RAND DOCUMENT

THE RELATIVE IMPACT OF U.S. MILITARY ACTIONS ON VIET CONG/NORTH VIETNAMESE ATTACKS IN SOUTH VIETNAM: PRELIMINARIES TO A COMPUTER SIMULATION

[FAR 11725]

Raymond Tanter

June 17, 1970

~~March 1970~~

~~D-20150-ARPA~~

D-20150-1-ARPA

CONTENTS

INTRODUCTION*

This is a preliminary report on the constuction of a Vietnam Computer Simulation (VCS). Purposes of such a simulation are (1) to specify causal models that characterize limited aspects of the war in South Vietnam; and, (2) to manipulate experimentally variables that assess probable consequences of alternative U.S. military actions. Verbal propositions from the literature concerning probable effects of alternative military actions in guerrilla and/or conventional warfare helped specify the mathematical models used; nineteen months of data--July 1966-January 1968-- in the I Corps Tactical Zone (CTZ) on military actions and outcomes form the initial data base. The problem of the relative importance of large and small unit military actions *vis à vis* enemy activity serves as the topic to illustrate the progress in the construction of a computer simulation. The tentative results suggest that increases in U.S. small unit actions are three to four times as effective as increases in U.S. battalion days of operation in decreasing Viet Cong/North Vietnamese attacks later. Here follows a summary of some prior research, quality control considerations, the design of the present inquiry, results as well as some caveats.

*Acknowledgements to Richard Cady, Daniel Ellsberg, Alvin Harman, Fred C. Ikle, Jeffrey Milstein, William Mitchell, Anthony Russo, Thomas Sanders and Paul Werbos for comments, and to Rita McDermott and Phyllis Glazer for administrative support and research assistance.

PRIOR RESEARCH

There are prior studies that try to develop quantitative predictors of Vietnam war outcomes both within and outside the defense analysis community. Within the community, analysts perform studies at such places as RAND, Research Analysis Corporation (RAC) and the Institute for Defense Analyses (IDA). Outside this community, much less work occurs; some of it takes place at Yale University, for example. Relevant to the present study are such works at RAND as: C.V. Sturdevant et al., An Examination of the Viet Cong Initiated Capability and Recent Performance, RM-4304, October 1964 and Additional Insight on the Military Situation in South Vietnam Relative to Prior Years, RM-4402, July 1965; R.D. Jones, Factor Analysis of Operational Data on Counterinsurgency, RM-4299, December 1964 and F.H. Denton, Trends in Viet Cong Attacks on Hamlets, RM-4604, December 1964; Edward Mitchell, Land Tenure and Rebellion: A Statistical Analysis of Factors Affecting Government Control in South Vietnam, RM-5181-ARPA, June 1967; Anthony Russo, Economic and Social Correlates of Government Control in South Vietnam, Draft RM-6008; Marvin B. Schaffer, Correlation of HES Scores with Independently Obtained Military Indicators, D-17655, October 18, 1968. Rather than summarizing these efforts, the following review briefly touches on aspects of some of these studies of relevance to the present inquiry.

Denton presents one of the first systematic analyses of Vietnam data. He tries to discover what factors occur in Viet Cong (VC) attacks on hamlets that are successful or not. Though pioneering, this study does not go far enough in assessing the more strategic implications of the trends in VC attacks. For example, consider the relation of U.S. military actions to enemy attacks, a dependent variable in the present study. Also, how do VC attacks relate to population control? Several RAND studies do address themselves to the population control question. For example, Edward Mitchell tries to explain population control--the percentage of hamlets under the Saigon government's (GVN) control, 1965. Controlling for some other possible explanations, he concludes that land inequality covaries directly with higher levels of GVN control of the population--the higher the land inequality, the greater the government control. Mitchell tries to account for this "surprising" finding by equating land inequality with lower peasant income. That is, he assumes that poorer peasants are more docile and thus are more subject to government control.

Russo, on the other hand, challenges Mitchell's assumption. Also using regression analysis across Vietnam provinces, Russo finds that land inequality covaries directly with higher peasant incomes; he concludes that the better-off peasant is more likely to be under government control. Both the Mitchell and Russo works regress variables across provinces and/or hamlets rather than over time as does the present inquiry. A striking feature of their spatial studies of interest here, however, is that both are able to explain over half the variation in 1965 population control without reference to military operations, a primary causal variable in this study. It may be that the military activity data are necessary to explain changes over time in population control whereas socio-economic variables do an adequate job spatially.

Evidence in favor of the relevance of military variables comes from a regression analysis by Schaffer. He tries to explain a somewhat different type of population control--changes over time in the number of people in categories of secure or insecure hamlets from the U.S. Hamlet Evaluation System (HES). Moreover, Schaffer's study includes military actions and outcomes but excludes environmental attributes in contrast to the Mitchell and Russo efforts. Schaffer aggregates the hamlet data to each of the four Corps Tactical Zones (CTZs) and uses regression analysis across 14 months--January, 1967-March, 1968. A principal finding of relevance here is that in the II, III and IV CTZs the number of small-unit actions is generally the most important predictor of population security. Appearing in several forms (small-unit actions without contact in II and IV CTZ, and small-unit night contacts in III CTZ), ~~conventional military operations by the Army of the Republic of Vietnam (ARVN)~~ small unit actions generally is the dominant positive predictor of the security variables.

Problems with the Schaffer study, however, may decrease the validity and reliability of his findings. As regards validity, it is questionable to compare the four CTZ's. For example, during January, 1967-March, 1968 over 90% of the small-unit actions occur in I CTZ and the remainder are spread across II, III and IV CTZs. Thus, a comparison of I CTZ with the others may not be valid. Furthermore, there is a problem with the reliability of the reporting system. II CTZ ceased reporting small unit actions after the third quarter, 1966 and resumed again in the third quarter, 1967 but these figures are too small to be credible. During this lapse in reporting, small-unit actions did occur. After February, 1968, the reporting appears to reflect more accurately the situation in the field. There is a similar but not as potentially biasing fluctuation in the reporting for III CTZ. Recall Schaffer's tentative conclusion that in the II, III and IV CTZs, the number of small-unit actions is one of the best predictors of population security. Before accepting this result, one should assess the degree to which it is an artifact of the erratic reporting and/or small frequencies of actual occurrence (since over 90% of the small unit actions take place in I CTZ).

In any event, Schaffer's tentative findings are of particular relevance for the present work because both studies draw on the same data source--the Office of the Assistant Secretary of Defense Systems Analysis, Southeast Asia Programs-(OASD/SA SEAPRO) Statistical Tables. Both studies use some of the same data on military actions, they are longitudinal, and both studies aggregate by CTZ units. The two studies differ in the process by which independent variables are allowed to explain effects, the way in which potential causal variables eventually can be manipulated experimentally and data quality control considerations. Schaffer allows any of his military variables "to cause" the population security measures. For example, he permits ARVN casualties and U.S. deaths due to hostile actions to explain the population security variables. The present study, on the other hand, considers the level of enemy attacks as outcomes of alternative U.S. military actions, e.g., large vs. small unit actions. Furthermore, in the present study, the variables thought to be causal predictors are candidates for an experimental manipulation. This exercise can assess the impact of alternate military actions on outcomes, e.g., later attacks. Data quality control concerns of the present study warn against accepting the reporting of military actions at face value and suggest procedures for assessing the degree to which random or systematic error bias the conclusions. These three steps--greater initial theoretical specification of the models, plans to manipulate their implications and data quality control--distinguish this study from those like Schaffer's.[1]

In addition to the RAND work, within the larger defense community there are a number of studies that seek to develop quantitative predictors of Vietnam war outcomes. Many of these are not explicit about the theoretical rationale for the inclusion of variables, their data bases generally are quite restricted in time and space, the data quality is unknown and the studies do not perform experimental manipulations of predictor variables to forecast probable consequences of alternative actions. Some of the studies outside RAND relevant here are such works as: James W. Johnson and Charles Anello, Measurement of Pacification Progress in Vietnam, McClean, Va.: Research Analysis Corporation (RAC-TP-309 September, 1968); William A. Niskanen, The Productivity of Major Military Forces in Vietnam: A Statistical Analysis of Three Years of War, Arlington: Institute for Defense Analyses (IDA, S-316, June, 1968) and Alfred I. Schwartz, A Method for Predicting Enemy Attack Locations (applied to the TET Attack in SVN), Arlington: (IDA/SWEG Staff Study November, 1968).

The Johnson/Annello study uses regression to study pacification "progress" in a similar manner as do Mitchell and Russo. The RAC authors measure pacification progress by changes in the proportion of the civilian population under the control of both sides from May, 1964-February, 1966 across 43 provinces. They aggregate the province data to the CTZ and also for the entire country. The Johnson/Annello results may be useful for predicting

---

[1] In addition to these RAND studies discussed above, consider the motivation and morale work by RAND analysts. This research may be of relevance to validate some of the relations obtained in the aggregate analyses above. It is important that aggregate analyses, such as the present work, validate their results by drawing upon the survey research and vice-versa.

future trends in population control as well as estimating force requirements. That is, they claim that one of the best predictors of population control is the force ratio--the number of friendly to enemy forces. This tentative conclusion fits in with prior work on force ratios as predictors of population control in Malaya. The emphasis on force ratios, however, overlooks the fact that the strategy of force use, e.g., a few massive sweeps versus many small coordinated ones, seems to be even more important regarding Vietnam outcomes. A further problem with this study, however, is that they transform the data to such an extent that the results may be too far removed from the data.

Possibly subject to a similar critique of "over-transformation" is the Niskanen regression analysis. In contrast to the Johnson/Annello effort, Niskanen's main dependent variables are casualty rates rather than population control. Niskanen does not find strong relationships between military actions and casualties. His study suggests that VC attacks and U.S. KIA, for example, do not covary systematically. (This relationship does seem to exist after the TET offensive in January, 1968). It appears, however, that trend and feedback effects in the data should be taken into account prior to accepting Niskanen's conclusions.

In contrast to both the Johnson/Annello and Niskanen efforts, the Schwartz work does not use regression analysis; rather, Schwartz employs a type of cluster analysis of VC attacks to discover regularities over time and space. For example, he concludes that enemy incidents during the 1968 TET offensive follow patterns prior to TET. Although he finds strong relationships in the data, their significance may be more for the tactical user in the field. Like the earlier Denton work, Schwartz does not relate VC incidents to U.S. military actions as does the present inquiry.

Moreover, the Schwartz study does not go into the more strategic aspects as the implications of the VC incident patterns for VC overall capabilities and/or intents. For example, if VC incident patterns are very stable over time and space but do not covary directly with US KIA until after TET, does this indicate a change in enemy intentions, e.g., to cause additional US KIA as a way of influencing U.S. domestic public opinion?

These studies discussed above demonstrate the feasibility of doing quantitative analyses of socio-economic conditions and/or military actions to predict outcomes such as population security and/or casualties. Three shortcomings cutting across these efforts are their general failure to make clear the theoretical basis for the modelling, their lack of experimental manipulations of independent variables to assess probable effects of alternative military actions and the absence of data quality controls.

Work outside the defense community makes more progress in theoretical specification and experimental manipulation. For example, while students at Stanford University, Jeffrey Milstein and William Mitchell explicitly specified their mathematical models on the basis of a priori theoretical ideas, and they conducted experimental manipulations of possible causal variables. Here follows a brief summary of their methodology and findings as they relate to the present inquiry.

Three goals of the Milstein/Mitchell work are to describe, explain and forecast the escalation of actions and reactions in the Vietnam war. They focus on what is said and done by each of the belligerents vis à vis their reported combat operations. Using data from January, 1965 through December, 1967, they test implications of models of the war derived from public statements of policy-makers in the National Liberation Front (NLF), North Vietnam and in the United States Milstein and Mitchell construct models of selected aspects of the war employing a "mediated stimulus-response" idea to predict actions, reactions and outcomes. The essence of this idea is that the responses of decision-makers of the U.S. constitute environmental stimuli for decision-makers in North Vietnam and the NLF; the decision-makers perceive U.S. threats in such a way that they increase their capabilities and express hostile statements to U.S. decision-makers, who perceive these responses as confirming their original suspicions. The U.S. policy-maker reacts to the North Vietnamese increase in capability and hostility by escalating from verbal threats to limited bombing of North Vietnam, which in turn causes the North Vietnamese to escalate even further. This theoretical framework forms a basis for specifying Milstein/Mitchell models.

Using regression analysis across 36 months, they report some of the following tentative findings:

1. Escalation of the bombing of North Vietnam provoked a subsequent counter-escalation of North Vietnamese troop commitments.
2. Bombing de-escalation, on the other hand, led to increased North Vietnamese and Viet Cong willingness to negotiate.
3. While sustained bombing inflicted physical damage on North Vietnam and did impede their troop and supply movements, the increases in the bombing rate provoked increases in infiltration by North Vietnam.

Based on some of the regression results, they perform an experimental manipulation of variables--exercises of their computer simulation. They state that:

Once we have developed a model of the Vietnam conflict that can make accurate predictions, we can use it to explore the probable effects of experimental changes in selected variables. Policy-makers, for example, might wish to determine probable outcomes of alternative strategies without actually implementing them on the battlefield or at the negotiating table. Simulation can be used to make forecasts of this kind so long as the relationships among the variables in the system remain constant.

What would have happened, for example, if President Johnson had continued the peace overtures of December, 1965-January, 1966, the period of the 38-day bombing halt and major diplomatic activity? We have sought to answer this question by exercising the simulation of the 36-month model. We explored this by experimentally changing those variables which President Johnson might have used to signal the North Vietnamese and Viet Cong of his desire to de-escalate the war. We considered the following four activities:

1) Halting the bobming of North Vietnam in January, 1966, and continuing that halt from that time on; 2) Reducing the bombing attack sorties in South Vietnam ten percent per month starting in January, 1966, and continuing thereafter; 3) Reducing the major ground operations (battalion-size or larger) undertaken by the U.S. in South Vietnam by ten percent per month starting in January, 1966, and continuing thereafter; and 4) Reducing the increases of U.S. troops in South Vietnam by 15,000 (approximately one division) per month from January, 1966 and thereafter to the present. These proposed changes are of special interest to us today given that the bombing of North Vietnam was halted on November 1, 1968, and that the reduction of combat operations and troop levels are likely steps to be taken in any further deescalation of the war.

The way in which we exercise the simulation is to make the experimental variable exogenous in the system, i.e., we allow it to have only the specified experimental values rather than the values it would otherwise take on as a dependent variable in the system. In the exercises reported herein, the initial data points for all variables are from January, February, and March, 1965 and the regression model used is that derived from 36 months of data from January, 1965, through December, 1967. These exercises make statements about what might have been. Before one can validly apply such exercises to similar changes in the future, he must be sure that the relationships among variables remain the same. The results of this exercise are presented in the following table.

Table 1

RESULTS OF SELECTED EXERCISES OF THE COMPUTER SIMULATION

Dependent Variable	Halt Bombing of NV	Reduce Bombing in SV	Reduce Ground Operations	Reduce U.S. Troops
VC Defectors	More; increased in 1967, rather than decreased; continue to increase after 1967.	Fewer in 1967.	More; increased in 1967, not decreased.	Increased last half of 1967.
VC Killing of Civilians	Decrease from March, 1966, on, rather than increase.	Fewer in late 1966, then increase to same level as exact values.	Much fewer throughout.	Decrease from mid-1967.
VC Abductions of Civilians	Decrease after May, 1967, rather than increase.	Fewer in late 1966, no change otherwise.	Fewer throughout.	Decrease from mid-1967.
Popular Support for LBJ	Greater overall.	Same downward trend, but greater support.	No change.	Greater.
Piastre Value (Confidence in GVN)	First increases, then collapses to nothing.	No change.	No systematic change.	Consistently lower, but not catastrophic.
U.S. Troop Commitments	Accelerate much faster.	No systematic change,	Greater.	---
U.S. Ground Operations	Decrease from Oct., 1966-March, 1967, then increase greatly to exact level and above.	Fewer.	---	Fewer in 1967.

Table 1 (continued)

RESULTS OF SELECTED EXERCISES OF THE COMPUTER SIMULATION

Dependent Variable	Halt Bombing of NV	Reduce Bombing in SV	Reduce Ground Operations	Reduce U.S. Troops
Bombing SV	Less in 1966-67; no difference in trend.	---	No change.	Less in late 1967.
Bombing NV	---	No change.	No change.	Fewer missions.
U.S. Killed in Action	Fewer in 1967 than exact.	No systematic change.	No change.	Fewer, late 1967.
NV + VC KIA	Much fewer; less than U.S.	No systematic change.	No change, 1966; fewer in 1967.	Much fewer.
ARVN KIA	More 1966, less 1967.	No change.	No change.	No change.

Of special interest for the present study is the Milstein/Mitchell methodology. They first evaluate alternative possible causal links in terms of their theoretical expectations, then they combine the most adequate relationships into a system of interacting variables--the simulation. Next Milstein and Mitchell experimentally manipulate selected variables to assess the probable consequences on criteria of interest. Also some of their substantive conclusions are relevant to the present inquiry. For example, the forecasts of the probable effects of a reduction in U.S. ground operations relate to the present study's focus on the contrast between U.S. large and small-unit actions. The Milstein/Mitchell effort includes U.S. large actions only--number of ground operations of battalion size or larger--whereas the present study has small unit actions as well. They aggregate the military operations data to South Vietnam as a whole, while the present inquiry begins with I CTZ and correlates across time. Recall their forecast that a reduction of large unit actions may result in fewer killings and abductions of civilians; the forecast could be replicated and validated in a study such as the current one. (Cf. J.S. Milstein and W.C. Mitchell, "Computer Simulation of International Processes: The Vietnam War and the Pre-World War I Naval Race," Peace Research Society International, Papers, XII (1969); also, "Dynamics of the Vietnam Conflict: A Quantitative Analysis and Predictive Computer Simulation," Peace Research Society International, Papers, X (1968); Milstein, Changes in U.S. Domestic Public Opinion Support and Alternative U.S. Policies and Military Actions in the Vietnam War, 1965-1967, East Lansing: Michigan State University, March, 1969, 37 pp. (mimeo); and Milstein, Inferring Policy Changes from Changes in Relationships Among Conflict Behaviors and Outcomes in the Vietnam War, 1965-1969, East Lansing: Michigan State University, April, 1969, 27 pp. Followin

Milstein and Mitchell's lead, here are some next steps; disaggregate by the four CTZ then by 44 provinces, using more detailed data; adopt an appropriate set of theoretical propositions given the particular focus; construct a simulation, make forecasts and validate them. These were, indeed, the goals of the Vietnam Computer Simulation project (VCS) at The RAND Corporation during 1968-69. The present study, however, only performs some of the above steps. The analysis is for I CTZ; this study pays more attention to theoretical ideas for specifying the models; finally, the study evaluates links within the models--a preliminary to the development of a computer simulation.

The extent to which these goals can be achieved depends on the quality of the ideas and data. Very few of the studies reviewed above provide theoretical specification of their models in advance of analysis nor do they discuss the possible effects of data quality. The following statement is a general description of the error problem, some consequences of error given the purposes here and some possible solutions.

DATA QUALITY CONTROL

Types of Error: Random or Systematic[2]

Any quantitative analysis of the Vietnam War must contend with the general impression that the official information and data misrepresent the true situation in the field. For example, two scholars state that:[3]

> There is widespread agreement that ... official figures on Vietcong casualties are exaggerated. In the first place, it is often impossible to differentiate between a dead peasant and a dead Vietcong soldier, and a significant number of what have been reported as Vietcong killed in action have been civilians. In addition, the means for determining the number of killed, even on the basis of a body count are often highly unreliable, tending to exaggeration rather than underestimates. With regard to this matter, see the instructive article by the New York Times correspondent Charles Mohr (June 27, 1966), who concluded: "Statistically the war has been won several times already."

The Kahin and Lewis statement regarding the inclusion of civilian deaths as a significant proportion of enemy casualties is a question of one type of validity: does the variable in fact tap what it purports to measure? One alternative is to assume that the proportion of civilian casualties is a constant value (e.g., 10%) across units of observations. If this were true then correlation coefficients of two variables across time or space would not be affected. Another alternative is to assume that the civilian proportion is a function of some known variable such as U.S. air-strikes. It is reasonable to assume that air-strikes kill a higher proportion of civilians than ground actions because of the difficulty in differentiating VC from civilians during air strikes. If this were true, one can control for the effect of air-strikes

---

[2] Acknowledgments to Rudolph Rummel whose paper, "Dimensions of Error in Cross National Data," forthcoming, 1969 (mimeo) forms an important point of departure for the present essay.

[3] G.M. Kahin, J.W. Lewis, The United States in Vietnam. (New York: Delta Press, 1967), p. 188 footnote on Table 4.

(statistically) and then look at the relation of U.S. ground actions and VC casualties. A problem with this alternative is that it overlooks the fact that the air-strikes may be instrumental in increasing VC casualties as well as civilians. If one has alternative sources of evidence on enemy casualties and can get a consensus as regards what it indexes, this should increase confidence in the validity of the measure.[4]

Kahin and Lewis also discuss the reliability of data: do the reported values correspond to the true values? There are two ways of viewing reliability: consider the error of measurement as either random or systematic. Random error might result from equal amounts of under-reporting and over-reporting vis-à-vis the true values. Systematic error is either under or over-reporting. There are some analysts who contend that the so-called body counts of enemy killed-in-action (KIA) under-estimate and at times over-estimate the true values. A November, 1968 survey of students and faculty returnees from Vietnam then at the Army War College provides some evidence in this regard: Lt. Col. R.A. McMahon, "The Vietnam 'Body Count': Is It Accurate or Inflated?" (Ann Arbor News, June 29, 1969, p. 43). All the respondents held positions in the the military which qualify them to comment from first-hand experience on the body count. Of the 65 officers in the survey, 43 commanded combat battalions and 5 commanded brigades. Questionnaires elicited information regarding whether enemy casualties reported by their units were:

1. An accurate, physical count of every body that could be found;
2. The best, most honest estimate that could be made at the time;
3. Or a combination of (1) and (2), depending upon circumstances.

[4] Another example of this type of validity is in the studies by Milstein and Mitchell. They use the black market value of the South Vietnamese currency (piastre) as an indirect indicator of the confidence the South Vietnamese people have in this regime. One also suspects that the piastre value is a measure of the VC economic behavior as they exchange piastres for dollars e.g., in order to purchase rice.

Twenty-two officers (approximately 35%) responded that enemy casualties in their units were reported by actual body count. Most of these officers thought that many more enemy casualties had been inflicted by their units than they reported. They claimed that many more enemy died than were counted by U.S. forces on the battlefield due to (1) inability to assess the results of artillery and air strikes in inaccessible places; (2) underreporting of those who died later of wounds and (3) loss of information because the enemy tries to recover his dead before breaking contact. Other results of the survey relevant here are as follows. Only three officers (5%) said that in their units "reported enemy casualties" were based exclusively on the best estimate that could be made at the time. Forty officers (about 60%) said that reported enemy casualties were based on a combination of body counting and estimates. Many of these officers agreed that even though they were estimated, enemy casualties were probably understated for the same reasons as above. On the other hand many of these officers reported various kinds of pressures from their headquarters to turn in high body counts. Consider the following excerpts from their questionnaires.

1. Body counts led to "grave-digging" by some units to get numbers.

2. Sometimes the score was in the enemy's favor and this is when the cry came in from higher headquarters for another body count.

3. ... I usually made an estimate of the "body count" to please higher headquarters.

4. (A senior officer addressing the officers of a unit that had just suffered very heavy casualties: "Well, we lost just about all of (company size unit), so let's have a good body count so we can tell the survivors they did well."

The responses and comments by officers tend to support the view that the body count figures are sometimes under and over estimates of the true values, e.g., that they contain <u>random error</u>. In agreement with Kahin and Lewis regarding the tendency toward exaggeration, however, are many U.S. newspapers and magazine reporters with experience in Vietnam. Typical views with respect to the <u>systematic error</u> of overstatement are as follows:

1. ...almost every reporter in Vietnam had his own personal example of inflated reports of enemy dead in battles that he himself had observed. (E.G. Martin, <u>Newsweek</u>, February 12, 1968, as cited in McMahon.)

2. Military observers here tended to look askance at the estimated total of 7,500 killed over the last eight days, since both the South Vietnamese and American military commands . . .at times . . .exaggerate enemy casualties. (D. Robinson, <u>New York Times</u>, August 28, 1968, as cited in McMahon.)

In short, the statements of U.S. military officers above contend that random error characterizes enemy casulty figures. The scholars and press statements hold that the systematic error of overstatement inflates the casulty figures. What are some consequences for this preliminary study if random [redacted]or systematic error are in the data?

<u>Consequences of Error: Univariate or Bivariate</u>

Consequences of random or systematic error depend upon the purpose of the particular inquiry. For example, if the concern is with analyzing a single monthly time series of capability such as North Vietnamese (NVA) infiltration, error in the data collected on this variable may be <u>univariate</u> random or systematic.

Univariate random error is:

$$x = X + e \qquad E(e) = 0 \tag{1}$$

where e denotes random error in the single variable, X is the true value, x is the empirical value, and E the expected value. In other words, an observation, x, is a sum of a true value and an error; errors for all the true values vary in such a way that their average value is zero. For infiltration, this would mean that one could get a yearly total or average for that variable which would be very close to the true value because random error "cancels out" when the monthly statistics are aggregated. Univariate systematic error is:

$$x = X + e \qquad E(e) \neq 0 \tag{2}$$

The error on infiltration, for example, would overestimate or underestimate the true value. The claims by the officers in the survey summarized above imply univariate random error; the assertions by Kahin and Lewis as well as the reporters imply univariate systematic error. If one knows the degree of and direction of systematic error, corrections are relatively simple. Without such knowledge, estimates of error ranges can accompany the empirical values in order to give the user a feel for the quality of the evidence, given his particular purposes and need for accuracy. Univariate random or systematic error will not, in general, distort the results of the multiple regressions in this study. Bivariate systematic error is the only serious danger here.

This problem of bivariate random or systematic error is more difficult to handle statistically. Bivariate random error is:

$$x = X + e^{*} \qquad y = Y + f^{*} \tag{3}$$

$$\mathrm{Cov}(e^{*}, Y) + \mathrm{Cov}(f^{*}, X) = \mathrm{Cov}(e^{*}, f^{*}) = 0 \tag{4}$$

where the asterisk distinguishes univariate random error, e, from bivariate random error, e^{*}; Cov is covariance or the cross products of two variables taken as deviations from their respective means across the observations, N.

If the investigator's purpose is to obtain unbiased estimates of population regression coefficients, he assumes that the random error is in the dependent variable and that the independent variable is relatively error-free.[5] For example, the military officers assess the body count as producing estimates which are sometimes under and at other times overstatements of the true value--random error. The use of enemy casualties as a dependent variable with random error can yield unbiased estimates of regression coefficient if the independent variables such as U.S. battalion size and larger operations and total numbers of sorties flown are relatively error-free. This study assumes that the data on U.S. actions are relatively error-free in comparison with the dependent variable, e.g., enemy attacks.[6]

---

[5]When a more complex analysis is done, the assumptions required are sometimes different. For example, one can use a three-step model, with the independent, intervening and dependent variables. Here again, there should be minimum error in the independent variables and minimum systematic error in the dependent variables. Ordinary errors in the intervening variables do not tend to distort the relations between the independent and dependent variables. Thus one can work with "body counts" as an intervening variable, without high risk of distorting the conclusions.

[6]When error is in both independent and dependent variables, the investigator works with a structural model and the empirical regression coefficients underestimate the true regression. Techniques are available to correct the coefficients. (cf. Rummel, op. cit., 1969; M.G. Kendall, and Alan Stuart, The Advanced Theory of Statistics, II, New York, Hafner Publishing Co., 1961); (Curbin, "Errors in Variables." Revue of the International Statistical Institute, 22 (1954), 23-32); (A. Mandansky, "The Fitting of Straight Lines When Both Variables are Subject to Error," Journal of the American Statistical Association, 54 (1959), 173-205); (R. Geary, "Non-Linear Functional Relationship Between Two Variables When One Variable is Controlled," Journal of the American Statistical Association, 48 (1953), 94-103_; (H. Scheffe, "Fitting Straight Lines When One Variable is Controlled," Journal of the American Statistical Association, 53 (1958), 106-117). Consider one statistics text regarding error: "If the errors affect only values of the dependent factor, and if they are not correlated with the true values, their presence tends to lower the correlation and to increase the standard error of estimate, but does not tend to change the slope for the universe. If, however, uncorrelated errors are in the independent factor, that not only tends to lower the correlation and to increase the standard error of estimate, but also tends to decrease the regression below the true value." See M. Ezekiel and K.A. Fox, Methods of Correlation and Regression Analysis (Rev. ed.) N.Y.: Wiley, 1959.

If the investigator is not concerned with the regression of enemy casualties and/or attacks on U.S. actions but with correlations, errors have different consequences. In the correlation model, the effect of bivariate random errors, e^* and f^* in X and Y is to reduce the correlation. For example, a correlation of .69 between U.S. small-unit actions and VC/NVA attacks would underestimate the true correlation if the error in both variables were random. If statistical inference is of interest and one wants to avoid rejecting a true hypothesis about these two variables, the investigator might raise the level of significance from .05 to perhaps .20 depending upon the degree of error and the costs of rejecting the true hypothesis.

Bivariate systematic error is a situation where assumption (4) is violated, e.g., when

$$\begin{aligned} &\mathrm{Cov}\ (e^*, Y), \text{ and/or} \\ &\mathrm{Cov}\ (f^*, X), \text{ and/or} \\ &\mathrm{Cov}\ (e^*, f^*) \neq 0 \end{aligned} \tag{5}$$

That is, when one or more of the error covariances in (4) does not equal zero, one has bivariate systematic error. Whether the effect of bivariate systematic error is to lower the empirical correlation from the true correlation depends on two ratios: the ratio of the standard deviation of the error to that of the true value; the ratio of the covariance of its error with the other variable to the covariance of the true values, e.g.,

$$r_{xy} < r_{XY} \tag{6}$$

$$\frac{\sigma_{e^*}}{\sigma_x} > \frac{\mathrm{Cov}\ (e^*, Y)}{\mathrm{Cov}\ (X, Y)} \tag{7}$$

Where r_{xy} is the empirical correlation and r_{XY} is the true correlation; σ is a standard deviation. For example, a correlation of .69 between U.S. small unit actions and VC/NVA attacks would be a conservative estimate of the population correlation if the bivariate systematic error assumptions in (7) hold.

Regression and correlation are the principal techniques here. As regards the correlation coefficients, one could employ measures of random and systematic error to ascertain the degree to which they co-vary with substantive variables of interest. For example, here is one type of systematic error measure relevant to correlational analysis that would be of interest were data available: scales of whether U.S. forces are present or not in a province during the data collection effort. The assumption is that their presence, size, etc., might co-vary with higher quality reporting than would otherwise be the case. Related measures are their length of stay or the frequency of visits. Scales could be devised to see if the presence, size, length and frequency of visits of U.S. forces correlate positively or negatively with the substantive variables of interest. The present study does not develop measures of systematic error because it assumes that random error is the most likely type of error in the dependent variable-- VC/NVA attacks, and that the independent variables are relatively error free.

In short, the consequences of random or systematic error depend upon the purpose of the investigator. Some policy concerns require estimates of single indicators of "progress" over time. Error in such data may be univariate random or systematic. One can evaluate and minimize such error whenever the interest is in the univariate distributions. Other policy considerations may need estimates of the movement of two or more indicators over time such as the effect of U.S. small-unit actions on enemy attacks across months. Error in such data may be bivariate random or systematic. One could anticipate the consequences of bivariate random error regarding conservative estimates of coefficients and increase significance levels if appropriate. Moreover, it is possible to devise measures of both bivariate random and systematic error in order to ascertain their effects on policy conclusions. Though bivariate systematic error may bias the present preliminary results, there is an explicit awareness of this possibility which may mitigate against misleading results.

Specifically regarding the present study, available evidence suggests that random error characterizes such variables as enemy attacks. Bivariate systematic error might lower or raise the correlations of enemy attacks and U.S. small-unit actions, but this study concludes that the most probable errors in the data are random. The most likely possibility is that the correlation coefficients will be under rather than overestimates and that the regressions will be reliable. These statements on types of error--random or systematic--and some consequences of error regarding univariate or bivariate purposes illustrate a concern with data quality control. Though important, the concern with the reliability and validity of data is secondary to evaluating the validity of alternative ideas.

RESEARCH DESIGN

Theoretical Orientation and Military Strategies

Two broad sets of ideas guide this inquiry. The first asserts that the Vietnam War is similar to any conventional war, such as the Korean effort in the 1950's. A main task then is to search for and then destroy the enemy. The strategy of attrition tries to locate the enemy's main force units and seeks to destroy them. The strategy of attrition generally involves major allied military operations of battalion size and larger. Primary indicators of progress in such a strategy are the body count of enemy casualties and the so-called kill ratio--a ratio of friendly to enemy casualties. A focus of attention of the attrition strategy is on infiltration routes and enemy supply routes. The interdiction allegedly slows down and/or raises the cost of infiltration and supply operations. A further aspect of the focus on infiltration and supply routes concerns the need for allied outposts in the remote areas. The dispersion of allied forces geographically is, in part, a function of the infiltration and supply routes. This emphasis on a positional rather than a more mobile role is supposedly slight because of the role of helicopters in ferreting troops to trouble spots quickly. A major aim is to locate enemy jungle bases where he keeps his supplies and trains his men.

A second set of ideas claims that the Vietnam war is unlike a conventional war. It is a "People's Revolutionary War" (PRW). The PRW may go through three phases. The build-up or defensive phase; the guerrilla war phase to achieve some favorable ratio of force vis à vis government troops; the offensive phase to take over the country. Throughout a country, there may be an overlap between the phases at any given time. The primary weapon of the insurgents is an underground political organization within the population. A secondary weapon is the guerrilla force which depends on the underground organization for support. A first aim of the insurgency is to gain momentum by capturing more weapons, ammunition and explosives. Attacks on isolated government posts aid this goal. Optimally, these attacks should result in government forces extending themselves into the countryside. Government response may be to look for the regular insurgents' units at company and battalion strength or local armed units at platoon and company strength. Generally overlooked are the communist cells and supporters in the population with the accompanying district committees. Thus, the government priority is to defeat the military insurgency rather than the political subversion.

The People's Revolutionary War is largely a political effort designed to reach a long-range outcome without being too dependent on winning military victory in conventional battles. Sir Robert Thompson's books--No Exit From Vietnam, New York: McKay (1969), and Defeating Communist Insurgency, London: Chatto (1966)--make this argument well. In No Exit . . ., he suggests that "Even in defeat an insurgent movement will willingly sacrifice all its military guerrilla units if only it can preserve some of its underground organization intact within a country to await a more favourable opportunity" (p. 33). Furthermore, he asserts that "the so-called bases within the jungles and swamps are operating bases for the guerrilla units where supplies can be stocked and troops trained. If the latter are captured, the success is only temporary because they can be rapidly replaced, even in the same area as before, when the forces who captured them withdraw" (p. 34).

Thompson traces the history of the search and destroy strategy of attrition, which began about late Spring, 1966. Thompson argues that American strategy in Vietnam failed for two major reasons. First, attrition and bombing took the direct line of greatest expectation and toughest resistance for the Viet Cong and North Vietnamese. These forces were fully prepared and the strategy failed to destroy the main force units or to reduce infiltration. A second reason for the failure of American strategy is that the organization and structure of the enemy was never threatened. This second reason is the principal one for Thompson, who asserts that unless the government cracks the underground organization the guerrilla units and main line forces will not be defeated (p. 144; p. 166). He argues that ". . . the enemy underground organization should be ruthlessly weeded out, for which purpose the (government) intelligence organization needs to be steadily improved and a greater number of small unit actions initiated with emphasis on ambushes at night" (p. 198). In Thompson's other book, Defeating Communist Insurgency, he points out that "In an insurgency it is not the major operations that defeat the insurgents. It is a high ratio of contact in minor operations, based on good intelligence . . ." (p. 88).

In addition to Thompson, other authors provide a theoretical rationale for the political nature of an insurgency. Consider some of the ideas in Mao Tse-Tung, On Guerrilla Warfare, New York: Praeger (1967); F.M. Osanka (ed.), Modern Guerrilla Warfare, New York: The Free Press (1962); Roger Trinquier, Modern Warfare, New York: Praeger (1964); George Tanham, War Without Guns, New York: Praeger (1966) and Communist Revolutionary Warfare, New York: Praeger (1966). For example, Mao states that, in general, guerrilla units disperse to operate. When the enemy is in overextended defense and sufficient force cannot be concentrated against him, guerrillas must disperse, harass and demoralize the enemy. When guerrillas are encircled by the enemy, they should disperse to withdraw, moreover. Mao also hints that the main danger to guerrillas comes from small-unit actions. He argues that each guerrilla base has its own peculiar problems of attack and defense. In general, the enemy tries to extinguish these bases by dispatching <u>numerous bodies over a number of different routes</u> (emphasis added). This has to be anticipated and the encirclement broken by counterattack. When the enemy is dispersed without reserves, the main forces should attack him. Mao seems to imply that small-unit actions force guerrillas to fight, whether they planned to fight or not. Mao gives strong emphasis to the need for guerrillas to fight only on their own terms. For example, he advises guerrillas to retain the initiative, to fight no unprepared engagements and to fight no engagements in which there is no assurance of victory (cf. Mao (1967), p. 102, 111 and in Osanka, pp. 270-272; also cf. Tanham, <u>Communist . . .</u>, p. 11, p. 17).

These ideas of Mao seem to apply to an early stage of guerrilla warfare rather than to conventional war where the concentration of attacking forces is an accepted necessity of battle. Thus, an important problem is to identify the stage of warfare and apply the strategies which promise to be most effective for that phase. Roger Trinquier's ideas also are relevant within this guerrilla phase of fighting. He claims that large-unit sweeps temporarily disperse guerrilla bands. These sweeps usually attempt a surprise encirclement of a well-defined zone in which guerrillas are thought to be located while mobile elements conduct a mopping-up operation. Trinquier argues that only a longer operation in the countryside which permits police-type operations can succeed in rooting out guerrillas. Like Thompson, Trinquier concludes that an important objective is the destruction of the enemy politico-military organization. It seems reasonable to expect that small units with adequate logistical support and integration with indigenous personnel would be more able to destroy the enemy organization than the large sweeps.

The general discussion above acknowledges two principal alternative ideas regarding Vietnam. One is a view that stresses large sweeps to destroy the enemy; the other emphasizes rooting out enemy political organization and keeping him off balance with large numbers of small actions.

Design Decisions

With this brief theoretical discussion in mind, now consider the following specific design decisions: selection of independent, dependent and intervening variables; discovery of the most appropriate time lags and a way to decrease trend effects in the data; the selection of the spatial-temporal domain; the modelling technique to use; the derivation and specification of the models.

The theoretical literature suggests that the research design divide U.S. military actions in such a way that the more conventional search and destroy operations separate themselves from the more politically oriented small unit activity. A problem with this dichotomy is that search and destroy operations often combine small- and large-unit actions into coordinated efforts. Consider the following scenario. Small units of platoon and company size go into enemy territory to engage their main forces. Once the small units make contact, allied commanders call in air support and large-scale reinforcements. Hence, it may not be possible with the current data to assume that the small-versus large-unit actions represent alternative strategies. The analysis, however, assumes that the frequency of small-unit actions and battalion days of operation may indicate the relative preferences of U.S. generals regarding a more political strategy (small units) or a more conventional military strategy (large units).

Thus, the two main independent variables are the frequency of U.S. small-unit actions and the number of battalion days of operation.

There are several candidates to be dependent variable in this inquiry. One could try to explain enemy casualties, population control, defectors, etc. The variable that seems linked most to the theoretical assumptions, however, is enemy activity. That is, recall Thompson's assertion in Defeating Communist Insurgency that "In an insurgency it is not the major operations that defeat the insurgents. It is a high ratio of contact in minor operations, based on good intelligence..." (p. 88). In addition, recall Mao's implication that small-unit actions force guerrillas to fight, whether they planned to fight or not. Finally, consider Trinquier's assumption that large unit sweeps temporarily disperse guerrilla bands. Based on the reasoning of Thompson, Mao and Trinquier, it seems that one should evaluate the relative effectiveness of small- versus large-unit actions on the basis of which variable has the most impact on enemy activity. Note that a common yardstick of force effectiveness is enemy casualties. The present inquiry considers the casualty yardstick to be inadequate because it is not as theoretically-based as is the attack variable.

In short, there are three primary variables in the present work: total U.S. small-unit actions with or without contact; U.S. battalion days of operation; VC/NVA attacks. A theoretical assumption links these variables

Table 2: Correlations with

Time Lags, N $\leq$ 19 $\geq$ 15 Months*

	X_1	X_2	X_3	X_4	X_5	Z_6	Z_7	Z_8	Z_9	Z_{10}	Y_{11}	Y_{12}	Y_{13}	Y_{14}	Y_{15}
X_1(t-4)		96	91	81	66	92	90	91	90	90	89	84	70	49	32
X_2(t-3)			96	91	79	88	89	88	90	86	84	89	76	60	46
X_3(t-2)				96	89	82	82	84	84	83	85	84	82	68	55
X_4(t-1)					95	72	72	73	76	73	75	85	78	77	60
X_5(t-0)						60	59	60	61	63	65	75	80	74	69
Z_6(t-4)							88	82	84	86	76	80	77	50	22
Z_7(t-3)								89	84	87	78	71	57	52	44
Z_8(t-2)									90	86	87	74	50	35	47
Z_9(t-1)										91	87	84	55	30	29
Z_{10}(t-0)											92	80	60	31	26
Y_{11}(t-4)												75	57	28	36
Y_{12}(t-3)													70	50	23
Y_{13}(t-2)														66	34
Y_{14}(t-1)															55
Y_{15}(t-0)															

X = U.S. small unit actions

Z = U.S. battalion days of operation

Y = VC/NVA attacks

t = time lags in months

* The values in the cells are product moment correlation coefficients multiplied by 100 to remove decimals.

in the following way. Earlier VC/NVA attacks cause later U.S. small-and/or large-unit responses, which in turn affect the VC/NVA attacks even later. Note that this sequence has three independent variables predicting to one dependent variable. The extra independent variable is, of course, prior values of the dependent variable. Thus, the analysis assumes that U.S. activity intervenes between prior and later VC/NVA attacks.

Given these variables, next design questions involve the most appropriate time-lags to employ and how to decrease trend effects in these variables. Table 2 contains a full list of fifteen variables based on the three original ones. Variables X_1 - X_5 refer to small-unit actions; variables Z_6 - Z_{10} refer to battalion days of operation; variables Y_{11} - Y_{15} refer to VC/NVA attacks. Variables X_1, Z_6 and Y_{11} each refer to a four-month earlier period than X_5, Z_{10} and Y_{15}. Likewise, X_2, Z_7 and Y_{12} refer to a period, t-3. See the row of Table 2 for a full listing of the time-lags associated with each variable number.

Consider the following examples to illustrate how to use Table 2. What is the correlation of U.S. small unit actions and VC/NVA attacks at time-zero (t_o)? Find the cell where variables X_5 and Y_{15} intersect; the product-moment correlation, r, = .69. What is the correlation of U.S. battalion days of operation and VC/NVA attacks at time zero? Find the cell where variables Z_{10} and Y_{15} intersect; r = .26. Note that U.S. small-unit actions explain almost 50% of VC/NVA attacks while U.S. battalion days of operation explain less than 7%. (The square of the correlation coefficient, r, provides the proportion of variance that two variables share.)

The correlations in Table 2 allow one to discover the optimal time-lags to employ in the models. For example, for which time-lags is there the highest correlation between U.S. small-unit actions or battalion days of operation and VC/NVA attacks? Figure 1 contains a plot of the correlations between VC/NVA attacks at time-zero (t_o) versus U.S. small-unit actions and U.S. battalion days of operation at t - 4 through t + 4. For example, recall from Table 2 the correlation of .69 between VC/NVA attacks and U.S. small-unit actions at time zero. One also can find the .69 value in Figure 1. See the intersection of time zero on the vertical axis and about .7 on the horizontal axis for the line plotting the correlations of U.S. small-unit actions.

Ideally, the lines in Figure 1 should peak twice. One peak should be within the periods t - 4 to t_o; the other peak should be between t_o and t + 4. In fact, however, the correlations between U.S. small-unit actions and VC/NVA attacks increase almost linearly from t - 4 through t_o to t + 2. For negative lags, the highest correlation between U.S. small-unit actions and VC/NVA attacks is at t_o, where $r = .69$. For positive lags, the highest correlation between U.S. small-unit actions and VC/NVA attacks is at t + 2, where $r = .80$. Note that the positive lags refer to a situation where VC/NVA attacks is an exogenous variable. In this case, prior attacks may be thought of as a predictor of later U.S. small-unit actions. For negative lags, prior U.S. small-unit actions predict to later VC attacks, which is now an endogenous variable.

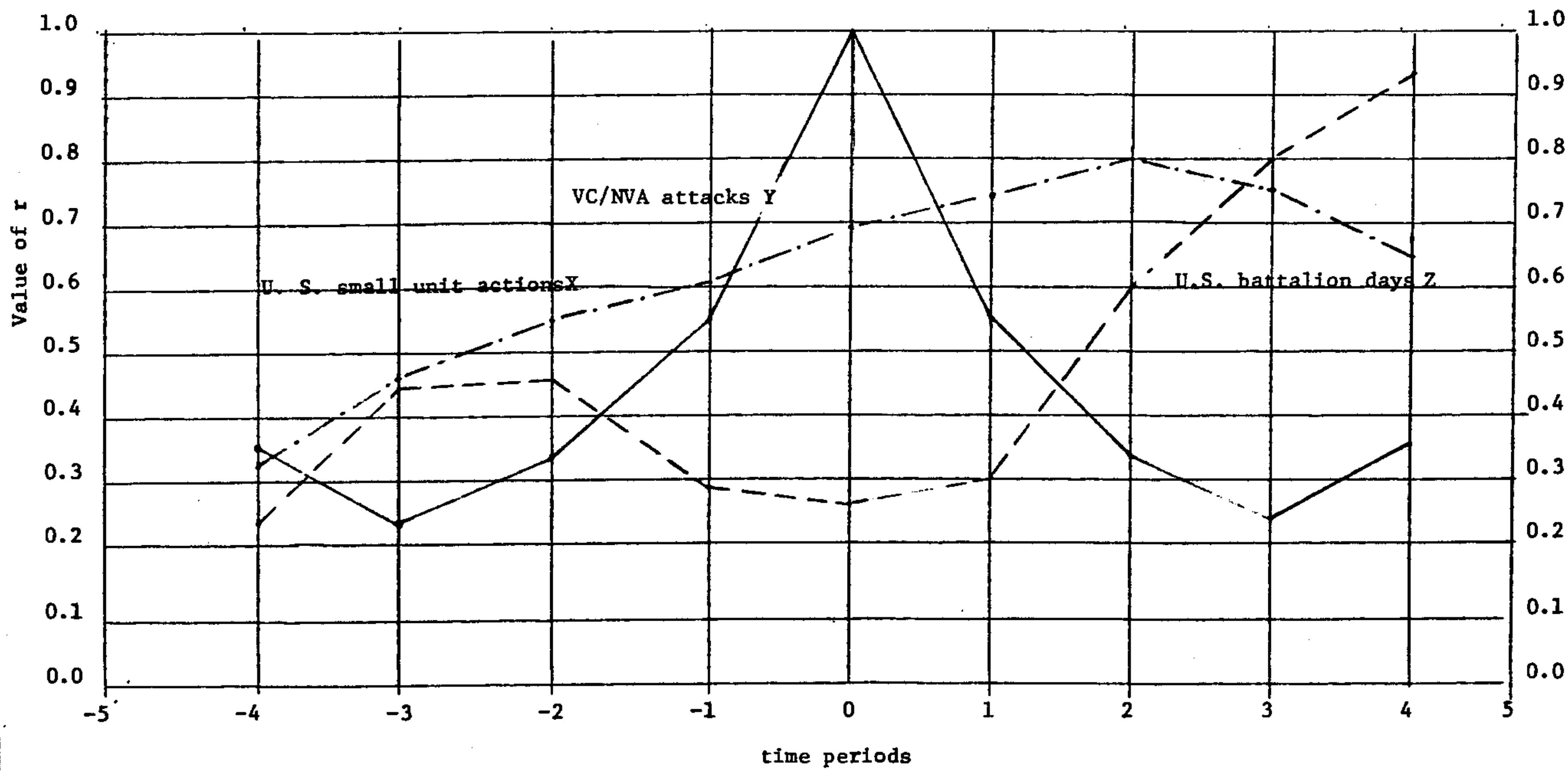
As regards U.S. battalion days of operation, it peaks at t - 2 and t + 3. That is, U.S. battalion days predict best to VC/NVA attacks two months later, likewise, VC/NVA attacks predict best to U.S. battalion three months later.

Figure 1

Correlations Between VC/NVA attacks at Time

Zero versus attacks, U.S. small-unit actions and U.S. battalion days of operation at

Time Points t - 4 through t + 4



The results of Figure 1 one combined with an intuitive sense of the variables in question yield sufficient information as regards the most appropriate time-lags. (See below under the discussion of models to find the time-lags actually employed).

In addition to the time-lag selection, there is the problem of decreasing the trend effect in the three substantive variables: U.S. small-unit actions, battalion days of operation, and VC/NVA attacks. Plots of each variable against time generally reveal high linear correlations with time. An exception is that VC/NVA attacks have only a modest correlation with time, r = .40. Moreover, the correlation of VC/NVA attacks with attacks four months later (autocorrelation) is only .36.

The technique for decreasing the trend effects of U.S. small-unit actions and U.S. battalion days of operation is to calculate first differences. For example, to compute the first difference, subtract the values of small-unit actions at t - 2 from t -3 across each of the time periods. The first difference or rate of change is like velocity, whereas the second difference is similar to acceleration--a changing rate of change. The correlation of the first difference of small-unit actions with time is only -.13. Likewise, the correlation of the first difference of U.S. battalion days of operation with time is zero. Thus, the calculation of first differences decreases the trend effect in the bivariate correlations.

The models analyzed below, however, are multivariate ones. Although the trend effects may be reduced in the bivariate case, they may not decrease in the multivariate models. Thus, there are models that include time as an explicit variable and those that do not.

A further design decision involves the selection of the spatial-temporal domain. An initial goal was to compare I CTZ with the rest of South Vietnam--II, III and IV CTZs. Recall the fact that some 90% of the U.S. small-unit actions reported in South Vietnam occur in I CTZ. With such a concentration in I CTZ, it may not be valid to compare I CTZ with the rest of South Vietnam. Hence, the present inquiry includes I CTZ only. The time period of nineteen months, July, 1966-January, 1968, stems from the data available in the Statistical Tables, OASD/SA at the time the present study started, June, 1968. The end-date of January, 1968 makes sense because of the 1968 TET offensive, which issued in a new period of combat activity. A replication of the present study might consider continuing the analysis beginning with January, 1968.

Another design decision consists of the modelling technique to use. The study employs path analysis. In recursive models, path analysis relies on standardized regression coefficients (beta-weights) to assess the relative value of previously hypothesized causal links. A recursive model is one that allows only one-way flows of causation. Casual modelling in general does not contain a scheme for assigning causality; the modelling, however, does offer a concise method for analyzing previously assumed causal links.

In this regard, one first postulates a model as a series of causal arrows linking a set of variables. At this stage, the presence of a causal arrow between two variables represents the explicit hypothesis that a positive or negative correlation exists between the two variables, after one controls for the effects of intervening variables. Likewise, the absence of a causal arrow between two variables represents the implicit assumption that the association between the two variables either washes out or remains zero when one takes into account the effects of the intervening variables. (Also, the ommission of a variable assumes that its effects are random).

Given these assumptions, one can subject causal models to two forms of evaluation. In the first place, one evaluates only the explicit links to see if they are non-zero as hypothesized. The second type of evaluation focuses on all possible arrows within the model to see if the arrows that are implicitly set equal to zero are,in fact, zero. In other words, there is a shift from allowing only the hypothesized arrows to affect the endogenous (dependent) variables to a model that allows all possible variables to affect the endogenous variables. This shift should not significantly alter the values determined in the first stage of testing <u>if</u> the implied links are approximately zero.

Thus, causal modelling offers the researcher two distinct advantages. First, it assumes a set of symbols and assumptions well suited to causal thinking, consequently, the researcher must specify his <u>a priori</u> theoretical assumptions in a concise manner. Second, the mathematics of causal modelling allow the researcher to make a relatively straightforward test of his assumptions.

So far the discussion treats the causal arrows as if they must be either zero or non-zero. This dichotomy however, is not the case. The present analysis assesses the <u>relative importance</u> of a given set of variables within a causal system. Fortunately, path analysis is ideally suited to this task.

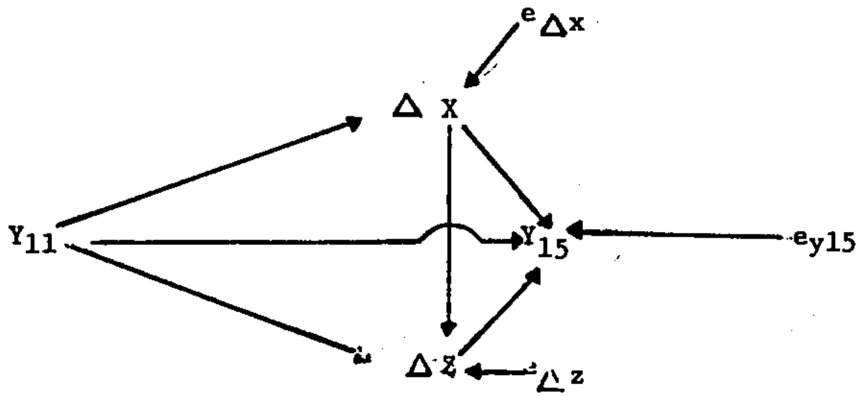
The absolute value of the path coefficients (beta-weights) is a direct indicator of the functional relationship between variables. In this context, one can ignore the absolute value of the path coefficient between two variables and focus on the value of this path relative to all other paths influencing the same dependent variable. In this manner, one simultaneously assesses the relative importance of two or more causal variables as they influence a given endogenous variable.

Regarding the recursiveness assumption, it may be unrealistic for the researcher to restrict his models to one-way causal flows. The nature of the data and the purpose of the present research suggests that the recursive models are not unrealistic for several reasons. First the data refer to monthly observations over a 19-month period. This time span allows the appropriate use of time-lags throughout the evaluations. The time-lags, moreover, make the recursive assumptions more legitimate. Second, the purpose of this research is to evaluate the relative impact of two particular causal variables on a third variable. It is possible to obtain a better fit with reality if other causal variables were introduced and the recursive restraints abandoned. The purpose of the present study, however, is to evaluate the relative importance of two causal variables rather than to construct a complete model of the war. The rather simplistic model serves this evaluation purpose well. A final reason for opting for recursive models is the relative ease with which recursive models may be evaluated. As previously mentioned in recursive models, path coefficients are estimated by standardized regression coefficients--beta weights. Many social science computer programs contain options that allow the direct computation of the beta-weights. And even with those programs that do not have this option, it is usually a simple matter to transform the variables into standard form (mean = zero; standard deviation = unity) and then compute the standardized regression coefficients.

One of the major advantages of standardized statistics, such as beta weights, is the dimensionless nature of the units. This is due to the fact that the value of a standardized variable is a measure of its deviation from its mean relative to the overall dispersion of the variable. Measuring variables in this manner removes the effects of different scaling techniques and strips the variable of its original dimension such as days of operation versus individual military actions. Consequently, the use of standardized variables makes it possible to compare changes in variables that were originally measured on totally different scales. The disadvantage inherent in standardized statistics, particularly the beta weight, is the loss of its predictive capability. An unstandardized regression coefficient can be used to estimate a line that describes the "best guess" as to later enemy activity, given alternative U.S. actions. Since the beta weight deals with standardized variables, it cannot be used in this fashion unless it is converted to its unstandardized form. Consequently the beta weight can best be interpreted in the relative sense. That is, the beta weight offers an excellent way in which to assess the relative importance of two or more causal variables-even when they are measured on different scales. Thus, if one desires to make forecasts based on the present analysis, he should use unstandardized regression coefficients, rather than the beta weights. These coefficients could be computed quite simply from the raw rather than the standardized data.

A final design area consists of the derivation of the structural equations and the specification of the models. The structural equations of a recursive causal model represent each endogenous variable in turn as a dependent variable and allow each of its incoming arrows to act as an independent predictor. Consider the following model and accompanying structural equations in this regard:

Model 1



1.) $\Delta X = P_{\Delta xy11} \; Y_{11} + e_{\Delta x}$

2.) $\Delta Z = P_{\Delta zy11} \; Y_{11} + P_{\Delta z \Delta x \Delta} X + e_{\Delta z}$

3.) $Y_{15} = P_{y15y11} \; Y_{11} + P_{y15 \Delta x \Delta} X + P_{y15 \Delta z \Delta} Z + e_{y15}$

Read the notation Pzx as the path value from X to Z. Note that there are three structural equations, or one equation for each endogenous variable. Now focus on the second equation because it contains two independent variables and the derivation of beta-weights is relatively straight forward as compared with equation number three. A multiplication of the second equation by either of its independent variables produces:

4.) $\Delta ZY_{11} = P_{zy11} \; Y_{11}^{2} + P_{\Delta z \Delta x \Delta} XY_{11} + Y_{11}e_{z}$ and

5.) $\Delta Z \Delta X = P_{zy11} \; Y_{11} \; \Delta X + P_{\Delta z \Delta x \Delta} X^{2} + \Delta X_{ez}$

By taking the expectation of both sides of equation 4, one obtains:

$$6.)\quad E(\Delta ZY_{11}) = E(P_{\Delta zy11} Y_{11}^2) + E(P_{\Delta z\ x}\Delta XY) + E(Y_{11}\ e_z)$$

$$r_{\Delta zy11} = P_{\Delta zy11}(1) + P_{\Delta z\Delta x}(r_{\Delta z\Delta x}) + 0$$

$$r_{\Delta zy11} = P_{\Delta zy11} + P_{\Delta z\Delta x}\ r_{\Delta z\Delta x}$$

Note that one assumes the expectation of $Y_{11}e\ z$ to be zero. This assumption, that the error term does not covary with the independent variables, is common to most regression-based analysis. Returning to equation 6, it has only two unknowns ($P_{\Delta zy11}$ and $P_{\Delta z\Delta x}$) since one can compute the r's or zero order correlations. Applying the expectation technique to equation 5 produces:

$$7.)\quad r_{\Delta z\Delta x} = P_{zy11}\ r_{y11\Delta x} + P_{\Delta z\Delta x}$$

Now equations 6 and 7 represent two independent equations, each with the same two unknowns. Solving this system produces:

$$8.)\quad P_{\Delta zy11} = \frac{r_{\Delta zy11} - r_{\Delta xy11}\ r_{\Delta x\Delta z}}{1 - r^2_{\Delta xy11}} \quad \text{and}$$

$$9.)\quad P_{\Delta z\Delta x} = \frac{r_{\Delta z\Delta x} - r_{\Delta xy11}\ r_{\Delta zy11}}{1 - r^2_{\Delta xy11}}$$

Since equation 8 is equivalent to an equation for the standardized regression coefficient of ΔZ on Y_{11} controlling for ΔX, one shows that the standardized regression coefficient is a legitimate estimator of path coefficients in a recursive model.[7]

---

[7]For a discussion of causal modelling in general and path analysis in particular see: Hayward Alker, "Causal Inference and Political Analysis," in C.L. Taylor (ed.) Aggregate Data Analysis. New Haven: Yale University Press, 1968, pp. 209-242; H.M. Blalock, Causal Inference in Non-experimental Research. Chapel Hill: University of North Carolina Press, 1964; D.R. Heise, "Problems in Path Analysis in E.F. Borgotta (ed.) Sociological Methodology. San Francisco: Jossey Bass, 1969, pp. 38-73; K.C. Land, "Principles of Path Analysis," in E.F. Borgotta (ed.). op. cit., pp. 3-37, R. Boudon, "A Method of Linear Causal Analysis: Dependence Analysis," American Sociological Review, XXX, No. 3 (June, 1965), pp. 365-373. R. Tanter, "Toward a Theory of Political Development," Midwest Journal of Political Science, XI, No. 2 (May, 1967), pp. 145-172; M. Midlarsky, R. Tanter, "Toward a Theory of Political Instability in Latin America," Journal of Peace Research (September, 1967), pp. 202-227.

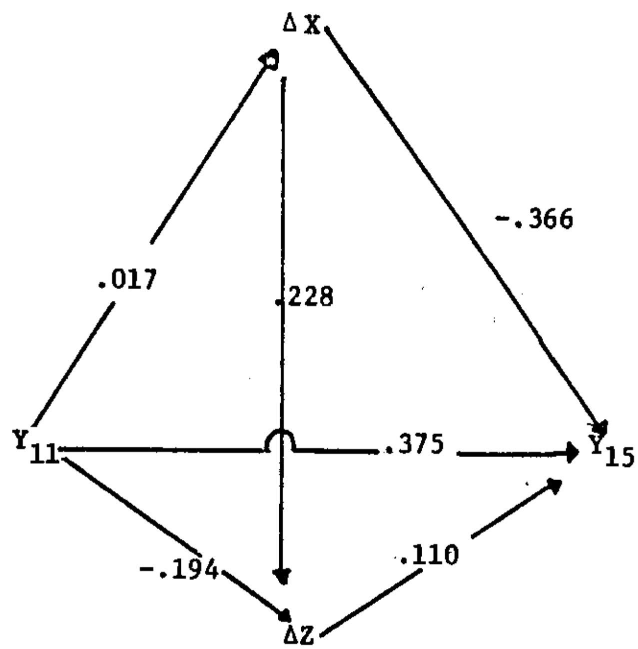
RESULTS

Recall that the main purpose of this research is to evaluate the relative impact of U.S. small- versus large-unit actions vis à vis VC/NVA attacks. In keeping with this purpose, the actual analysis of the general model involves a comparison of path coefficients across models 1-3. below.

Recall the coefficients in Table 2. Preliminary analysis reveals that U.S. small unit-actions and battalion days of operation covary positively, possibly due to their concurrent changes over time. Consequently, Model 1 successfully employs the first differences of both X and Z. The first differences decrease the compounding influence of time. Model 1 suggests that the rate of change in small-unit actions is over three times as important as the change in battalion days in influencing VC/NVA attacks. Moreover, note the signs of the coefficients, negative for small unit actions and positive for battalion days.

The purpose of Model 2 is to see if the confounding influences of time disappear as a result of first differences (ΔX and ΔZ). Note that in Model 2, time still covaries positively with VC/NVA attacks. In Model 2, there are no arrows between time and the X and Z terms. The implicit assumption here is that there is no connection between time and the first differences, e.g., time rates of change in small-unit actions and battalion days of operation respectively. Model 2 suggests that the rate of change in small-unit actions is about four and a half times as important as change in battalion days in influencing VC/NVA attacks. Indeed the coefficient for battalion days is zero. Again, note the negative sign for small-unit action changes.

Figure 2: Model 1



Y_{11} = V.C. attacks @ t - 4

Δ X = (X @ t - 2) - (X @ t - 3)

or $X_3 - X_2$

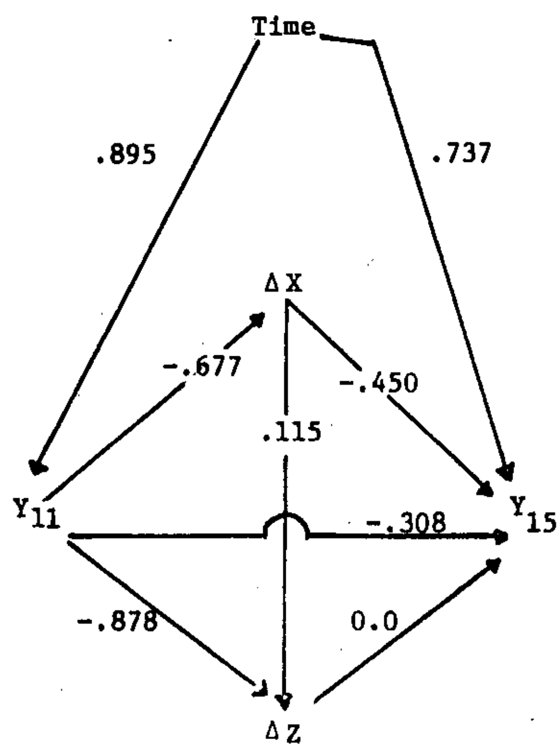
Δ Z = (Z @ t - 2) - (Z @ t - 3)

or $Z_8 - Z_7$

Y15 = ~~V.C.~~ VC/NVA attacks @ t - 0

N = 15 months

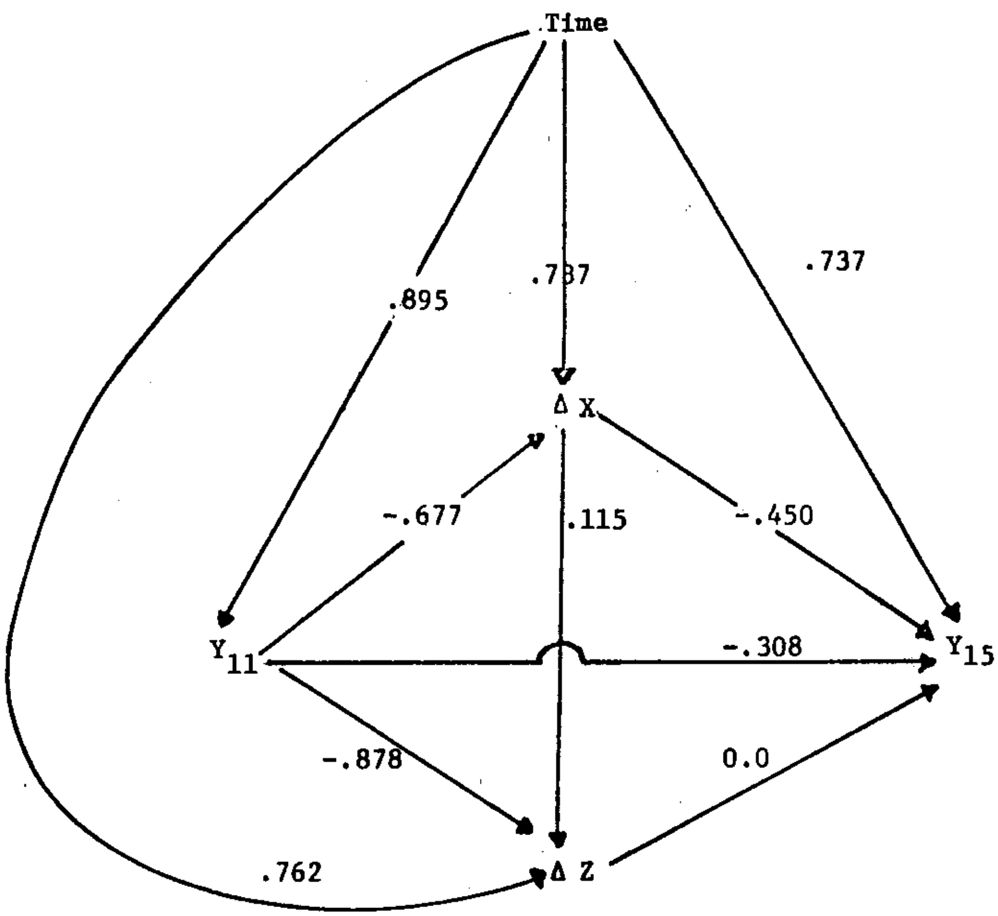
Figure 3: Model 2



Y_{11}, ΔX, ΔZ, and Y_{15} same as Model 1

Time = time in months; N ≥ 15 ≤ 19 months

Figure ': Model 3

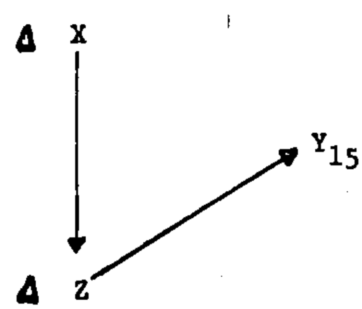


All variables same as Model 1; N ≥ 15 ≤ 19 months

Model 3 shows that the first differences of both X and Z still covary with time - and thus with each other. A possible explanation of this is that both X and Z change at a changing rate. Allowing time to enter into Model 3 does not change the tentative finding about the systematic negative effect of changes in small unit actions versus the random effect of changes in battalion days.

Perhaps it should be reemphasized at this point that Models 1-3 are not meant to be complete simulations of reality--instead these models are convenient tools for evaluating the relative impact of small-unit actions and battalion days on VC/NVA attacks. In this regard, it is interesting to recall that the absolute values of the paths connecting small-unit actions (X) with VC/NVA attacks (Y_{15}) are consistently higher than the corresponding values for battalion days of operation (Z). Also the rather high negative path coefficients linking the X term to Y_{15} (Models 1, 2, 3) suggest that an increase in Δ X covaries with a decrease in Y_{15}. It would seem that Δ Z has little or no impact on Y_{15} - particularly in Models 2 and 3 where the path value is approximately zero.

In summary, every model examined in this analysis suggests that changes in small-unit actions (X) have a stronger influence on VC/NVA attacks (Y_{15}) than does change in battalion days (Z). Furthermore, these conclusions are independent of the effects of X on Y_{15} that mediate through Z, that is, the path:



If this path were taken into account it would reduce further the causal effect produced by Z.

What are the implications of these tentative results for theory and policy? These initial findings seem to support the ideas of theorists such as Thompson. The U.S. small-unit actions <u>may</u> bring about temporary increases in the level of VC/NVA attacks but <u>perhaps</u> decreases in attacks later. The battalion operations appear to have no systematic effect on VC/NVA activity. Recall from Table 2 that there are no negative correlations among the variables because all are increasing over time. The calculation of first differences in small-unit actions and battalion days of operation results in a negative path coefficient between rate of change in small-unit actions and VC/NVA attacks. That is, the greater the rate of change in small-unit actions, the lower the level of VC/NVA attacks later. The preceding statement controls for prior levels of VC/NVA attacks as well as the rate of change in battalion days of operation. In short, the tentative finding of virtually no effect from battalion days of operation accords nicely with the theoretical expectations of this study.

As regards policy, the results suggest that the small-unit actions should increase relative to battalion days of operation. As Thompson states in <u>No Exit From Vietnam</u>, "It should be possible to reduce costs by cutting down all those activities which produce a negligible return for the effort expended". "With a reduction in search and destroy operations and with the enemy being forced to come out and fight on grounds not of his own choosing, the casualty rate should be significantly reduced though there is bound to be an occassional bad week" (p. 199).

The implications of the results and Thompson's theoretical expectations seem to be validated by the actual events in South Vietnam. With the change in the leadership of the U.S. Military Assistance Command Vietnam (MACV) from General William Westmoreland to General Creighton Abrams, there was a gradual shift from large search and destroy operations to a strategy using more small-unit actions. K.P. Buckley writes in the New York Times that before General Abrams (May, 1967), "Tactics were based on finding enemy forces, preferably in large groups, and fighting them. The kill count was all-important." Furthermore, U.S. battalions virtually joined arms with each other and swept through large areas, often finding nothing, or just a few enemy soldiers who could not get out of the way. The target was always the enemy soldier and not so much the political system which sustained him. Abrams broke down the patrolling units into companies of about 100 men or even squads of as few as six men. Buckley, moreover, concludes that Abram's tactics keep casualties at a minimum (pp. 124-125).

In addition to the change in U.S. strategy from one emphasizing battalions to a strategy based on small-unit actions, there may be a similar change in North Vietnam strategy. An article by Ted Szulc in the New York Times of December 27, 1969 suggests such a change in North Vietnamese strategy. "General Vo Nguyen Giap, North Vietnam's Defense Minister, has declared that Hanoi is concentrating on the development of 'high quality' and highly mobile strike units--instead of massed forces as in the past--to solve major strategic problems in the present stage of the Vietnam war." "The North Vietnamese have largely avoided big-scale engagements with United States and South Vietnamese forces during 1969, ..." It may be the case that both sides finally see the increased effectiveness that comes with a strategy based on small- versus large-unit operations.

By relying more on small-unit actions, Hanoi needs fewer men to fight in the South. Since the infiltration rate has been an indicator of Hanoi's battlefield intentions, the decreased rate may be interpreted in Washington as a conciliatory gesture. Indeed, Secretary of State William Rogers provided a positive interpretation of the decrease when he announced on December 23, 1969 that infiltration had tapered off to 60% of what it was a year before. Just as Hanoi needs fewer men to fight a "protracted war" with small units, the U.S. would need fewer men as well.

The Nixon Administration's policy of Vietnamization has a better chance of working if small- rather than large-unit activity characterizes U.S. operations. That is, U.S. forces can withdraw at an increasing rate consistent with the faster reduction in large operations in favor of small-unit actions. Such a deescalation of U.S. military activity could be interpreted as a conciliatory move by Hanoi, who may respond accordingly. Recall the theoretical expectations of Milstein and Mitchell in this regard. They forecast that a reduction of large-unit operations should result in fewer killings and abductions of South Vietnamese civilians by VC/NVA troops.

Qualifications and Summary

Recall the discussion of data quality control. The assumption here is that random error characterizes the dependent variable, VC/NVA attacks, and that the independent variables, U.S. small-unit actions and battalion days of operation, are relatively error-free. The fact that lagged values of the dependent variable, VC/NVA attacks, act as an independent variable does not invalidate the preceding assumption. The purpose here is to evaluate the strength of the two U.S. military action variables *vis à vis* VC/NVA attacks, *not* to evaluate the autocorrelation of VC/NVA attacks over time. Thus, the error assumptions appear be valid given the limited purposes of this inquiry.