MTR-6617

HIGH-IMPACT ANTI-CRIME PROGRAM

A METHODOLOGY FOR CONDUCTING A POLICE HYPOTHESIS TEST

U. S. DEPARTMENT OF JUSTICE
Law Enforcement Assistance Administration
National Institute of Law Enforcement and Criminal Justice

THE MITRE CORPORATION

MTR-6617

NATIONAL IMPACT PROGRAM EVALUATION

A METHODOLOGY FOR CONDUCTING A POLICE HYPOTHESIS TEST

by

J. A. SOISSON
and
M. BROWN

The MITRE Corporation

MARCH 1974

U.S. DEPARTMENT OF JUSTICE
LAW ENFORCEMENT ASSISTANCE ADMINISTRATION
NATIONAL INSTITUTE OF LAW ENFORCEMENT AND CRIMINAL JUSTICE

This document has been prepared by The MITRE Corporation, Washington Operations, under Contract F19628-73-C-0001 for the Law Enforcement Assistance Administration.

MITRE Department
and Project Approval: L L Holmes

ABSTRACT

This document presents a plan for conducting research to evaluate the effectiveness of highly visible police presence as a deterrent to crime. This police hypothesis test is being undertaken by the National Institute of Law Enforcement and Criminal Justice and The MITRE Corporation as part of a nation-wide evaluation of the High Impact Anti-Crime Program now in operation in eight cities across the country.

The document discusses the research issues involved in the test, the approach to be taken in evaluating the key variables, measurement alternatives and the analysis strategy.

FOREWORD

J. A. Soisson maintained overall responsibility for development of the plan to conduct the proposed research described in this document. This responsibility included: definition of the research issues; coordination of project information and liaison with appropriate law enforcement authorities; general approach to the research; and proposed application of model-generated information. M. Brown was responsible for the formulation and analysis of the proposed statistical models.

PREFACE

The High Impact Anti-Crime Program was designed by the Law Enforcement Assistance Administration (LEAA) to demonstrate in eight large cities the effectiveness of comprehensive, crime specific programs in reducing stranger-to-stranger crime and burglary.

The National Institute and The MITRE Corporation are engaged in an effort to conduct a National Level Evaluation (NLE) of the High Impact Anti-Crime Program. The NLE provides for the examination of a range of program processes and effects, both intra-city and inter-city in the areas of program planning, project selection and implementation, and evaluation activities. In addition, the NLE includes the examination of several underlying assumptions on which a number of anti-crime efforts, both in the Impact Program and nationwide, have been based. Hypothesis tests in the areas of the police, corrections, and the courts have been designed and will be conducted as part of the NLE. These hypothesis tests focus on a number of basic issues of interest to the criminal justice community. It is hoped that the findings of this research will be useful to criminal justice agencies in producing better designed and more effective anti-crime activities and that questions raised by this research can form the basis for further applied research in the area of criminal justice administration.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

1.0 INTRODUCTION

1.1 Hypothesis Testing in Criminal Justice Research

The rapid increase in reported rates of urban crime since the middle '60s has made policy makers, urban residents and members of the criminal justice community alike, increasingly aware of the need for effective anti-crime programs in our urban areas.

More money is now being spent on anti-crime activities than ever before. And more interest is being paid to the outcomes of the programs being funded in this area of increasing concern.

With the advent of the High Impact Anti-Crime program, large scale evaluation became a reality. Evaluation guidelines* at the federal level were designed to assist agencies in better utilizing limited resources. Project evaluation has allowed criminal justice planners to assess relative project effectiveness and to determine which anti-crime tactics bring the greatest impact in their local areas. Evaluation design has already produced a clearer understanding of what different criminal justice agencies are attempting to accomplish; evaluation results should provide new information on which strategies can be most effective in reaching these goals.

Hand-in-hand with evaluation has come the introduction of crime-oriented planning within the Impact program. Crime-oriented planning is an approach to the administration of criminal justice which focuses

---

*1) National Institute's Planning Guidelines and Programs to Reduce Crime.
2) National Institute Memorandum, Information Needs and Impact Funds, dated 24 August 1972.
3) National Institute Memorandum, Guidelines for Regional Office Review of Evaluation Components of Impact City Project Proposals, dated 23 February 1973.
4) National Institute/MITRE Document, Evaluation in Criminal Justice Programs: Guidelines and Examples, dated May 1973.

on the specific crime problems of a local area. Basic data analysis on levels of criminal activity, geographical areas of greatest concern and characteristics of the crime victim and offender (crime-specific analysis) provides the guidelines for an individual area's approach to crime reduction. Within these area-specific guidelines, problems can be identified and prioritized, broad strategies chosen and specific tactics selected to carry out the strategies.

Evaluation results not only assist planners and operators in judging the effectiveness of selected tactics or strategies; they also provide feedback to the data analysis stage of the crime-oriented planning process and influence future choices of tactical or strategic approaches. Further, the process of evaluation itself brings program assumptions into question. As in other fields of research or social action, the more criminal justice or law enforcement activities are examined for their true impact on crime, the more the need is felt to pinpoint the underlying principles which drive the choice of these activities, since it is the soundness of such assumptions which will determine a project's capacity to generate the desired results.

Experimental tests of these underlying assumptions are often thought to be the only option for assessing their validity. It is true that experiments or quasi-experiments do offer the best opportunities for clear and definitive results; they are not, however, the only investigative route open. Individual project evaluations are in fact a first step in the right direction.

The difference between an assumption or hypothesis test and an individual project evaluation is one of focus rather than kind. While individual project evaluations focus on finding out what happened in the project target area or target group in an immediate way, hypothesis tests generally seek to find out what changes in the underlying process have occurred, whether those changes are area-specific or whether they apply across the board.

Not all projects are amenable to hypothesis testing. What is essentially needed to conduct a project level hypothesis test is a project with clear objectives dependent on a unitary coherent strategy. There are a number of projects presently being proposed, planned, and operated in the Impact program which fall into this category. Projects amenable to hypothesis testing are thus being designed and funded through the traditional criminal justice system which not only serve the direct needs of the recipient community but which can fulfill research needs as well.

Much of the data necessary for conducting a hypothesis test is normally collected as part of a project evaluation. The police hypothesis test (PHT), the topic of this document, utilizes data provided in conjunction with project evaluations for patrol projects implemented as part of the Impact program in a number of the Impact cities. The test was designed with data availability for the sample projects well in mind. Working within these constraints, it was found that although restrictions were unquestionably placed on the type of methodology which could be implemented, it was nevertheless possible to design a workable methodology which will produce the type of information necessary to examine the hypothesis. This methodology is one example of the type of research which large scale project evaluation can make possible. Research of the type outlined in this document can be conducted at the local, state, and regional levels as well as at the national level, utilizing the data resources developed through area-wide crime-specific analysis and program/project evaluation. Conducting hypothesis tests at a national level allows the researcher to look across cities at similar types of projects using a single test strategy. This, however, is also possible at the regional and state levels, and local areas could accomplish a similar end by working in cooperation with one another in the design and conduct of a hypothesis test. Given the data resources now being developed across the nation as part of the on-going criminal justice process, the outlook is good for future evaluation research of this kind within the framework of project level evaluation.

1.2 Selection of Hypothesis

Police have long been considered a crime-deterrent factor by the community. The mere presence of a police officer is believed by many to be sufficient to deter crime in the area where he is located. The American Bar Association, for example, in Standards for Criminal Justice Relating to the Urban Police Function, listed, in order of priority, eleven major police responsibilities; the second one listed was "to reduce the opportunities for the commission of some crimes through preventive patrol and other measures."

Preventive patrol (i.e., the deterrence strategy most often used by police officers) is steeped in tradition and there is little evidence that there is a trend away from it. Although the true deterrent value of police has been questioned, the National Advisory Commission on Criminal Justice Goals and Standards stated, as recently as January, 1973, that "unless conclusive data is obtained establishing that preventive patrol is not the best utilization of patrol resources in controlling crime, the practice should be continued" and that every police executive should set forth written policies on patrol services, with an emphasis on "the need for preventive patrol to reduce the opportunity for criminal activity."

What follows from the commonly-held belief that police presence deters crime is the corollary assumption that a change in police presence, or in the nature of that presence, will affect crime, that an increase in the visible presence of police manpower in a certain area, for example, will cause a decrease in crime levels in that area. Yet, it has not been definitively shown that visible police presence does indeed have a meaningful relationship to crime levels, and if so,

under what circumstances and within what limits. Consequently, the Impact program effort has been viewed as an opportunity to augment existing knowledge in this area, and the following hypothesis was selected for testing:

> "An increase in the visible presence of police in a given area will result in a decrease in crime rates in that area."

1.3 Purpose of this Document

The purpose of this document is to present the approach to be taken in this examination of the police hypothesis. A previous paper, "A Plan for Conducting a Police Hypothesis Test--Task 4 of the National Level Evaluation," (WP-10296), outlined the possible approaches to conducting a PHT and the range of factors which may be relevant to such a test. This document brings together the relevant issues previously discussed and presents a methodology for conducting this research.

The document includes discussion of the conceptual framework on which the research is based and the general research issues involved in the hypothesis. The research questions to be addressed in this test are outlined in detail as well as the methods to be used in measuring the factors involved and in analyzing the data collected. The document is intended to give the reader a clear picture of the research package which has been designed to test the police hypothesis.

2.0 RESEARCH FRAMEWORK

2.1 Overall Plan for Conducting the Police Hypothesis Test

The test of the police hypothesis being conducted by MITRE/ National Institute focuses on the use of projects implemented in the Impact program which have been designed to reduce crime in target areas through increasing visible police presence in those areas. Data to be used in conducting the test is to be provided in conjunction with the evaluations of the individual projects.

Projects will be utilized in the study which fit the operational needs of conducting the PHT. These sample cases will be selected from a large number of Impact projects which are based at least in part on the police visibility assumption. Possible candidate projects are briefly described in Appendix I.

The projects selected all involve an increase in visible police presence (VPP) as defined:

Visible Police Presence

Sworn police officers, including auxiliaries and cadets in full uniform. (Excludes private police, metermaids, tenant patrols and block watchers.)

Operating in an overt mode (excludes undercover, disguises, stakeout, etc.).

On foot or in a vehicle, but only if a sworn officer is present in the vehicle (excluding helicopters).

Plainclothesmen and unmarked sedans are included if they are engaged in overt operations. (It is thought that unmarked police sedans and plainclothesmen in overt operations are generally recognized by most citizens, especially in high crime areas; hence, their deterrent effect will be included in this study.)

In each project to be utilized as a sample case, the increase in VPP is in a clearly defined target area and is continued in that target area for a substantial length of time (at least three months, preferably six months of more). Results in terms of the outcome variable, crime levels, will then be monitored and analyzed.

What the study is essentially planning to do then is to address selected issues central to the police hypothesis using sample police patrol projects whose success depends directly on the police visibility assumption. A single analysis strategy will be used across all cases which addresses the changes in the crime levels affected by the introduction of increases in VPP.

Each research question will be dealt with as an individual issue within this unitary strategy. As will be discussed in greater detail in later sections, the results found in the sample cases will then be synthesized to provide a general picture of possible changes in crime levels which are apparent after levels of police visibility have been increased.

2.2 Conceptual Framework

A large number of political, demographic and socio-economic forces operate together in a community to create its basic underlying social environment. Certain social environments tend to be more associated with crime problems than others. Typically, high crime areas are characterized by populations possessing low socio-economic and educational status, by poor housing conditions and by social disorganization. There is little conclusive information, however, about the dynamics of these factors in creating specific crime climates.

Criminal justice activities including police activities are superimposed upon different crime climates with the expectation of reducing crime. The present research is investigating the effects of police visibility across varying crime climates.

As the hypothesis states, it is expected that the introduction of increased VPP into a community will trigger a change in the crime climate and crime will be reduced.

Figure 1 (see page 9) shows those factors which may be directly involved in the impact of VPP on the crime climate of a community. In general this schematic lays out a number of factors which could play a role or exhibit changes which could be realized with the introduction of VPP into an area.

There are a number of factors which characterize the nature of police visibility which may have an effect on the impact of VPP in the crime environment of a target community. These factors include: the level of manpower and type, the type of deployment, the modes of patrol utilized and the attitudes of the participating patrolmen.

There are also a number of factors outside the realm of direct police activity which may affect the impact of VPP in the area. These include these community characteristics which contribute to the crime environment of the area such as the physical environment of the target area, the socio-economic make-up of the target community and the attitudes of the target community towards the police or the patrol project. In addition, other criminal justice projects in the area may have an effect on the area where VPP has an impact and on the degree to which the impact of VPP is felt.

Finally, as Figure 1 shows, there are a number of areas where the impact of VPP on the crime picture in the target area may be realized. VPP may affect a number of things: the level of crime, the amount of

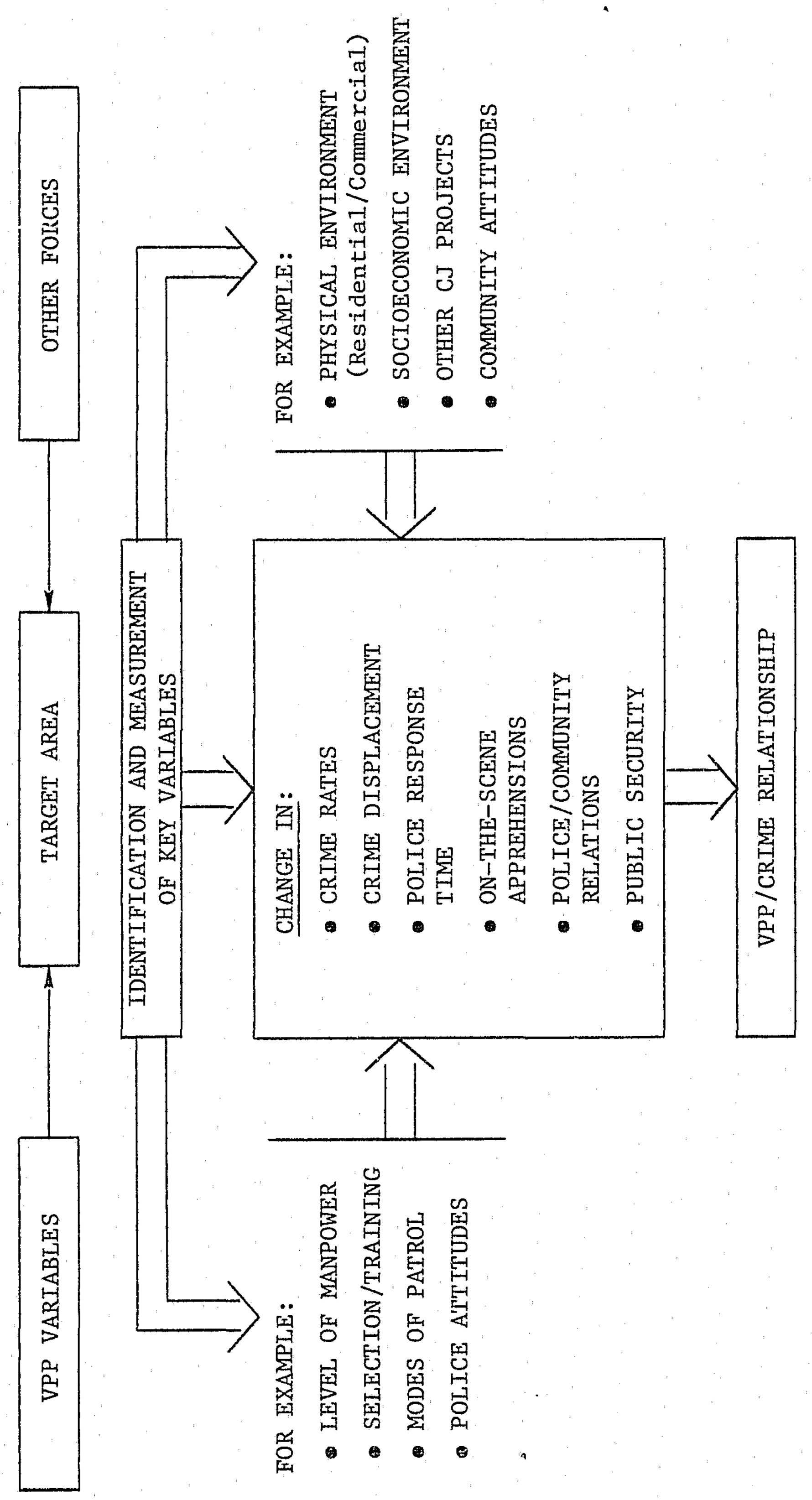


FIGURE 1

CONCEPTUAL FRAMEWORK FOR EXAMINING HYPOTHESIS crime displacement, certain types of police performance such as the police response time and the rate of on-the-scene apprehensions. It may affect police or community attitudes or the community's perception of crime.

2.3 Operational Model

As the hypothesis states, it is expected that there is an inverse relationship between VPP and crime, which is to say that an increase in visible police presence will mean a decrease in crime. Taking this as the focus, the relevant factors already discussed in the description of the conceptual approach to examining the hypothesis can be reorganized into an operational model for testing the central research issue in the test: the relationship between VPP and crime. Figure 2 displays this operational model.

The independent variable in the hypothesis test is visible police presence; the central dependent variable is the level of crime, (i.e., the crime rates) given the hypothesis that high VPP will result in lower crime rates.

The other outcome variables indicated in the methodology (see Figure 1) serve an explanatory function in the test. Changes in police response time, rates of on-the-scene apprehension, etc., which may accompany changes in crime rates, in essence can help to explain the effect of VPP. For instance, if in a number of cases, increases in VPP with resulting decreases in crime rates are accompanied by decreases in police response time, it might be assumed that the crime rate decreases could be linked to the intervening effect on response time.

Similarly, the factors grouped in Figure 1 as "other forces," (physical make-up of the target area, socio-economic character of population, etc.) can also be viewed as possible explanatory variables.

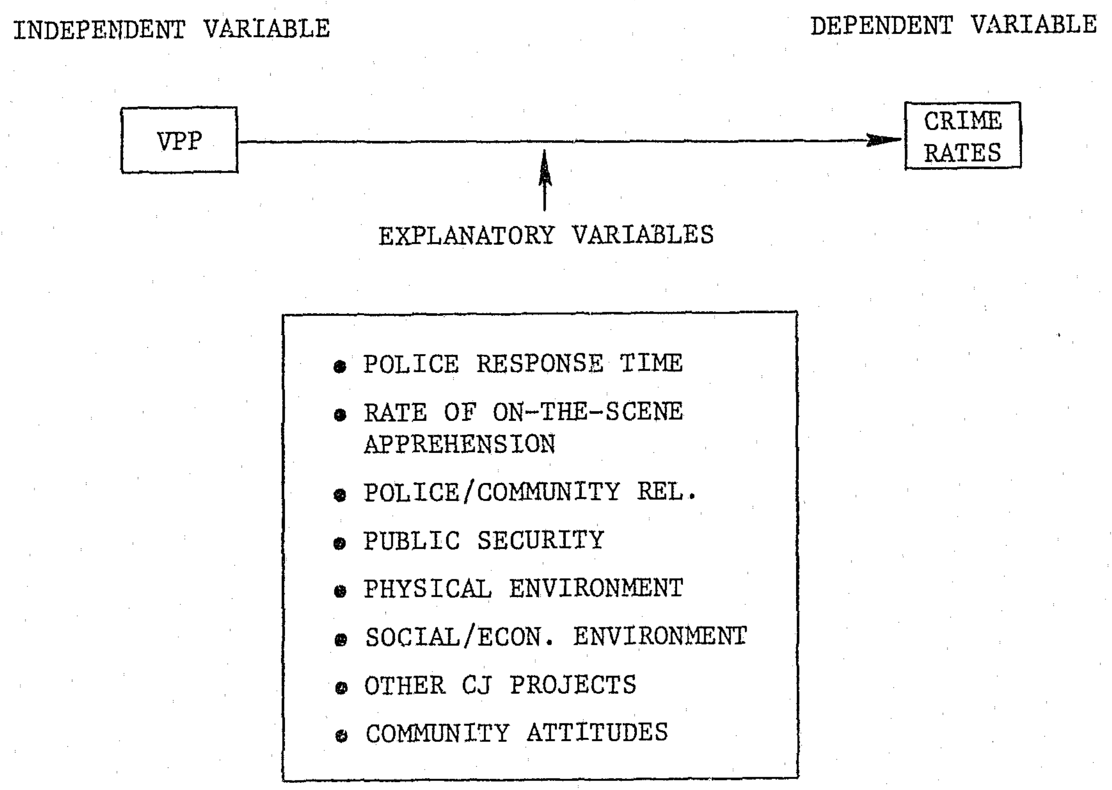


FIGURE 2

OPERATIONAL MODEL FOR CONDUCTING A POLICE HYPOTHESIS TEST

These variables characterize contextual factors which may influence the central relationship (VPP and crime rates) under examination.

This operational model then frames the approach of this study. The central research issue, the relationship between VPP and crime, and other research issues involved in this relationship will be discussed in greater detail below.

3.0 RESEARCH ISSUES AND QUESTIONS

Based on the operational model presented in the previous section, there are a number of research issues which are related to a test of the police hypothesis. In the section below (3.1), the range of research questions suggested by the model will be presented. Each question involves certain key factors which need to be measured in order for the question to be adequately addressed. The possible ways of measuring these factors will be outlined as well as the reliability and validity of alternative possibilities of measurement. Finally, the feasibility of addressing each question in this test of the police hypothesis will be discussed.

Those questions which are deemed researchable in the context of this test of the police hypothesis as based on the above discussion, will then be laid out in the following section (3.2) and will define the scope of the police hypothesis test.

3.1 Broad Research Issues

The central research question in a police hypothesis test asks what is the relationship between visible police presence and the occurrence of crime. If the hypothesis is true, then the level of crime should decline when VPP is increased.

The first major factor involved in this research question is the level of crime. Crime is generally measured in one of two ways: (1) from point of view of the police (reported crime rates), or (2) from point of view of the victim (victimization crime rates). Each type of measurement has both advantages and disadvantages.

Reported crime rates have the main advantage of being readily available. As well, they are the basis for most police operations, planning and policy making decisions. The major disadvantage is that reported crime rates measure only one portion of the crime which actually takes place.

Victimization crime rates on the other hand are believed to be a better estimate of crime. There are, however, a number of problems associated with their use. Data collection is time consuming and expensive. Because of survey duration, victimization data can be made available for only a small number of points in time surrounding and during a test period. Thus, the use of extensive historical or baseline data for the time periods preceding a project is precluded if analysis is to utilize victimization crime estimates.

Due to the availability of police data (i.e., reported crime figures), the expense of victimization data and the resource constraints of this research project, reported crime data will have to be used.

As will be discussed in greater detail below, use of reported crime rates alone will leave questions unanswered as to the effect of VPP on actual (as opposed to reported) crime rates. An increase in VPP may initiate different changes in the reporting of crimes versus the commission of crimes. Thus, monitoring only the reported crime levels may give a skewed picture of the effect of VPP on actual levels of crime. Hence, given this constraint, what will be tested in this study is the effect of VPP on reported crime rates.

The second major factor involved in this primary research question concerning the relationship between VPP and crime, is that of the police visibility. VPP can be measured from two perspectives: (1) from the point of view of the police and/or (2) from the community's point of view.

From the police point of view, measurement of VPP involves the number of men allocated by the police department in a given area as well as the deployment pattern and the mode of patrol.

Unfortunately, this information is available in our sample areas only for the manpower added by the projects. In most areas, no record is kept on a timely or area specific basis of the regular police officers operating in the target area. Similarly, little if any historical data is available on the levels of visible police presence in the target areas before the project was implemented.

What can and will be done is this: the visible police manpower introduced by a project can be measured and this gives us an indicator of the increase in manpower which has occurred. For each sample area VPP profiles will be constructed which include: the number of patrolmen, the type of deployment, the mode of patrol, the size of the area patrolled and the length of patrol time. The profiles will be utilized in comparing the results in the crime level analysis. (This will be discussed in Section 4.3.)

Approaching the question of measurement of VPP from the community's point of view involves measurement of the level of VPP as experienced by the population of the area. People's awareness of police presence may vary from area to area and it has been suggested that it is in fact the perception of police presence that is the key issue in VPP deterrence. The methods available for assessing the community awareness of VPP involve surveying either people on the streets or the resident population and/or placing observers in the target area to record the number of patrolmen that are seen in the area during certain time periods. Because of resource constraints, no surveys will be conducted as part of this research project. It is hoped that for some projects, observers can be placed in target areas to monitor the level of VPP from the public standpoint.

It should be noted that VPP is directed toward the deterrence of criminals or potential criminals. There is no assurance that this target group is identical to the members of the local community or that the perception of VPP of these two groups is the same. Thus, there may be no direct relationship between the public perception of VPP and the levels of crime in the area. It is the criminal's reaction or change in action resulting from his perception in terms of changes on the level of crime in which we are most interested in this research endeavor. Clearly, there is no viable way of judging or assessing the perception of the criminal of different levels of VPP except by monitoring the outcomes in terms of crime levels.

Within this central question involving the relationship between VPP and crime there are a number of subsidiary questions.

Which specific crimes are most affected by VPP? It is plausible to expect that certain crimes may be more susceptible to a deterrent effect than others both because of the nature of visibility and the nature of the specific crime.

Specific crime rates are the most direct measure of this, of course. The choices for data sources are again victimization and reported data. In this study, for the reasons cited above, the reported crime rates will be used.

Another question subsidiary to the central issue in the police hypothesis test is that of the differential effect of VPP on outdoor crimes. It has been suggested, that outdoor* (crimes committed out-of-doors, within possible sight of an officer on patrol) are more strongly affected by VPP than are other crimes. If this is true, it would be expected that the rate of outdoor crime occurrences would have a greater tendency to decline with an increase in VPP than would the rate of other crimes.

The location of a criminal offense (indoor, outdoor, unknown) is a data item recorded on offense reports by many police departments. Using location data, overall crime levels and specific crime levels (again relying on reported crime) can be partitioned into indoor and outdoor crimes. These additional crime categories can then be examined in light of increases in VPP to provide additional information on the differential deterrent effect of police visibility.

---

*Outdoor crimes are often termed "suppressible crimes."

Related to this central question, it has been postulated that VPP may affect police performance. Certain changes in police performance (on-the-scene apprehensions, police response time, identification of crime) may help to explain the effect of VPP on crime. The possible intervening effects of changes in police performance are discussed below.

a. <u>Increased Identification of Crime</u>

It has been suggested that increased police visibility and concomitantly increased police manpower on the streets may lead to a higher level of police detection of crimes. More police-detected crimes with higher manpower levels would mean a larger number of reported crimes with an increase in VPP.

Higher reported crime rates accompanying increases in VPP would then be an effect of greater police detection of crimes rather than an increase in the actual number of crimes being committed.

Theoretically, there are several ways to assess this effect. If reported crimes could be partitioned into crimes detected by the police and other crimes (either reported by citizens or unknown), the relative changes in these categories which accompany increases in VPP could be assessed. However, this type of information is rarely recorded by police departments on a regular basis or otherwise. Even if this data were available, there are a large number of cases in which a crime in progress is detected by the police but the warrant is filed by the victim. This is particularly frequent in assault cases. For these reasons, this approach to addressing this question is infeasible.

An alternate way of approaching this issue is to examine the relative changes in victimization crime rates and reported crime rates. Reported crime rates would reflect the inflated number of police detected crimes whereas victimization data would not (since unreported crimes are already included in this estimate). However, based on our decision to use only reported crime rates, this alternative must also be ruled out in this test of the hypothesis and this question will not be addressed.

b. Police Response Time

It has also been suggested that an increase in VPP may result in a decrease in police response time. In turn, rapid police response time is believed to act as a deterrent to crime. This, however, is still unproven.

Rapid response to calls for service, especially for crimes in progress, is believed to be essential to apprehension of the criminal suspect because of the increased likelihood of evidence and increased availability of witnesses. As well, rapid response to calls may increase the probability of on-the-scene apprehension. Thus, it would be interesting to look at changes in response time which accompany increases in VPP as a possible intervening factor which could help explain possible reductions in crime levels.

This question raises certain problems. First of all, there is very little reliable data available on police response time. This is understandable because in order to record the time involved in answering a call, valuable time must be allocated to data recording. Thus, by implementing data collection, police response time would be decreased. Hence, there is a great deal of hesitancy on the part of many police departments to collect this type of data.

Also, there is little conclusive information as to the actual deterrent effect of rapid response time. Several studies are being conducted or are planned which examine the relationship between police response time and crime, notably one by the police department in Kansas City. Without information on the link between police response time and crime, it is premature to try to utilize police response time as an explanatory factor in this police hypothesis test.

c. Rate of On-The-Scene Apprehensions

Again, an increase in visible manpower is thought to lead to an increase in the rate of on-the-scene apprehensions.

A high rate of on-the-scene apprehension may lead to increased conviction of criminal suspects due to improved evidence and the availability of witnesses. Thus, it is believed to have a deterrent effect on crime.

Again, as in the question of police response time, there is a problem of data availability. Few police departments collect this information on a regular basis.

Also, as in the case of police response time, the evidence on the relationship between the rate of on-the-scene apprehension and crime is not conclusive, making the measure less than useful as an explanatory factor. There are factors other than police activity involved in the rate of on-the-scene apprehension; in point of fact, only a small proportion of on-the-scene apprehensions result from the patrol officer detecting the crime in progress. Rather, they are a product of the police being available in close proximity at the time of a call for service for a crime in process. This process involves other factors such as the citizen cooperation in contacting the police in a timely fashion. This is not directly affected by an increase in VPP. In addition, if the mode of visible patrol utilized is foot patrol, it is likely that the average rate would decline, confounding the issue.

For the above reasons, this question of the role of on-the-scene apprehension will not be addressed in this test of the PHT.

It has also been postulated that the attitudes of both the police participating in the patrol activities and of the population in the patrolled area may have an impact on the effectiveness of visible patrol as a crime deterrent.

Police attitudes could affect the relationship between VPP and crime levels in a number of ways. Police officers could, through a good rapport with the community in which they are patrolling, increase community support for and cooperation with the police. This could mean a better response from citizens in terms of timely calls to police about crime problems either in progress or suspected. Police, familiar with the local area, may be better able to apprehend suspects. In addition, police with a positive attitude toward their work may perform better in their duties which could mean faster response times, more on-the-scene apprehensions and more thorough crime investigations. All these things could work together to maintain a police presence in the area which is more effective in deterring crime.

Similarly, community attitudes towards the police could have an effect on the impact of VPP as a crime deterrent. If having the police in the community engenders a greater level of confidence in the police force, citizens may lend increased cooperation to the patrolmen in terms of calls to police about potential problems.

Further, seeing more patrolmen in their community may make residents more aware of the crime problems in their area and this can lead them to initiate steps on the part of the population to take additional measures on their own to protect themsleves.

However, it may also be true that increased police presence in an area may give the community a greater sense of security and residents may take fewer anti-crime precautions based on this change in perception. If the crime level has not actually decreased, these residents may be inadvertently leaving themselves open to risks in ways that they would have not done otherwise.

It is also possible that a community may perceive police presence as an infringement on their personal space and react negatively to increases in patrolmen in their neighborhood. This could lead to decreases in public cooperation with the police and a decrease in the reporting of crimes by residents.

These are only hypotheses without conclusive evidence to support them. The only way to get the data necessary to address the questions raised in these suggestions is through opinion surveys both of the police and the target area populations. The major constraints of this study, manpower and resources, again preclude examination of these issues. These are issues which deserve attention and hopefully they will be addressed in a comprehensive manner in the near future. Some studies are in progress which are examining public attitudes, most notably the National Level Victimization Survey conducted by the Bureau of the Census for LEAA.

An issue which has been getting increased attention both in the High Impact Anti-Crime Program and in other anti-crime programs and research is the problem of crime displacement.

The question involved here is whether, if a decrease in crime is realized in a specified area or a specified time slot, crime has actually been deterred or if the crime has simply been shifted to a different geographical area or a different time slot. Similarly, if the level of certain crimes or types of crime has decreased, it is not clear whether a mere shift to different crime alternatives has occurred or whether a true deterrent effect has been realized.

There are several types of possible displacement effects which may be directly related to the police visibility question:

1. <u>Localized Geographical Displacement</u>

The transfer of criminal activity to an area immediately adjacent or peripheral to a specified target area.

2. Temporal Displacement

The transfer of criminal activity from one time slot in which VPP has been increased to time slots not receiving such treatment.

3. Inward Displacement

The transfer of on-street criminal activity to indoor locations potentially out of sight of a patrol officer on duty.

4. Crime-to-Crime Displacement

Shifting of criminal activities from the commission of specific crimes or types of crime to other crime types.

Measurement of the factors involved in this issue, the occurrence of various crimes in various geographical areas or time slots is fairly straightforward. The same alternatives are available and here again reliance will be placed on reported crime rates.

However, actually assessing the significance of shifts in crime levels among crime categories or time-spaces, poses some difficult problems for which there are no answers.

Very little is known about the dynamics of crime occurrences. If crime levels decline in one area and increase in an immediately adjacent area, it is difficult to determine if or to what extent a transfer has taken place. It is plausible that a change in the process has occurred within one or both of the areas which accounts for the variation in crime levels. Or it is possible that the differential changes are attributable to a change in the overall crime generation/distribution process in the city as a whole which could affect different areas of the city to varying degrees based on perhaps economic or socio-demographic characteristics of the neighborhoods.

It is clear that the question of displacement must be addressed but the above problems preclude analysis which can address the question directly.

In this study, the approach that will be taken is to monitor and assess changes in the possible "displaceable" time-spaces and crime categories. Based on the regularity of crime transfers in the sample cases which are presented by this assessment, certain inferences can be made as to the potential types of displacement resulting from VPP.

In addressing the issue of localized geographic displacement, several rings of "displaceable" areas will be examined for projects able to provide data in the necessary form. This will allow us to monitor the possibility of a first ring spread effect (i.e., the deterrent effect of increased VPP extending over the target area boundaries into the immediately adjacent areas) with displacement possibly occurring past that point into the areas peripheral to the first surrounding ring.

3.2 Research Questions to be Addressed

As is discussed above, there are a number of research issues which are related to the police hypothesis test. Based on feasibility constraints (data, time, manpower, financial resources) a group of the issues involved in a test of the police hypothesis have been selected for examination in this study.

The central question to be addressed is:

> Is an increase in visible police presence in an area accompanied by a decrease in the reported crime levels in that area?

Two questions subsidiary to the central question in the test will also be addressed:

> Do the reported crime levels of certain crimes or types of crime show a decrease while others do not? and,
>
> Are reported levels of outdoor crime affected more than those of indoor crime?

Several questions related to the central question will also be addressed:

Is a decrease in crime in a target area accompanied by an increase in crime in the areas immediately adjacent to that area?

Is a decrease in crime in a certain time slot accompanied by increases in crime in the surrounding time slots? and,

Is a decrease in outdoor crime accompanied by an increase in indoor crimes?

4.0 ANALYSIS STRATEGY

4.1 Analytical Approach

The above research questions will be addressed utilizing the following analysis strategy.

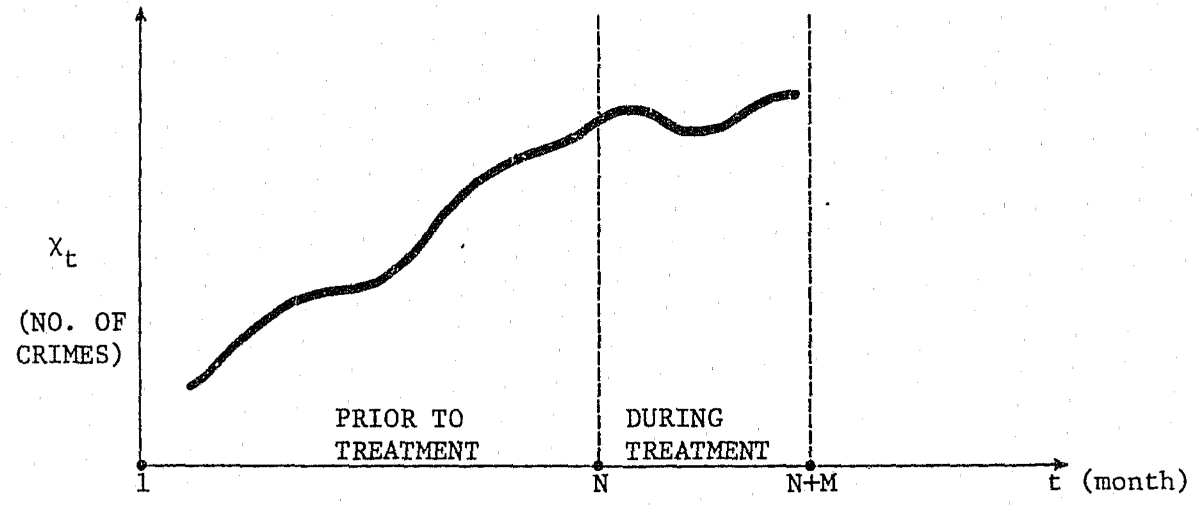
Each research question to be addressed involves an assessment of changes in levels of certain crimes or categories of crime in a certain time-space as related to changes in levels of police visibility. For example, to address the question of the impact of increases in police visibility on the levels of Impact crimes in a target area, the changes in levels of Impact crimes which correspond to the changes in the levels of visibility must be assessed.

A model has been developed which describes the levels of crime in a given time-space. Using historical or baseline data, this model can be utilized to describe the levels of crime occurring before increases in VPP were introduced. The same model can be used to describe the crime levels after the increase in VPP. These two descriptions can then be compared to assess both the direction of changes in crime levels which have occurred since VPP was increased and the nature of these changes.

The model is described in the following section (4.2) and the application of the model in addressing the research questions in the test is discussed in Section 4.3.

4.2 Analytical Model

For each space-time slot and each crime type, we can obtain data as to numbers of crimes committed (i.e., reported) each month. These will form a time series: $X_1, X_2, X_3, \ldots, X_N, X_{N+1}, \ldots, X_{N+M}$ where N is the number of data points prior to treatment and M is the number of data points during treatment.



Each such series is to be analyzed to determine the confidence it engenders in the hypothesis that the treatment has reduced the crime level to less than what it would have been in the absence of treatment.

To test the hypothesis, it is necessary to model the process that generates the χ_t. It seems plausible to assume that the data are generated as a sum of the following components:

1. A "reference" level of crime, denoted by "a", a constant.
2. A "long term trend", represented by "bt", where b is a constant.
3. A "annual cyclic component", represented by c sin $\left(\frac{\pi t}{6}\right)$ + d cos $\left(\frac{\pi t}{6}\right)$, where c and d are constants.
4. A purely random, or "noise" component, denoted by ε_t.

Thus, before treatment (i.e., t = 1, 2, ..., N),

$$\chi_t = a + bt + c \sin\left(\frac{\pi t}{6}\right) + d \cos\left(\frac{\pi t}{6}\right) + \varepsilon_t$$

It is assumed that the effect of increasing police visibility is to change the crime rate by some factor, denoted by Θ. Thus, during treatment (i.e., t = N + 1, N + 2, ..., N + M),

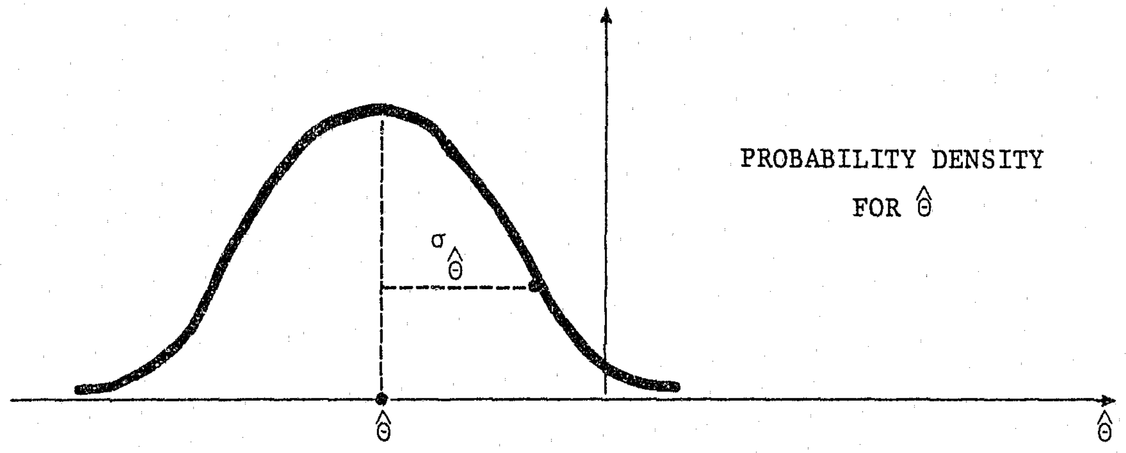
$$U_t = \left[a + b\tau + c \sin\left(\frac{\pi t}{6}\right) + d \cos\left(\frac{\pi t}{6}\right) + \eta_t\right] \Theta$$

where for notational convenience the t, χ_t and ε_t are denoted by τ, U_t and η_t respectively, thus distinguishing them from the pretreatment values.

The hypothesis, that the crime level has been reduced by the treatment to a level below what it would have been without treatment, is then mathematically equivalent to: $\Theta < 1$.

The time series data is to be analyzed to estimate the quantity Θ.

The estimate, $\hat{\Theta}$, will be a random variable (since it is computed from data), and will contain an uncertainty, which can be depicted as:



The uncertainty is measured by the standard deviation, $\sigma_{\hat{\Theta}}$, of the estimate $\hat{\Theta}$.

The area under the probability density curve, for $\hat{\Theta} < 0$, measures the percentage confidence that the data accords to the hypothesis, $\hat{\Theta} < 0$.

What is necessary, then, are formulas for computing $\hat{\Theta}$ and $\hat{\sigma}_{\hat{\Theta}}$ (the estimated value of $\sigma_{\hat{\Theta}}$) from the data.

The required formulas are derived in Appendix II, and can be most efficiently expressed in matrix notation. The results are as follows:

$$\hat{\Theta} = \frac{\underline{U}^T J \underline{\chi}}{\underline{\chi}^T K \underline{\chi}}$$

$$\hat{\sigma}^2_{\hat{\Theta}} = \frac{\underline{\chi}^T R \underline{\chi}}{\underline{1}^T \text{ diag } R} \left\{ \frac{1}{\underline{\chi}^T K \underline{\chi}} + \frac{\underline{\chi}^T P \underline{\chi}}{(\underline{\chi}^T K \underline{\chi})^2} \right\} \frac{(\underline{U}^T J \underline{\chi})^2}{(\underline{\chi}^T J \underline{\chi})^2}$$

where $\underline{\chi}$, $\underline{U}$ are vectors of crime data

superscript T indicates the transpose

superscript -1 indicates the inverse

$\underline{1}$ is a vector whose components are all ones

and
$$\begin{cases} J = H(G^TG)^{-1}G^T \\ K = G(G^TG)^{-1}H^TH(G^TG)^{-1}G^T \\ P = G(G^TG)^{-1}H^TH(G^TG)^{-1}H^TH(G^TG)^{-1}G^T \\ R = I-G(G^TG)^{-1}G^T \end{cases}$$

where

$$G = \begin{pmatrix} 1 & 1 & \sin \frac{\P}{6} & \cos \frac{\P}{6} \\ 1 & 2 & \sin \frac{2\P}{6} & \cos \frac{2\P}{6} \\ 1 & 3 & \sin \frac{3\P}{6} & \cos \frac{3\P}{6} \\ \cdots & & & \\ \cdots & & & \\ 1 & N & \sin \frac{N\P}{6} & \cos \frac{N\P}{6} \end{pmatrix}$$

$$H = \begin{pmatrix} 1 & N+1 & \sin \left[\frac{(N+1)\P}{6}\right] & \cos \left[\frac{(N+1)\P}{6}\right] \\ 1 & N+2 & \sin \left[\frac{(N+2)\P}{6}\right] & \cos \left[\frac{(N+2)\P}{6}\right] \\ \cdots & & & \\ \cdots & & & \\ 1 & N+M & \sin \left[\frac{(N+M)\P}{6}\right] & \cos \left[\frac{(N+M)\P}{6}\right] \end{pmatrix}$$

and
$$\begin{cases} I = \text{identity matrix} \\ \text{diag } R = \text{a vector whose components are the major diagonal} \\ \qquad \text{elements of the square matrix } R. \end{cases}$$

AN ALTERNATIVE MODEL

Another representation of the process by which the χ_t are generated is (for t = 1, 2, ..., N):

$$\chi_t = \sum_{i=1}^{12} a_i \nu_{ti} + bt + \varepsilon_t$$

where the ν_{ti} are 0-1 indicator variables that specify whether month t is January, February, etc. For example, if the data started in January, one would have:

$$\left.\begin{aligned} \nu_{11} &= 1 \\ \nu_{12} &= \nu_{13} = \ldots = \nu_{12} = 0 \end{aligned}\right\}$$
$$\left.\begin{aligned} \nu_{22} &= 2 \\ \nu_{21} &= \nu_{23} = \ldots = \nu_{2\ 12} = 0 \end{aligned}\right\}$$

...

...

$$\left.\begin{array}{l} \nu_{12\ 12} = 1 \\ \nu_{12\ 1} = \nu_{12\ 2} = \dots = \nu_{12\ 11} = 0 \end{array}\right\}$$

$$\left.\begin{array}{l} \nu_{13\ 1} = 1 \\ \nu_{13\ 2} = \nu_{13\ 3} = \dots = \nu_{13\ 12} = 0 \end{array}\right\}$$

etc.

The advantage of this representation is that the seasonal variations, while still repeating cyclically from one year to the next, are not restricted by assumption to be sinusoidal. The disadvantage is that 13 parameters, rather than 4 (as in the sinusoidal representation assumed earlier), are required to determine the χ_t. This may be expected to lead to statistical errors in curve fitting the parameters when the number of data points (i.e., N) is sparse.

A posteriori tests of goodness-of-fit can help to determine which of these (or other) representations provides a best description of available data, in individual cases.

It is assumed, as before, that the effect of increasing police visibility is to change the crime rate by some factor Θ, to be estimated. Thus, during treatment (i.e., t = N+1, N+2, ..., N+M):

$$U_t = \left[\sum_{i=1}^{12} a_i \, \mu_{ti} + b\tau + \eta_t \right] \Theta$$

where, again for notational convenience, the t, ν_{ti}, and ε_t have been replaced by τ, μ_{ti}, and η_t respectively, to distinguish them from pretreatment values.

Assuming this as the appropriate representation, the formulas required to estimate $\hat{\Theta}$ and $\hat{\sigma}_{\hat{\Theta}}$ are derived in Appendix III. As expressed in matrix form, the results are:

$$\hat{\Theta} = \frac{\underline{\chi}^T Y \underline{U}}{\underline{\chi}^T W \underline{\chi}}$$

$$\hat{\sigma}^2_{\hat{\Theta}} = \frac{\underline{\chi}^T V \underline{\chi}}{\underline{1}^T \text{ diag } V} \left\{ \frac{1}{(\underline{\chi}^T W \underline{\chi})} + \frac{(\underline{\chi}^T Z \underline{\chi})}{(\underline{\chi}^T W \underline{\chi})^2} \right\} \frac{(\underline{\chi}^T Y \underline{U})^2}{(\underline{\chi}^T W \underline{\chi})^2}$$

where $\underline{\chi}$, $\underline{U}$ are vectors of crime data

superscript T indicates the transpose

$\underline{1}$ is a vector whose components are all ones

diag V = a vector whose components are the major diagonal elements of the square matrix V

and, using superscript -1 to indicate the inverse,

$$\left\{\begin{array}{lcl} Y & = & \Omega\ (\Omega^T\Omega)^{-1}\ \Gamma^T \\ W & = & \Omega\ (\Omega^T\Omega)^{-1}\ \Gamma^T\Gamma\ (\Omega^T\Omega)^{-1}\ \Omega^T \\ Z & = & \Omega\ (\Omega^T\Omega)^{-1}\ \Gamma^T\Gamma\ (\Omega^T\Omega)^{-1}\ \Gamma^T\Gamma\ (\Omega^T\Omega)^{-1}\ \Omega^T \\ V & = & I - \Omega\ (\Omega^T\Omega)^{-1}\ \Omega^T \end{array}\right.$$

where

I = the identity matrix

and Ω, Γ are the partitioned matrices

$$\left\{\begin{array}{lcl} \Omega & = & (\nu\ \underline{t}) \\ \Gamma & = & (\mu, \underline{\tau}) \end{array}\right.$$

where $\underline{t}$ and $\underline{\tau}$ are the vectors

$$\left\{\begin{array}{l} \underline{t} \;=\; \begin{pmatrix} 1 \\ 2 \\ \cdots \\ \cdots \\ N \end{pmatrix} \\ \\ \underline{\tau} \;=\; \begin{pmatrix} N+1 \\ N+2 \\ \cdots \\ \cdots \\ N+M \end{pmatrix} \end{array}\right.$$

and ν, μ are the matrices of the ν_{ti}, $\mu_{\tau i}$ respectively.

4.3 Use of Model Results

The central research question in the test of the PHT involves the relationship between police visibility and crime levels. Related and subsidiary questions include the differential effects of VPP on crime types, crime categories, and crime situations. By applying the model described above to each crime type, crime category and situation of interest, results can be obtained which describe the confidence one has that the level of each crime type or crime category has decreased in the time period following the introduction of increased visibility.

For each sample project, target area crime data for Index crimes, person crimes, specific Impact crimes and outdoor or suppressible crimes will be analyzed using the model. The results obtained, describing the confidence that crime has decreased in the target area since VPP was introduced, speak to these research questions. A high level of confidence that a crime type or category has decreased indicates that the hypothesis is valid, that since VPP was introduced a decline in the level of that crime type or category has been realized.

However, if a low level of confidence is obtained, it would appear that the hypothesis is not valid, that since the increase in VPP was implemented no decline in crime levels has occurred.

These results for the sample areas in the study can be displayed as is shown below in Figure 3:

			Specific Types of Crime			
	Index Crime	Person Crime	Robbery	Burglary	•••	Outdoor Crime
Project A						
Project B	*					
Project C ⋮ Project N						

*% confidence that crime has <u>decreased</u>.

FIGURE 3

FORMAT FOR MODEL RESULTS FOR SAMPLE AREAS

This display will allow for the examination of results across projects and for the comparison of results for each crime category within and across sample areas.

Using both profiles of the individual target areas which describe the context into which VPP was introduced and descriptions of the nature and extent of the VPP increase in each area, these results can be sorted out as to similarities and differences in environmental factors and measures of VPP. The results can then be interpreted in this context. In addition, any project-originated information on police performance and/or police and public attitudes will also be utilized in interpreting results of this crime level analysis.

A similar procedure will be followed to assess changes in crime level in adjacent geographical areas and during non-treatment time slots. Figure 4 shows a display possible for the presentation of these analysis results:

Areas adjacent to:	Index Crime	Person Crime	Specific Types of Crime	Outdoor Crime
Project A				
Project B	*			
Project C ⋮ Project N				

*% confidence that crime has increased.

FIGURE 4

FORMAT FOR MODEL RESULTS FOR ADJACENT AREAS

The results presented in the above manner speak to the subsidiary and related research questions in the test plan. These results can be interpreted as follows: A high percentage of confidence that a crime type or category has increased in adjacent areas for a particular case would indicate that the increase in VPP in the target area had been accompanied by an increase in crime in peripheral areas. This is an indication of a possible displacement effect. If the opposite results were obtained and there was a low percentage of confidence that the level of a crime type or category had increased, this would indicate that possible displacement effects were absent and that a possible ramified effect is in operation.

This type of information can be generated using any one of the several models suggested in the previous section. It is planned that each of the models discussed above will be applied using the data from our sample cases. Goodness-of-fit tests will be applied to determine which of the alternative models best describes the situation. This strategy will provide information on the nature of the crime level changes which accompany increases of visible police presence.

In summary, it is hoped that the strategy described in this document will generate useful information which will allow for evaluation of a number of central issues involved in the hypothesis that: "An increase in the visible presence of police in a given area will result in a decrease in crime levels in that area."

APPENDIX I

STRATEGY FOR IMPLEMENTING THE POLICE HYPOTHESIS TEST

Impact Police Patrol Projects are to be selected for inclusion in the PHT on the basis of amenability to the test plan. Selected projects will involve increases in visible police presence in defined target areas.

For each project in the test, data on crime levels in target and adjacent areas will be collected for both the project duration and for the time periods preceding the project. These data will be used in the manner described in the analysis section (4.0) of the methodology. Other information being collected as part of the project level evaluations will be used in interpreting the results of the crime level analysis.

Data on the police patrol (level of manpower, mode of patrol...) are available in most cases at the project level. Information on the characteristics of the target community when not included in project materials will be collected from relevant city agencies.

Brief project descriptions of possible candidate projects are included to give the reader a picture of the type of projects to be utilized in this test of the police hypothesis. Whether the projects listed above are actually included as sample cases in the test will depend on the availability of crime data for the project target areas. Several other Impact projects which appear to be amenable to the test plan are in a planning phase and hopefully will be included in the test.

The possible candidate projects for the PHT include:

Overtime Patrol Program: Atlanta

This project will provide additional patrol units in two high crime areas to act as preventive patrols that are unencumbered by routine duties. The number of man hours devoted to aggressive preventive patrol in these high crime areas will be markedly increased, and it is estimated that there will be an increase in the apprehension of criminal suspects and a decrease in major crimes. Additionally, there are at least two by-products that are anticipated from this project. They are: (1) a reduction in fear on the part of the residents and businessmen in the areas concerned; and (2) the achievement of a better citizen image of the police.

Sixty-Four Foot Patrolmen: Baltimore

The objective of the project is to reduce Impact crimes by providing 64 patrolmen to supplement the motorized police force and improve police service. The foot patrol force is to be assigned to those areas determined by analysis. It is theorized that the force, in conjunction with motorized units, will be successful in deterring and preventing Impact target crimes being committed in their respective areas.

Concentrated Crime Prevention Patrol: Cleveland

The objective of the project is to deter crime through high visibility patrols. The project consists of hiring and training new police officers to facilitate the deployment of a crime prevention patrol. This patrol will be concentrated in high-crime areas making itself highly visible in an effort to prevent crime. The crime prevention patrol will insure a rapid response to all emergency calls, improving the response time, the rate of apprehension of offenders and the feeling of well-being of the citizens.

Special Crime Attack Team: Denver

The Special Crime Attack Team is formed as an integral part of the Denver Police Department. The team consists of thirty-three personnel including a commander and a mix of patrolmen, detectives and evidence technicians. The team is deployed utilizing "computer analysis of crime data." The three major tactics employed are: prevention, interception and investigation. An important aspect of the SCAT concept is that of a highly mobile group which can be rapidly deployed to meet daily responses to crime data analyses.

Foot Patrol (6 month pilot effort): St. Louis

The objective is to provide foot patrol in high crime areas during high incidence times in an effort to reduce Impact crime. The foot patrol will be performed on an overtime basis in six Pauly areas chosen for their high rate of crime. A secondary objective is to promote good community relations by reassuring the citizens that the St. Louis Police Department is providing a preventive force against crime. The foot patrol assigned to an area will be in addition to the regular police units in that area.

APPENDIX II

DERIVATION OF FORMULAS FOR $\hat{\theta}$, $\hat{\sigma}_{\hat{\theta}}$: MODEL I

Consider some particular type of crime, committed in some specified neighborhood during some specified time-of-day interval. Let X_t ($t=1, 2, \ldots, N$) and U_τ ($\tau = N+1, N+2, \ldots, N+M$) be monthly crime rates for the N months prior to treatment and during M months of treatment, respectively. The X_t and U_τ are assumed available as data. It is assumed that the X_t and U_τ are generated according to:

$$\begin{cases} X_t = a + bt + c \sin\left(\frac{\pi t}{6}\right) + d \cos\left(\frac{\pi t}{6}\right) + \varepsilon_t \\ U_\tau = \left[a + b\tau + c \sin\left(\frac{\pi \tau}{6}\right) + d \cos\left(\frac{\pi \tau}{6}\right) + \eta_\tau\right]\theta \end{cases}$$

where $\begin{cases} \text{a, b, c, d are parameters, to be determined;} \\ \theta \text{ is a parameter measuring the effect of treatment;} \\ \varepsilon_t, \eta_\tau \text{ are random variables, independent, and are normal} \\ \text{with zero mean and variance } \sigma_\varepsilon^2. \end{cases}$

The notational distinction of t vs. τ, X vs. U, and ε vs. η to indicate pre-treatment vs. during-treatment values is made merely for convenience in the analysis which follows. The above equations are compactly expressed in matrix form by defining:

$$\underline{X} = \begin{pmatrix} X_1 \\ X_2 \\ \cdots \\ \cdots \\ X_N \end{pmatrix}, \quad \underline{U} = \begin{pmatrix} U_{N+1} \\ U_{N+2} \\ \cdots \\ \cdots \\ U_{N+M} \end{pmatrix}, \quad \underline{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \cdots \\ \varepsilon_N \end{pmatrix}$$

$$\underline{\eta} = \begin{pmatrix} \eta_{N+1} \\ \eta_{N+2} \\ \cdots \\ \cdots \\ \eta_{N+M} \end{pmatrix}$$

$$\underline{\alpha} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$$

$$G = \begin{pmatrix} 1 & 1 & \sin\left(\frac{\pi}{6}\right) & \cos\left(\frac{\pi}{6}\right) \\ 1 & 2 & \sin\left(\frac{2\pi}{6}\right) & \cos\left(\frac{2\pi}{6}\right) \\ \cdots & & & \\ \cdots & & & \\ 1 & N & \sin\left(\frac{N\pi}{6}\right) & \cos\left(\frac{N\pi}{6}\right) \end{pmatrix}$$

$$H = \begin{pmatrix} 1 & N+1 & \sin\left[\frac{(N+1)\pi}{6}\right] & \cos\left[\frac{(N+1)\pi}{6}\right] \\ 1 & N+2 & \sin\left[\frac{(N+2)\pi}{6}\right] & \cos\left[\frac{(N+2)\pi}{6}\right] \\ \cdots & & & \\ \cdots & & & \\ 1 & N+M & \sin\left[\frac{(N+M)\pi}{6}\right] & \cos\left[\frac{(N+M)\pi}{6}\right] \end{pmatrix}$$

whereupon it is readily verified that

$$\begin{cases} \underline{X} = G\,\underline{\alpha} + \underline{\varepsilon} \\ \underline{U} = (H\,\underline{\alpha} + \underline{\eta})\theta \end{cases}$$

It appears preferable to estimate $\underline{\alpha}$ from pre-treatment data alone, since such data will not be contaminated by any transient terms present in $\underline{U}$ not accounted for in the essentially steady-state model above. Thus, the estimate $\hat{\underline{\alpha}}$ is to be chosen to minimize

$$\sum_t\left(X_t - \sum_j G_{tj}\hat{\alpha}_j\right)^2$$

whereupon

$$\sum_t\left(X_t - \sum_j G_{tj}\hat{\alpha}_j\right) G_{tk} = 0 \qquad \forall k$$

i.e., $\sum_t\sum_j G_{tj} G_{tk}\hat{\alpha}_j = \sum_t X_t G_{tk}$

Denoting the transpose of a matrix by superscript T and the inverse of a matrix by superscript -1, it follows that

$$G^T G\underline{\hat{\alpha}} = G^T\underline{X}$$

whereupon

$$\underline{\hat{\alpha}} = \left(G^T G\right)^{-1} G^T\underline{X}$$

Using this $\hat{\alpha}$, an estimate $\hat{\theta}$ is to be determined to minimize

$$\sum_\tau\left(U_\tau - \sum_j H_{\tau j}\hat{\alpha}_j\theta\right)^2$$

The condition for a minimum is

$$\sum_\tau\left[U_\tau - \sum_j H_{\tau j}\hat{\alpha}_j\hat{\theta}\right]\sum_k H_{\tau k}\hat{\alpha}_k = 0$$

whereupon

$$\sum_\tau U_\tau \sum_k H_{\tau k}\hat{\alpha}_k = \sum_\tau\sum_j\sum_k \hat{\theta} H_{\tau j} H_{\tau k}\hat{\alpha}_k\hat{\alpha}_j$$

In matrix form,

$$\underline{U}^T H\underline{\hat{\alpha}} = \hat{\theta}(H\underline{\hat{\alpha}})^T H\underline{\hat{\alpha}}$$

$$\therefore \quad \hat{\theta} = \frac{\underline{U}^T H\underline{\hat{\alpha}}}{(H\underline{\hat{\alpha}})^T (H\underline{\hat{\alpha}})}$$

$$= \frac{\underline{U}^T H\left(G^T G\right)^{-1} G^T\underline{X}}{\underline{X}^T G\left(G^T G\right)^{-1} H^T H\left(G^T G\right)^{-1} G^T\underline{X}}$$

$$= \frac{\underline{U}^T J\underline{X}}{\underline{X}^T K\underline{X}}$$

where $\begin{cases} J = H\left(G^T G\right)^{-1} G^T \\ K = G\left(G^T G\right)^{-1} H^T H\left(G^T G\right)^{-1} G^T \end{cases}$

It remains to estimate $\sigma_{\hat{\theta}}$. Now,

$$\begin{aligned} \underline{U}^T J\underline{X} &= \theta\left(H\underline{\alpha} + \underline{\eta}\right)^T J\left(G\underline{\alpha} + \underline{\varepsilon}\right) \\ &= \theta\left(\underline{\alpha}^T H^T + \underline{\eta}^T\right)\left(JG\underline{\alpha} + J\underline{\varepsilon}\right) \\ &= \theta\left(\underline{\alpha}^T H^T JG\underline{\alpha} + \underline{\alpha}^T H^T J\underline{\varepsilon} + \underline{\eta}^T JG\underline{\alpha} + \underline{\eta}^T J\underline{\varepsilon}\right) \end{aligned}$$

Neglecting error terms higher than linear in $\underline{\eta}$, $\underline{\varepsilon}$ as "small,"

$$\underline{U}^T J\underline{X} \doteq \theta\left(\underline{\alpha}^T H^T JG\underline{\alpha} + \underline{\alpha}^T H^T J\underline{\varepsilon} + \underline{\eta}^T JG\underline{\alpha}\right)$$

But $JG = H\left(G^T G\right)^{-1} G^T G$

$= H$

so that

$$\underline{U}^T J\underline{X} = \theta\underline{\alpha}^T H^T H\underline{\alpha}\left\{1 + \frac{\underline{\alpha}^T H^T J\underline{\varepsilon}}{\underline{\alpha}^T H^T H\underline{\alpha}} + \frac{\underline{\eta}^T H\underline{\alpha}}{\underline{\alpha}^T H^T H\underline{\alpha}}\right\}$$

Again,

$$\begin{aligned} \underline{X}^T K\underline{X} &= \left(G\underline{\alpha} + \underline{\varepsilon}\right)^T K\left(G\underline{\alpha} + \underline{\varepsilon}\right) \\ &= \left(\underline{\alpha}^T G^T + \underline{\varepsilon}^T\right)\left(KG\underline{\alpha} + K\underline{\varepsilon}\right) \\ &\doteq \underline{\alpha}^T G^T KG\underline{\alpha} + \underline{\alpha}^T G^T K\underline{\varepsilon} + \underline{\varepsilon}^T KG\underline{\alpha} \end{aligned}$$

Thus

$$\frac{1}{\underline{X}^T K \underline{X}} = \frac{1}{\underline{\alpha}^T G^T K G \underline{\alpha} \left\{ 1 + \frac{\underline{\alpha}^T G^T K \underline{\varepsilon}}{\underline{\alpha}^T G^T K G \underline{\alpha}} + \frac{\underline{\varepsilon}^T K G \underline{\alpha}}{\underline{\alpha}^T G^T K G \underline{\alpha}} \right\}}$$

$$\doteq \frac{1}{\underline{\alpha}^T G^T K G \underline{\alpha}} \left\{ 1 - \frac{\underline{\alpha}^T G^T K \underline{\varepsilon}}{\underline{\alpha}^T G^T K G \underline{\alpha}} - \frac{\underline{\varepsilon}^T K G \underline{\alpha}}{\underline{\alpha}^T G^T K G \underline{\alpha}} \right\}$$

neglecting terms higher than linear in $\underline{\varepsilon}$. It follows that

$$\theta = \theta \frac{\underline{\alpha}^T H^T H \underline{\alpha}}{\underline{\alpha}^T G^T K G \underline{\alpha}} \left\{ 1 + \frac{\underline{\alpha}^T H^T J \underline{\varepsilon}}{\underline{\alpha}^T H^T H \underline{\alpha}} + \frac{\underline{\eta}^T H \underline{\alpha}}{\underline{\alpha}^T H^T H \underline{\alpha}} \right\} \left\{ 1 - \frac{\underline{\alpha}^T G^T K \underline{\varepsilon}}{\underline{\alpha}^T G^T K G \underline{\alpha}} - \frac{\underline{\varepsilon}^T K G \underline{\alpha}}{\underline{\alpha}^T G^T K G \underline{\alpha}} \right\}$$

Again neglecting terms higher than linear in $\underline{\varepsilon}$, $\underline{\eta}$ and using the facts that

$$\begin{aligned} G^T K G &= G^T G \left(G^T G\right)^{-1} H^T H \left(G^T G\right)^{-1} G^T G \\ &= H^T H \end{aligned}$$

and

$$\begin{aligned} G^T K &= G^T \left(G^T G\right)^{-1} H^T H \left(G^T G\right)^{-1} G^T \\ &= H^T H \left(G^T G\right)^{-1} G^T \\ &= H^T J \end{aligned}$$

it follows that

$$\hat{\theta} = \theta \left\{ 1 + \frac{\underline{\eta}^T H \underline{\alpha}}{\underline{\alpha}^T H^T H \underline{\alpha}} - \frac{\underline{\varepsilon}^T K G \underline{\alpha}}{\underline{\alpha}^T H^T H \underline{\alpha}} \right\}$$

Using E to denote expected value, it follows immediately that

$$E\left(\hat{\theta}\right) = \theta$$

So that $\hat{\theta}$ is (within the approximations assumed) an unbiased estimate of θ.

The variance of $\hat{\theta}$ is then given by:

$$\begin{aligned} \sigma^2_{\hat{\theta}} &= E\left[\hat{\theta} - E\left(\hat{\theta}\right)\right]^2 \\ &= \frac{\theta^2}{\left(\underline{\alpha}^T H^T H \underline{\alpha}\right)^2} E\left(\underline{\eta}^T H \underline{\alpha} - \underline{\varepsilon}^T K G \underline{\alpha}\right)^2 \end{aligned}$$

Thus

$$\frac{\left(\underline{\alpha}^T H^T H \underline{\alpha}\right)^2}{\theta^2} \sigma^2_{\hat{\theta}} = E\left\{ \left[\sum_{\tau} \eta_{\tau} \left(H\underline{\alpha}\right)_{\tau} - \sum_{t} \varepsilon_t \left(KG\underline{\alpha}\right)_t \right] \left[\sum_{\tau'} \eta_{\tau'} \left(H\underline{\alpha}\right)_{\tau'} - \sum_{t'} \varepsilon_{t'} \left(KG\underline{\alpha}\right)_{t'} \right] \right\}$$

$$= \sum_{\tau} \sum_{\tau'} \left(H\underline{\alpha}\right)_{\tau} \left(H\underline{\alpha}\right)_{\tau'} E\left(\eta_{\tau} \eta_{\tau'}\right) + \sum_{t} \sum_{t'} \left(KG\underline{\alpha}\right)_t \left(KG\underline{\alpha}\right)_{t'} E\left(\varepsilon_t \varepsilon_{t'}\right)$$

Since terms involving $E\left(\varepsilon_t \eta_{\tau}\right)$ are zero, by assumption. It follows that

$$\begin{aligned} \frac{\left(\underline{\alpha}^T H^T H \underline{\alpha}\right)^2}{\theta^2} \sigma^2_{\hat{\theta}} &= \sigma^2_{\varepsilon} \left[\sum_{\tau} \left(H\underline{\alpha}\right)^2_{\tau} + \sum_{t} \left(KG\underline{\alpha}\right)^2_t \right] \\ &= \sigma^2_{\varepsilon} \left[\left(H\underline{\alpha}\right)^T H \underline{\alpha} + \left(KG\underline{\alpha}\right)^T KG\underline{\alpha} \right] \\ &= \sigma^2_{\varepsilon} \underline{\alpha}^T \left[H^T H + G^T K^T K G \right] \underline{\alpha} \end{aligned}$$

But

$$\begin{aligned} G^T K^T K G &= G^T G \left(G^T G\right)^{-1} H^T H \left(G^T G\right)^{-1} \left(G^T G\right) \left(G^T G\right)^{-1} H^T H \left(G^T G\right)^{-1} G^T G \\ &= H^T H \left(G^T G\right)^{-1} H^T H \end{aligned}$$

$$\therefore \sigma^2_{\hat{\theta}} = \frac{\theta^2 \sigma^2_{\varepsilon}}{\left(\underline{\alpha}^T H^T H \underline{\alpha}\right)^2} \underline{\alpha}^T H^T \left[I + H \left(G^T G\right)^{-1} H^T \right] H \underline{\alpha}$$

where I is the identity matrix. The desired variance, $\sigma^2_{\hat{\theta}}$, can be estimated from data by using $\underline{\hat{\alpha}}$ in place of $\underline{\alpha}$, $\hat{\theta}^2$ in place of θ^2, and a suitable estimator (to be derived) for σ^2_{ε}. Thus,

$$\sigma^2_{\hat{\theta}} = \frac{\hat{\theta}^2 \; \hat{\sigma}_{\varepsilon}^{\;2}}{\left(\underline{\hat{\alpha}}^T H^T H \underline{\hat{\alpha}}\right)^2} \, \underline{\hat{\alpha}}^T H^T \left[I + H\left(G^TG\right)^{-1} H^T \right] H \underline{\hat{\alpha}}$$

But $\underline{\hat{\alpha}}^T H^T H \underline{\hat{\alpha}} = \underline{X}^T G\left(G^TG\right)^{-1} H^T H \left(G^TG\right)^{-1} G^T \underline{X}$

$$= \underline{X}^T K \underline{X}$$

and $\underline{\hat{\alpha}}^T H^T H \left(G^TG\right)^{-1} H^T H \underline{\hat{\alpha}} = \underline{X}^T G\left(G^TG\right)^{-1} H^T H \left(G^TG\right)^{-1} H^T H \left(G^TG\right)^{-1} G^T \underline{X}$

$$= \underline{X}^T P \underline{X}$$

where $\quad P \equiv G\left(G^TG\right)^{-1} H^T H \left(G^TG\right)^{-1} H^T H \left(G^TG\right) G^T$

Thus

$$\hat{\sigma}^2_{\hat{\theta}} = \hat{\theta}^2 \hat{\sigma}^2_{\varepsilon} \left\{ \frac{1}{\left(\underline{X}^T K \underline{X}\right)} + \frac{\left(\underline{X}^T P \underline{X}\right)}{\left(\underline{X}^T K \underline{X}\right)^2} \right\}$$

$$= \hat{\sigma}^2_{\varepsilon} \frac{\left(\underline{U}^T J \underline{X}\right)^2}{\left(\underline{X}^T K \underline{X}\right)^3} \left\{ 1 + \frac{\left(\underline{X}^T P \underline{X}\right)}{\left(\underline{X}^T K \underline{X}\right)} \right\}$$

It remains to estimate σ^2_{ε}.

Assume that this is to be done from the pre-treatment data alone.

Now, $\quad \underline{\varepsilon} = \underline{X} - G\underline{\alpha}$

so that $\quad \underline{\hat{\varepsilon}} = \underline{X} - G\underline{\hat{\alpha}}$

$$= \underline{X} - G\left(G^TG\right)^{-1} G^T \underline{X}$$

$$= \left[I - G\left(G^TG\right)^{-1} G^T \right] \underline{X}$$

Let $R = I - G\left(G^TG\right)^{-1} G^T$

Then $\underline{\hat{\varepsilon}} = R\underline{X}$

$$= R\left(G\underline{\alpha} + \underline{\varepsilon}\right)$$

$$= RG\underline{\alpha} + R\underline{\varepsilon}$$

But $RG = \left[I - G\left(G^TG\right)^{-1} G^T \right] G$

$$= G - G\left(G^TG\right)^{-1} G^T G = 0$$

Thus, $\underline{\hat{\varepsilon}} = R\underline{\varepsilon}$

whereupon $E\left(\underline{\hat{\varepsilon}}\right) = \underline{0}$.

Define $S = \sum_t \varepsilon_t^2$

$$= \sum_t \left(R\underline{X}\right)_t^2$$

$$= \left(R\underline{X}\right)^T R\underline{X}$$

$$= \underline{X}^T R^T R \underline{X}$$

Now, it is readily verified that

$$R^T R = R$$

Thus, $S = \underline{X}^T R \underline{X}$

But $\quad S = \sum_t \hat{\varepsilon}_t^2$

$$= \sum_t \left(R\underline{\varepsilon}\right)_t^2$$

$$= \left(R\underline{\varepsilon}\right)^T R\underline{\varepsilon}$$

$$= \underline{\varepsilon}^T R^T R \underline{\varepsilon}$$

$$= \underline{\varepsilon}^T R \underline{\varepsilon}$$

$$= \sum_t \sum_{t'} R_{tt'} \; \varepsilon_t \varepsilon_{t'}$$

so that

$$E(S) = \sum_t \sum_{t'} R_{tt'} E\left(\varepsilon_t \varepsilon_{t'}\right)$$

$$= \sigma^2_{\varepsilon} \sum_t R_{tt}$$

$$= \sigma^2_{\varepsilon} \, \underline{1}^T \text{ diag } R$$

wnere $\underline{1}$ is a vector whose components are all ones and diag R is a vector whose components are those of the principal diagonal of the square matrix R.

Define $\quad \hat{\sigma}^2_{\varepsilon} = \dfrac{\underline{X}^T R \underline{X}}{\underline{1}^T \text{ diag } R}$

Then $\quad E\left(\hat{\sigma}^2_{\varepsilon}\right) = E \, \dfrac{S}{\underline{1}^T \text{ diag } R}$

$$= \frac{E(S)}{\underline{1}^T \text{ diag } R}$$

$$= \sigma_\varepsilon^2$$

so that $\hat{\sigma}_\varepsilon^2$ is an unbiased estimator of σ_ε^2.

Summary of Results for Model I

$$\hat{\theta} = \frac{\left(\underline{U}^T J \underline{X}\right)}{\left(\underline{X}^T K \underline{X}\right)}$$

$$\hat{\sigma}_{\hat{\theta}}^2 = \frac{\left(\underline{X}^T R \underline{X}\right)}{\underline{1}^T \text{ diag } R} \frac{\left(\underline{U}^T J \underline{X}\right)^2}{\left(\underline{X}^T K \underline{X}\right)^3} \left\{ 1 + \frac{\left(\underline{X}^T P \underline{X}\right)}{\left(\underline{X}^T K \underline{X}\right)} \right\}$$

where
$$\begin{cases} J = H\left(G^T G\right)^{-1} G^T \\ K = G\left(G^T G\right)^{-1} H^T H \left(G^T G\right)^{-1} G^T \\ P = G\left(G^T G\right)^{-1} H^T H \left(G^T G\right)^{-1} H^T H \left(G^T G\right)^{-1} G^T \\ R = I - G\left(G^T G\right)^{-1} G^T \end{cases}$$

APPENDIX III

DERIVATION OF FORMULAS FOR $\hat{\theta}$, $\sigma_{\hat{\theta}}$: MODEL II

The general assumptions, approach, and notation in this Appendix are the same as in Appendix II, except where otherwise indicated.

In Model II, it is assumed that the X_t and U_τ are generated by:

$$\begin{cases} X_t = \sum_{i=1}^{12} a_i \nu_{ti} + bt + \varepsilon_t \\ U_\tau = \left[\sum_{i=1}^{12} a_i \mu_{\tau i} + b\tau + \eta_\tau \right] \Theta \end{cases}$$

where the ν_{ti}, $\mu_{\tau i}$ are indicator variables that distinguish month t or τ as January, February, etc. For example, if the pre-treatment data started (t = 1) in January, the matrix ν would appear as:

$$\nu = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & & & & \\ \dots & & & & \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & & & & \\ \dots & & & & \\ 0 & 0 & 0 & \dots & 1 \\ \dots & & & & \\ \dots & & & & \end{bmatrix}$$

The dimensions of ν are N x 12, and the dimensions of μ, defined similarly, are M x 12.

Define the vectors

$$t = \begin{pmatrix} 1 \\ 2 \\ \cdots \\ \cdots \\ N \end{pmatrix}$$

$$\tau = \begin{pmatrix} N+1 \\ N+2 \\ \cdots \\ \cdots \\ N+M \end{pmatrix}$$

$$\underline{a} = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ \cdots \\ a_{12} \end{pmatrix}$$

$$\underline{A} = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ \cdots \\ a_{12} \\ b \end{pmatrix}$$

and the partitioned matrices

$$\Omega = (\nu \ \underline{t})$$

$$\Gamma = (\mu \ \underline{\tau})$$

Then

$$\begin{cases} \underline{X} = \Omega \underline{A} + \underline{\varepsilon} \\ \underline{U} = (\Gamma \underline{A} + \underline{\eta}) \ \Theta \end{cases}$$

Proceeding as in Appendix II.

$$\hat{\underline{A}} = (\Omega^T \Omega)^{-1} \Omega^T \underline{X}$$

and

$$\hat{\Theta} = \frac{(\Gamma\hat{\underline{A}})^T \underline{U}}{(\Gamma\hat{\underline{A}})^T (\Gamma\underline{A})}$$

Combining,

$$\Theta = \frac{\underline{X}^T \ \Omega(\Omega^T\Omega)^{-1} \ \Gamma^T \ \underline{U}}{\underline{X}^T \ \Omega(\Omega^T\Omega)^{-1} \ \Gamma^T\Gamma(\Omega^T\Omega)^{-1} \ \Omega^T \ \underline{X}}$$

Define $\quad Y \equiv \Omega(\Omega^T\Omega)^{-1} \ \Gamma$

Then $\quad \hat{\Theta} = \dfrac{\underline{X}^T \ Y \ \underline{U}}{\underline{X}^T \ YY^T \ \underline{X}}$

It remains to compute $E(\hat{\Theta})$ and $\sigma^2_{\hat{\Theta}}$.

Now,

$$\begin{aligned} \underline{X}^T Y \underline{U} &= (\Omega\underline{A} + \underline{\varepsilon})^T \ Y \ (\Theta\Gamma\underline{A} + \Theta\underline{\eta}) \\ &= \Theta \ (\underline{A}^T\Omega^T + \underline{\varepsilon}^T) \ Y \ (\Gamma\underline{A} + \underline{\eta}) \\ &= \Theta \ (\underline{A}^T\Omega^T + \underline{\varepsilon}^T)(Y\Gamma\underline{A} + Y\underline{\eta}) \\ &\doteq \Theta \left[\underline{A}^T\Omega^T Y\Gamma\underline{A} + \underline{\varepsilon}^T Y\Gamma\underline{A} + \underline{A}^T\Omega^T Y\underline{\eta}\right] \end{aligned}$$

after neglecting terms higher than linear in $\underline{\varepsilon}$, $\underline{\eta}$.

Again,

$$\begin{aligned}\underline{X}^T YY^T \underline{X} &= (\Omega\underline{A} + \underline{\varepsilon})^T \; YY^T \; (\Omega\underline{A} + \underline{\varepsilon}) \\ &= (\underline{A}^T\Omega^T + \underline{\varepsilon}^T)(YY^T\Omega\underline{A} + YY^T\underline{\varepsilon}) \\ &= \underline{A}^T\Omega^T YY^T\Omega\underline{A} \\ &\qquad + \underline{A}^T\Omega^T YY^T\underline{\varepsilon} \\ &\qquad\quad + \underline{\varepsilon}^T YY^T\Omega\underline{A}\end{aligned}$$

after neglecting terms higher than linear in $\underline{\varepsilon}$.

Thus,

$$\frac{1}{\underline{X}^T YY^T \underline{X}} = \frac{1}{\underline{A}^T\Omega^T YY^T\Omega\underline{A}\left\{1 + \dfrac{\underline{A}^T\Omega^T YY^T\underline{\varepsilon}}{A^T\Omega^T YY^T\Omega\underline{A}} + \dfrac{\underline{\varepsilon}^T YY^T\Omega\underline{A}}{\underline{A}^T\Omega^T YY^T\Omega\underline{A}}\right\}}$$

$$\doteq \frac{1}{\underline{A}^T\Omega^T YY^T\Omega\underline{A}} \left\{1 - \frac{A^T\Omega^T YY^T\underline{\varepsilon}}{\underline{A}^T\Omega^T YY^T\Omega\underline{A}} - \frac{\underline{\varepsilon}^T YY^T\Omega\underline{A}}{\underline{A}^T\Omega^T YY^T\Omega\underline{A}}\right\}$$

It follows that, for small $\underline{\varepsilon}$, $\underline{\eta}$:

$$\hat{\Theta} = \frac{\Theta\underline{A}^T\Omega^T Y\Gamma\underline{A}}{\underline{A}^T\Omega^T YY^T\Omega\underline{A}} \left\{1 + \frac{\underline{\varepsilon}^T Y\Gamma\underline{A}}{\underline{A}^T\Omega^T Y\Gamma\underline{A}} + \frac{\underline{A}^T\Omega^T Y\underline{\eta}}{\underline{A}^T\Omega^T Y\Gamma\underline{A}}\right\}$$

$$\left\{1 - \frac{\underline{A}^T\Omega^T YY^T\underline{\varepsilon}}{\underline{A}^T\Omega^T YY^T\Omega\underline{A}} - \frac{\underline{\varepsilon}^T YY^T\Omega\underline{A}}{\underline{A}^T\Omega^T YY^T\Omega\underline{A}}\right\}$$

$$= \Theta \; \frac{\underline{A}^T\Omega^T Y\Gamma\underline{A}}{\underline{A}^T\Omega^T YY^T\Omega\underline{A}} \left\{1 + \frac{\underline{\varepsilon}^T Y\Gamma\underline{A}}{\underline{A}^T\Omega^T Y\Gamma\underline{A}} \right.$$

$$\left. + \frac{\underline{A}^T\Omega^T Y\underline{\eta}}{\underline{A}^T\Omega^T Y\Gamma\underline{A}} - \frac{\underline{A}^T\Omega^T YY^T\underline{\varepsilon}}{\underline{A}^T\Omega^T YY^T\Omega\underline{A}} - \frac{\underline{\varepsilon}^T YY^T\Omega\underline{A}}{\underline{A}^T\Omega^T YY^T\Omega\underline{A}}\right\}$$

But $\quad Y^T\Omega = \Gamma(\Omega^T\Omega)^{-1}\Omega^T\Omega$

$\therefore \quad Y^T\Omega = \Gamma$

and $\quad \Omega^T Y = \Gamma^T$

Substituting, it follows that

$$\hat{\Theta} = \Theta \left\{1 + \frac{\underline{A}^T\Gamma^T\underline{\eta}}{\underline{A}^T\Gamma^T\Gamma\underline{A}} - \frac{\underline{A}^T\Gamma^T Y^T\underline{\varepsilon}}{\underline{A}^T\Gamma^T\Gamma\underline{A}}\right\}$$

Thus, for small $\underline{\varepsilon}$, $\underline{\eta}$:

$$E(\hat{\Theta}) = \Theta$$

so that $\hat{\Theta}$ is an unbiased estimate of Θ. It follows that

$$\sigma^2_{\hat{\Theta}} = \frac{\Theta^2}{(\underline{A}^T\Gamma^T\Gamma\underline{A})^2} \; E\left[\underline{A}^T\Gamma^T(\underline{\eta} - Y^T\underline{\varepsilon})\right]^2$$

Thus

$$\frac{(\underline{A}^T\Gamma^T\Gamma\underline{A})^2}{\Theta^2}\,\sigma^2_{\hat{\Theta}} = E\left\{\left[\sum_\tau (\underline{A}^T\Gamma^T)_\tau\eta_\tau - \sum_t (\underline{A}^T\Gamma^T Y^T)_t\,\varepsilon_t\right]\right.$$

$$\left.\left[\sum_{\tau'} (\underline{A}^T\Gamma^T)_{\tau'}\eta_{\tau'} - \sum_{t'} (\underline{A}^T\Gamma^T Y^T)_{t'}\varepsilon_{t'}\right]\right\}$$

$$= \sum_\tau \sum_{\tau'} (\underline{A}^T\Gamma^T)_\tau \; (\underline{A}^T\Gamma^T)_{\tau'} \; E\,(\eta_\tau\eta_{\tau'})$$

$$+ \sum_t \sum_{t'} \; (\underline{A}^T\Gamma^T Y^T)_t \; (\underline{A}^T\Gamma^T Y^T)_{t'} \; E\,(\varepsilon_t\varepsilon_{t'})$$

$$= \sigma^2_\varepsilon \left[\sum_\tau (\underline{A}^T\Gamma^T)^2_\tau + \sum_t (\underline{A}^T\Gamma^T Y^T)^2_t\right]$$

$$= \sigma_{\varepsilon}^{2} \left[(\underline{A}^{T}\Gamma^{T})(\underline{A}^{T}\Gamma^{T})^{T} + (\underline{A}^{T}\Gamma^{T}Y^{T})(\underline{A}^{T}\Gamma^{T}Y^{T})^{T} \right]$$

$$= \sigma_{\varepsilon}^{2} \left[\underline{A}^{T}\Gamma^{T}\Gamma\underline{A} + \underline{A}^{T}\Gamma^{T}Y^{T}Y\Gamma\underline{A} \right]$$

Then $\hat{\sigma}_{\hat{\Theta}}^{2}$ can be estimated from

$$\hat{\sigma}_{\hat{\Theta}}^{2} = \frac{\hat{\Theta}^{2}\ \hat{\sigma}_{\varepsilon}^{2}}{(\hat{\underline{A}}^{T}\Gamma^{T}\Gamma\hat{\underline{A}})^{2}} \left[\hat{\underline{A}}^{T}\Gamma^{T}\Gamma\hat{\underline{A}} + \hat{\underline{A}}^{T}\Gamma^{T}Y^{T}Y\Gamma\hat{\underline{A}} \right]$$

But $\hat{\underline{A}}^{T}\Gamma^{T}\Gamma\hat{\underline{A}} = \underline{\chi}^{T}\Omega(\Omega^{T}\Omega)^{-1}\Gamma^{T}\Gamma(\Omega^{T}\Omega)^{-1}\Omega^{T}\underline{\chi}$

and $\hat{\underline{A}}^{T}\Gamma^{T}Y^{T}Y\Gamma\hat{\underline{A}} = \underline{\chi}^{T}\Omega(\Omega^{T}\Omega)^{-1}\Gamma^{T}\Gamma(\Omega^{T}\Omega)^{-1}\Omega^{T}\Omega$

$$(\Omega^{T}\Omega)^{-1}\Gamma^{T}\Gamma(\Omega^{T}\Omega)^{-1}\Omega^{T}\underline{\chi}$$

$$= \underline{\chi}^{T}\Omega(\Omega^{T}\Omega)^{-1}\Gamma^{T}\Gamma(\Omega^{T}\Omega)^{-1}\Gamma^{T}\Gamma(\Omega^{T}\Omega)^{-1}\Omega^{T}\underline{\chi}$$

Let $$\begin{cases} W = \Omega(\Omega^{T}\Omega)^{-1}\Gamma^{T}\Gamma(\Omega^{T}\Omega)^{-1}\Omega^{T} \\ Z = \Omega(\Omega^{T}\Omega)^{-1}\Gamma^{T}\Gamma(\Omega^{T}\Omega)^{-1}\Gamma^{T}\Gamma(\Omega^{T}\Omega)^{-1}\Omega^{T} \end{cases}$$

Then $\hat{\underline{A}}^{T}\Gamma^{T}\Gamma\hat{\underline{A}} = \underline{\chi}^{T}W\underline{\chi}$

and $\hat{\underline{A}}\Gamma^{T}Y^{T}Y\Gamma\hat{\underline{A}} = \underline{\chi}^{T}Z\underline{\chi}$

and $\underline{\chi}^{T}YY^{T}\underline{\chi} = \underline{\chi}^{T}\Omega(\Omega^{T}\Omega)^{-1}\Gamma^{T}\Gamma(\Omega^{T}\Omega)^{-1}\Omega^{T}\underline{\chi}$

$$= \underline{\chi}^{T}W\underline{\chi}$$

It follows that

$$\hat{\Theta} = \frac{\underline{\chi}^{T}Y\underline{U}}{\underline{\chi}^{T}W\underline{\chi}}$$

whereupon

$$\hat{\sigma}_{\hat{\Theta}}^{2} = \hat{\sigma}_{\varepsilon}^{2} \frac{(\underline{\chi}^{T}Y\underline{U})^{2}}{(\underline{\chi}^{T}W\underline{\chi})^{3}} \left\{ 1 + \frac{(\underline{\chi}^{T}Z\underline{\chi})}{(\underline{\chi}^{T}W\underline{\chi})} \right\}$$

To estimate $\hat{\sigma}_{\varepsilon}^{2}$, define

$$\begin{aligned} \hat{\underline{\varepsilon}} &= \underline{\chi} - \Omega\hat{\underline{A}} \\ &= \underline{\chi} - \Omega(\Omega^{T}\Omega)^{-1}\Omega^{T}\underline{\chi} \\ &= \left[I - \Omega(\Omega^{T}\Omega)^{-1}\Omega^{T} \right] \underline{\chi} \end{aligned}$$

Let $V = I - \Omega(\Omega^{T}\Omega)^{-1}\Omega^{T}$

Then $\hat{\underline{\varepsilon}} = V\ \underline{\chi}$

so that

$$\hat{\underline{\varepsilon}} = V(\Omega\underline{A} + \underline{\varepsilon})$$

$$= V\Omega\underline{A} + V\underline{\varepsilon}$$

But $V\Omega = \left[I - \Omega(\Omega^{T}\Omega)^{-1}\Omega^{T} \right] \Omega$

$$= \Omega - \Omega$$

$$= o$$

Thus $\hat{\underline{\varepsilon}} = V\underline{\varepsilon}$

It is readily verified that

$$V^{T}V = V$$

Define $$\begin{aligned} S &= \sum_{t} \hat{\varepsilon}_{t}^{2} \\ &= \sum_{t} (V\underline{\chi})_{t}^{2} \\ &= (V\underline{\chi})^{T}\ V\underline{\chi} \\ &= \underline{\chi}^{T}V^{T}V\underline{\chi} \\ &= \underline{\chi}^{T}V\underline{\chi} \end{aligned}$$

Again, $S = \sum_t \hat{\varepsilon}_t^2$

$$= \sum_t (V\underline{\varepsilon})_t^2$$

$$= (V\underline{\varepsilon})^T V\underline{\varepsilon}$$

$$= \underline{\varepsilon}^T V\underline{\varepsilon}$$

Thus $E(S) = E\left[\sum_t \sum_{t'} \varepsilon_t V_{tt'} \varepsilon_{t'}\right]$

$$= \sigma_\varepsilon^2 \sum_t V_{tt}$$

$$= \sigma_\varepsilon^2\, \underline{1}^T \text{ diag } V$$

Define

$$\hat{\sigma}_\varepsilon^2 = \frac{\underline{X}^T V \underline{X}}{\underline{1}^T \text{ diag } V}$$

Then

$$E(\hat{\sigma}_\varepsilon^2) = E\, \frac{\underline{X}^T V \underline{X}}{\underline{1}^T \text{ diag } V}$$

$$= \frac{E(S)}{\underline{1}^T \text{ diag } V}$$

$$= \sigma_\varepsilon^2$$

so that $\hat{\sigma}_\varepsilon^2$ is an unbiased estimate of σ_ε^2.

Summary or Results for Model II

$$\hat{\theta} = \frac{\underline{X}^T Y \underline{U}}{\underline{X}^T W \underline{X}}$$

$$\hat{\sigma}_{\hat{\theta}}^2 = \frac{(\underline{X}^T V \underline{X})}{1^T \text{ diag } V} \; \frac{(\underline{X}^T Y \underline{U})^2}{(\underline{X}^T W \underline{X})^3} \left\{ 1 + \frac{(\underline{X}^T Z \underline{X})}{(\underline{X}^T W \underline{X})} \right\}$$

where $\left\{ \begin{aligned} Y &= \Omega(\Omega^T\Omega)^{-1}\Gamma \\ W &= \Omega(\Omega^T\Omega)^{-1}\Gamma^T\Gamma(\Omega^T\Omega)^{-1}\Omega^T \\ Z &= \Omega(\Omega^T\Omega)^{-1}\Gamma^T\Gamma(\Omega^T\Omega)^{-1}\Gamma^T\Gamma(\Omega^T\Omega)^{-1}\Omega^T \\ V &= I - \Omega(\Omega^T\Omega)^{-1}\Omega^T \end{aligned} \right.$