A. Skinner (eds.) *International Events and the Comparative Analysis of Foreign Policy*. Columbia, S.C.: University of South Carolina Press.

_____, and H. B. SHAPIRO (1973) *The Comparative Study of Foreign Policy*. Beverly Hills, Calif.: Sage.

MARCH, J. G. and H. A. SIMON (1958) *Organizations*. New York: John Wiley.

MARTIN, R. W. (1974) "Conditions and practices which inhibit cumulation in international relations research." Presented at the Conference on Cumulation in Scientific International Relations Research, Santa Monica, Calif., March 27-29.

NERLOVE, M. and K. F. WALLIS (1966) "Use of the Durbin-Watson statistic in inappropriate situations." *Econometrica* 34 (January): 235-238.

PAULIN, G. (1975) "Review of Modelling and Managing International Conflicts." *Etudes Internationales* 5 (September): 417-419.

PETERSON, S. (1975) "Research on research: events data studies, 1961-1972." *Sage International Yearbook of Foreign Policy Studies* 3: 263-309.

PHILLIPS, W. R. (1975) "Two view of foreign policy interaction," pp. 245-260 in Edward E. Azar and Joseph Ben-Dak (eds.) *Theory and Practice of Events Research*, New York: Gordon and Breach.

——— (1974) "Where have all the theories gone?" *World Politics* 26 (January): 155-188.

POOL, I. de SOLA (1959) "Trends in content analysis," pp. 189-233 in Ithiel de Sola Pool (ed.) *Trends in Content Analysis*. Urbana: University of Illinois Press.

RICHMAN, A. (1975) "Issues in the conceptualization and measurement of events data." *International Interactions* 2 (July): 19-32.

ROSENAU, J. N. (1973) "Success and failure in scientific international relations research: a report on a workshop." University of Southern California (mimeo).

——— (1973) *International Studies and the Social Sciences*. Beverly Hills, Calif.: Sage.

——— and G. H. RAMSEY (1975) "External and internal typologies of foreign policy behavior: testing the stability of an intriguing net of findings." *Sage International Yearbook of Foreign Policy Studies* 3: 245-262

SIGLER, J. H. (1972) "Reliability problems in the measurement of international events in the elite press," pp. 9-29 in John H. Sigler, John O. Field, and Murray L. Adelman (comps.) *Applications of Events Data Analysis*. Beverly Hills, Calif.: Sage.

____ (1974) "Problems in the development of cumulative knowledge in scientific international relations research." Prepared for the Conference on Cumulation in Scientific International Relations Research. Santa Monica, Calif., March 27-29.

SIMON, H. A. (1957) Models of Man. New York: John Wiley.

SIVERSON, R. M. (1974) "Some comments on cumulation in scientific international relations research." Prepared for the Conference on Cumulation in Scientific International Relations Research. Santa Monica, Calif., March 27-29.

SMOKER, P. (1969) "A time-series analysis of Sino-Indian relations." Journal of Conflict Resolution 13 (June): 172-191.

STARR, H. (1974) "The quantitative international relations scholar as surfer: riding the 'Fourth Wave'." Journal of Conflict Resolution 18 (June): 336-368.

TANTER, R. (1972) "International system and foreign policy approaches: implications for conflict modelling and management," pp. 7-39 in R. Tanter and R. H. Ullman (eds.) Theory and Policy in International Relations. Princeton, N. J.: Princeton University Press.

____ (1974) Modelling and Managing International Conflicts. Beverly Hills, Calif.: Sage.

____ (1966) "Dimensions of conflict behavior within and between nations, 1958-1960." Journal of Conflict Resolution 10 (March): 41-64.

____ and W. C. POTTER (1973) "Modelling alliance behavior: East-West conflict over Berlin." Peace Science Society (International) Papers 20: 25-41.

TUFTE, E. R. (1974) Data Analysis for Politics and Policy. Englewood Cliffs, N. J.: Prentice-Hall.

VINCENT, J. E. (1974) "Comments on scientific accumulation in international relations research." Prepared for the Conference on Cumulation in Scientific International Relations Research, Santa Monica, Calif., March 27-29.

WILDAVSKY, A. B. (1964) The Politics of the Budgetary Process. Boston: Little, Brown.

WILSON, F. D., F. L. SMOKE, and J. D. MARTIN (1973) "The replication problem in sociology: a report and a suggestion." Sociological Inquiry 43: 141-149.

ZINNES, D. A. (1974) "The problem of accumulation in the scientific study of international politics." Prepared for the Conference on Cumulation in Scientific International Relations Research, Santa Monica, Calif., March 27-29.

——— (1976) Contemporary Research in international relations, New York: Free Press.

END

5

08

2

FIN