Geographical Analysis ISSN 0016-7363

# Leading Indicators and Spatial Interactions: A Crime-Forecasting Model for Proactive Police Deployment

Jacqueline Cohen[1], Wilpen L. Gorr[2], Andreas M. Olligschlaeger[3]

[1]H. John Heinz III School of Public Policy and Management, Carnegie Mellon University, Pittsburgh, PA, [2]Department of Public Policy and Management Information Systems, H. John Heinz III School of Public Policy and Management, Carnegie Mellon University, Pittsburgh, PA, [3]TruNorth Data Systems, Freedom, PA

*We develop a leading indicator model for forecasting serious property and violent crimes based on the crime attractor and displacement theories of environmental criminology. The model, intended for support of tactical deployment of police resources, is at the microlevel scale; namely, 1-month-ahead forecasts over a grid system of 141 square grid cells 4000 feet on a side (with approximately 100 blocks per grid cell). The leading indicators are selected lesser crimes and incivilities entering the model in two ways: (1) as time lags within grid cells and (2) time and space lags averaged over grid cells contiguous to observation grid cells. Our validation case study uses 1.3 million police records from Pittsburgh, Pennsylvania, aggregated over the grid system for a 96-month period ending in December 1998. The study uses the rolling-horizon forecast experimental design with forecasts made over the 36-month period ending in December 1998, yielding 5076 forecast errors per model. We estimated the leading indicator model using a robust linear regression model, a neural network, and a proven univariate, extrapolative forecast method for use as a benchmark in Granger causality testing. We find evidence of both the crime attractor and displacement theories. The results of comparative forecast experiments are that the leading indicator models provide acceptable forecasts that are significantly better than the extrapolative method in three out of four cases, and for the fourth there is a tie but poor forecast performance. The leading indicators find 41–53% of large crime volume changes in the three successful cases. The corresponding workload for police is quite acceptable, with on the average 5.2 potential large change cases per month to investigate and with 31% of such cases being positives.*

Correspondence: Wilpen L. Gorr, Department of Public Policy and Management Information Systems, H. John Heinz III School of Public Policy and Management, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213-3890
e-mail: wg0g@andrew.cmu.edu

Submitted: June 26, 2004. Revised version accepted: January 6, 2005.

## Introduction

Geography has become increasingly important in law enforcement and crime prevention. Criminology has long focused on individual propensities toward crime, but it was only during the last few decades that the criminogenic features of *settings* began to take on importance in research and practice. Environmental criminology gained in development, empirical verification, and practical applications by police (Cohen and Felson 1977; Brantingham and Brantingham 1981, 1984; Cornish and Clarke 1986; Eck and Weisburd 1995). Places, besides persons, became targets for allocation of police resources, and fields including crime mapping (Harries 1999), geographic profiling (Rossmo 2000), and (most recently) crime forecasting (Gorr and Harries 2003) arose in support of the new-found law enforcement opportunities.

This article introduces a leading-indicator crime-forecasting model for proactive policing and crime prevention, building on the work of Olligschlaeger (1997, 1998). Police, like other professionals delivering services, generally know the current locations and intensities of demand for services. Indeed, crime mapping based on near-real-time input of police reports has made the current picture for police more complete, integrating data from various officers, shifts, and neighborhoods. With the current situation in hand, the next step and most difficult new information to obtain is making forecasts of large changes in crime. If it were possible to obtain such forecasts, in the short term of up to a month ahead, then police could focus crime analysts' activities and build up intelligence on highlighted areas, target patrols, reallocate detective squads, and carry out other police interventions to prevent crimes.

Attempting to make accurate forecasts of the relatively rare, large changes in crime from month to month is an ambitious and difficult undertaking; however, the expectations of police can adapt to accepting good leads mixed in with false positives. For example, if 50% of forecasted large changes actually have large changes, then we claim that this would be an excellent result. Such forecasts would provide an entirely new kind of valuable information for police. Police practices already involve following up on many leads before success.

It is not difficult to obtain accurate extrapolative forecasts of crime. Gorr, Olligschlaeger, and Thompson (2003), using the same case study data as this article, demonstrated that exponential smoothing methods and classical decomposition yield accurate 1-month-ahead forecasts for areas that have average historical crime volumes in excess of 25–35 crimes per month. Unfortunately, these "business-as-usual" forecasts cannot foresee the largest changes in crime levels, namely, those involving breaks in time series patterns such as step jumps up or down.

The leading indicator model presented in this article is promising for forecasting breaks. If leading indicators experience a break from previous patterns during the last month of the estimation data set, they are capable of forecasting a similar break in the dependent crime variable in the next month. The present article develops and evaluates crime-based leading indicators and spatial interactions as a means to forecast breaks in serious crime levels. Another article (Gorr and McKay 2005), applying tracking signals to identify breaks in crime trends, finds that there are roughly two breaks every 3 years in high crime volume, square grid cells (4000 feet on a side) in Pittsburgh. Note that the leading indicator model of this article has not been used in practice by police as yet.

The next section provides a literature review and model specification. We draw on crime theories and police requirements to build our model. A case study of Pittsburgh, Pennsylvania, with 141 grid cell locations and 96 months of crime data is presented in the third section. This section includes an experimental design for validation of the leading indicator model, drawing on the forecasting literature. The results are given in the penultimate section, and the last section concludes the article.

## Model specification

Our leading indicator forecasts for serious crimes are based on a lag model for panel data. This section specifies the dependent variables, time-lagged leading indicator variables, and spatial interactions in the form of space and time lags. While multiple time lags are possible, our preliminary research indicated that a single time lag of independent variables was often the most accurate forecasting model. While lag models with up to four or more lags may ultimately prove to be the most accurate forecast models for leading indicators, we chose to keep the model in this first article on crime spatial interactions simple. Hence, we limit attention to a single time lag model in this article.

The choice of dependent variables depends on police requirements and data limitations. Municipal police in the United States have widely implemented a management by objectives approach known as CompStat, first developed by the New York City Police Department (Henry and Bratton 2002). CompStat focuses on reducing serious crimes, which, in the United States, consist of the index or part 1 crimes in the FBI's Uniform Crime-Reporting program: murder, rape, robbery, aggravated assault, burglary, larceny, and motor vehicle theft. Part 1 crimes are the dependent variables of our model; however, police requirements and data limitations both argue for using aggregates of part 1 crimes instead of individual crimes.

Police desire information on crime for the smallest geographic areas possible in order to precisely target patrols and investigative efforts. The smallest administrative unit of police departments is the patrol district or car beat, which is the territory of a single unit (usually a patrol car). Clearly, areas studied need to be the size of car beats or smaller. We use square grid cells 4000 feet on a side (yielding approximately 100 city blocks) in our case study of Pittsburgh, Pennsylvania. Grid cells have the advantage of easy visual interpretation on maps, given their uniform size and shape. During the time of study, Pittsburgh had 42 car beats and 141 grid cells.

We experimented with several grid cell sizes and found 4000 feet to be roughly the smallest possible for Pittsburgh while still yielding reasonable model estimates.

A necessary concession to working at this level is to sum all property crimes (robbery, burglary, larceny, and motor vehicle theft) to a single dependent variable, P1P, and similarly all violent crimes (murder, rape, and aggravated assault) to P1V for forecasting.[1] This aggregation is necessary to yield monthly crime time series with sufficient data volumes for accurate model estimation. For example, it is impossible to forecast murder at the grid cell level, with 40–60 murders per year in Pittsburgh.

Nevertheless, use of P1P and P1V as the dependent variables is compatible with the top-down analysis process used in CompStat in which participants need to make monthly, jurisdictionwide scans for crime problems to allocate limited analytical, investigative, and patrol resources. Leading indicator forecasts help make such a scan, with areas having a large forecasted increase in crime getting priority and perhaps those areas with opposite forecasts getting resources withdrawn. With such decisions made, crime analysts can ''drill down'' into selected areas for diagnosis and tactical-level planning of targeted patrols, assignment of detectives, and so on. It is in the second stage of crime analysis that information on individual part 1 and leading indicator crimes is needed.

Lesser crimes and incivilities, represented by selected part 2 crimes in police offense reports and citizen 911 complaints, constitute the leading indicators in our model. In general, these variables are suggested by two crime theories on spatial interactions: crime attractors and crime displacement, which we discuss later. It is fortunate for police (and perhaps unique) that they collect their own transactional data on leading indicators, thus enabling timely forecasts. Other, well-known economic indicators (Klein and Moore 1983) that are related to crime at national or regional levels (Deadman 2003; Harries 2003) change too slowly and are not available at the micro-geographic levels and time frames needed for tactical law enforcement within municipalities. Today, police have real-time information systems and can process and aggregate individual crime incidents to any desired variables to support forecasting.

Lesser crimes, like serious crimes, also receive intense enforcement because they too are costly to the public and because it is believed that they are precursors to serious crimes. The Broken Windows theory of crime (Wilson and Kelling 1982; Kelling and Coles 1996) argues that tolerance of minor incivilities and infractions of the law in neighborhoods are attractors to criminals, signaling settings conducive to a wide range of criminal activities. In addition, certain land uses and other physical features serve as attractors, for example, bars, parking lots, sporting events, and concerts. A major law of geography—the distance decay of attractions—suggests that criminals generally do not travel far to commit crimes (Capone and Nichols 1975) and hence would be attracted from nearby areas to a ''broken windows'' neighborhood.

The distance decay law of attractions has been incorporated into the pattern theory of crime (Brantingham and Brantingham 1984) and is the basis of geographic profiling (Rossmo 2000). A fundamental tenet of Broken Windows theory is that tolerated ''soft crimes'' harden later to serious crimes. This belief has led to ''zero tolerance'' enforcement of lesser crimes as a means of protecting neighborhoods from crimes of both the lesser and serious varieties. A reduction in the volume of lesser crimes is expected to lead to a similar reduction in serious crimes.

Even if the attractor theory is not at work, an additional argument applies to forecasts of serious violent crimes, P1V. Under 15% of all part 1 crimes are violent crimes, and the remainder are property crimes. P1P crimes are as numerous as leading indicator crimes, but P1V crimes are much less prevalent than many of their leading indicators. (See descriptive statistics for the dependent and independent variables in Table 1.) If a new source of crimes moves into a neighborhood for reasons other than the attraction of broken windows, those offenders bring with them all of their bad habits and multiple law-breaking practices for both lesser and serious crimes. Hence, by chance alone, because of their large differences in

**Table 1** Crime Leading Indicator and Dependent Variables

| Data type | Crime code | Variable name | Leading indicators | | Mean | Standard deviation | Maximum |
|---|---|---|---|---|---|---|---|
| | | | Property | Violent | | | |
| Citizen 911 call types | Domestic | C_DOMES | | X | 10.7 | 15.8 | 132 |
| | Drugs | C_DRUGS | X | X | 1.9 | 4.6 | 95 |
| | Public disorder | C_PUBLIC | | X | 6.2 | 8.3 | 75 |
| | Shots fired | C_SHOTS | | X | 2.5 | 5.4 | 66 |
| | Truancy | C_TRUAN | X | | 0.0 | 0.2 | 4 |
| | Vice | C_VICE | X | X | 0.3 | 1.5 | 41 |
| | Weapons | C_WEAPO | X | X | 2.5 | 4.4 | 53 |
| Offense crime types | Criminal mischief | CRIMIS | X | X | 5.1 | 6.5 | 50 |
| | Disorderly conduct | DISORD | X | X | 2.7 | 5.1 | 97 |
| | Liquor law violation | LIQUOR | X | X | 0.4 | 1.5 | 34 |
| | Prostitution | PROST | | X | 0.4 | 2.1 | 54 |
| | Public drunkeness | PUBDRUN | | X | 0.4 | 1.6 | 46 |
| | Simple assault | SIMPASS | | X | 6.5 | 9.6 | 82 |
| | Trespass | TRESPAS | X | X | 0.7 | 1.4 | 17 |
| Dependent variables | Part 1 property | P1P | | | 10.3 | 14.6 | 115 |
| | Part 1 violent | P1V | | | 1.7 | 3.3 | 31 |

volume, we would expect to see evidence of more frequent lesser crimes earlier than less frequent serious violent crimes.

An opposing effect to crime attractors is crime displacement. Police have long believed that increased enforcement in one location merely displaces criminal activity to other nearby locations (Eck 1993; Ratcliffe 2002). For example, concern about crime displacement was the basis for the large drug market analysis program (DMAP) of the U.S. National Institute of Justice in the early 1990s, which supported development of crime mapping in the United States and in which we participated. In that program, we saw much anecdotal evidence of drug dealing displacement in the Pittsburgh Police Bureau's DMAP geographic information system (GIS) we developed. Subsequent empirical research on crime displacement more generally suggests, however, that crime displacement is less prevalent than thought. Twenty-two out of 55 studies where crime displacement has been studied found no evidence of it at all (Hesseling 1994).

There are some difficulties in modeling crime displacement. A primary one is that police rarely record sufficient data on crackdowns and other special enforcement activities to allow for systematic modeling of the enforcement effect. Consequently, much of the evidence on displacement is anecdotal. Another difficulty concerns geographic scale. Displacement is likely a behavior that occurs over small distances, so that either it is unobserved within geographic observation units or the needed units are too small to yield sufficient data volumes needed for reliable model estimation. The model in this article has the advantage of using reported incidents of lesser crimes as a surrogate for police intervention measures, thereby providing data on displacement, as well as using relatively small observation units.

Without much more theory to draw on for identifying specific leading indicators of serious crimes, we decided to rely on expert judgment for selecting particular lesser crimes as leading indicators. Our first step was to compile a list of all lesser offense types and all codes characterizing complaints in citizen 911 calls for police services. We then asked police crime analysts in two cities to select leading indicators from this list. With the initial selection in hand, we then asked two criminologists to further refine and classify the list. The resulting final list of leading indicators for P1P and P1V crimes is in Table 1.

Offense data are based on incidents reported to or discovered by police and recorded as crimes in police offense reports. Underreported crimes (like sex crimes and assaults between family and friends) or victimless crimes (like illegal drug use or prostitution) are underrepresented in police offense data. Citizen 911 calls for service are more inclusive of complaints about criminal and public disorder activities, but are vulnerable to overestimates arising from untrained observers or complainants attempting to manipulate the system (e.g., claiming a more serious problem than existing to get a quick response by police).

Table 1 includes descriptive statistics for all variables. Note that 5 of the 14 leading indicators (C_TRUAN, C_VICE, LIQUOR, PROST, PUBDRUN, and TRES-

PAS) have low means of under one incident per month per grid cell, and high standard deviations and maximums. We retained these variables as leading indicators under the expectation that relatively high numbers of these measures are concentrated in a few areas and would be discriminating for those areas.

Application of the theories discussed and expert-based efforts on our part thus led to a leading indicator forecast model with P1P or P1V as dependent variables and the two sets of independent variables in Table 1. The analysis includes 1-month time lags for each leading indicator in the same grid cell and averages of each leading indicator in neighboring contiguous grid cells (queens case spatial lags) also lagged by 1 month.

If attractor theory is correct, we expect the signs of coefficients for time-lagged independent variables in the same grid cell to be positive, reflecting the direct effects of lesser crimes on subsequent serious crimes in the same location. Operating under the same attractor mechanism, we expect the coefficients of the spatial lags to be negative, with nearby high-activity grid cells drawing offenders and their criminal activity away from the local grid cell. In contrast, if displacement is the operant mechanism, and police actively target high levels of lesser crimes for enforcement, then we expect the coefficient signs of local time lags to be negative and those of the spatial lags to be positive as criminal activity moves away from higher levels of enforcement directed against lesser crimes. Displacement of otherwise nonviolent offending could have an immediate effect on violent crimes if displacement results in turf wars between already established and newly arriving displaced offenders.

If both attractor and displacement processes are at work, and these operate locally and between neighbors, the results may be net zero changes in local levels of serious crimes. Our estimates will only be able to detect the presence of significant nonzero net effects, and any estimated significant effects will thus represent net effects of the dominant process.

Estimation of our model includes robust linear regressions and a nonlinear neural network. Early in our research, we compared results from Poisson regressions suitable for crime counts, and found coefficient estimates to be similar to those from linear regressions, and so we use linear regressions for forecasting. We use STATA software to estimate robust linear regressions with observations clustered within grid cells to estimate the standard errors of coefficient estimates. These standard errors relax the usual ordinary least squares (OLS) assumptions of independent and identically distributed errors to yield consistent estimates for arbitrary variance/covariance error structure. The nonlinear neural network model (Olligschlaeger 1998), with a single middle layer and standard feed-forward estimation, provides an exploratory, self-adjusting mechanism to find additional patterns in the independent variables beyond the linear regression specification.

## Case study and validation approach

We collected approximately 1.3 million individual crime incident data records (crime offense reports and 911 calls for service) for Pittsburgh, Pennsylvania, over the period 1991 through 1998. We used a GIS to geocode the points, with overall address match rates of 90% for offense records and 80% for 911 call records. Overall, these rates are at the U.S. national average for police data, which is on the order of 85%. With data points and grid cells on a GIS map, we used spatial joins to assign grid cell identifiers to crime points, and then used database queries to create monthly series for each grid cell.

Our forecast validation study uses the rolling-horizon experimental design (e.g., Swanson and White 1997), which maximizes the number of forecasts for a given time series at different times and under different conditions. This design includes two or more alternative parallel forecasts. For each forecast included in an experiment, we estimate models on training data, forecast 1 month ahead to new data not previously seen by the model, and then calculate and save the forecast errors. Next, we roll forward 1 month, adding the observed value of the previously forecasted data point to the training data, dropping the oldest historical data point, and forecasting ahead to the next month. This process repeats until all data are exhausted.

The regression model uses a 3-year estimation window, the extrapolative method explained requires a 5-year estimation window, and neural network estimations start with the earliest 5 years of data and retain all historic data as the horizon rolls forward. The rolling 3-year window for regression estimation allows estimated parameters to vary over time, thus capturing effects of unmeasured factors such as changes in police policies or innovations in crime. The extrapolative univariate method needs at least 5 years of data to estimate seasonal effects. In the data sample used here, the earliest forecast origin is December 1995, retaining January 1991 through December 1995 for estimation. One-month-ahead forecasts are available for January 1996 through December 1998 for a total of 36 months times 141 grid cells to yield 5076 forecast errors per forecast method.

We used Granger causality testing (Granger 1969) to determine the relative value of leading indicators for serious crimes. A variable X Granger-causes Y, if Y can be better predicted using the histories of both X and Y than it can, using the history of Y alone. Our use of this concept for leading indicators determines whether they forecast serious crimes significantly better than the best univariate, extrapolative method, especially for large crime changes. To develop benchmark accuracy measures, we first optimized over univariate methods to obtain the most accurate extrapolative forecasts (Gorr, Olligschlaeger, and Thompson 2003).

The forecast literature generally uses central tendency of forecast error measures as the criterion for comparing alternative forecast models or assessing the value of forecasts. For a rolling-horizon experiment using panel data, let $Y_{it}$ = crime count in grid cell $i$ at time $t$ ($i = 1, \ldots, m$ and $t = 1, \ldots, T$), the dependent variable of estimation data panel, $T = T_1, \ldots, T_n$, forecast origins (last estimation data points) $F_{i,T+k}$ = forecasted crime, $k$ steps ahead (we restrict $k = 1$), $e_{i,T+k}$ =

$F_{i,T+k} - Y_{i,T+k}$ = forecast error, then example criteria are

$$\text{MSE}\,(K) = \sum\sum(e_{i,\,T+k})^2/(mn) = \text{mean squared error}$$
$$\text{MAPE}\,(K) = \sum\sum abs(e_{i,T+k}/Y_{i,\,T+k})/(mn) = \text{mean absolute percentage error}$$

We determined, however, that such measures are inappropriate for the police application at hand namely, detecting large changes in crime. Measures such as the mean standard error and mean absolute percent error assess forecast accuracy across all crime levels and do not directly assess change. In contrast, the decision requirement of police is on forecasted change versus actual change:

$$F\Delta_{i,\,T+k} = F_{i,\,T+k} - Y_T = \text{forecasted change} \tag{1}$$

$$A\Delta_{i,\,T+k} = Y_{i,\,T+k} - Y_T = \text{actual change} \tag{2}$$

Hence, we use a forecast error criterion that contrasts (1) and (2).

A common practice of crime analysts, and the basis of our forecast performance measure, is the use of threshold crime levels as triggers for exception reports for possible action. An example rule using a threshold level is as follows: if P1V crimes are forecasted to increase by more than six in any given grid cell, then that cell merits attention. Hence, rather than assessing accuracy based on the performance of individual point forecasts for each grid cell, we examined forecast performance within ranges of changes for both decreases and increases. Using contingency tables based on measures (1) and (2), we contrast forecasted and actual changes within each range and designate correctly triggered decisions as positives and negatives, and incorrect decisions as false negatives and false positives. We apply pairwise comparison *t* tests within classes to determine whether leading indicator forecasts are significantly better than univariate forecasts.

Before proceeding to the results of forecasting experiments, we address two issues regarding the contingency table analysis for forecasted change, namely, that of outliers and the related issue of forecasting large crime volume decreases. First, it is necessary to remove outliers from the contingency table analysis (they are retained in estimation of models). By outlier, we mean a crime count in 1 month that is unusually higher or lower than the months preceding and following it. If we were to include outliers in the assessment of forecast accuracy using (1) and (2), there would mostly be good performance in forecasting large decreases in crime levels, but this is a mere artifact of forecasting outliers in a moving horizon. Most outliers are high, yielding a large increase in crime level during the outlier month. Forecast models, of course, cannot forecast the outlier accurately and hence have poor accuracy for corresponding increases. The month following the outlier has a large decrease, a return to the long-term trend. Forecast models adjust minimally to the outlier and thus still forecast at levels corresponding to the long-term trend. Hence, by default, they accurately forecast the return of crime levels after an outlier. We identify and remove outliers from assessments of forecast accuracy to avoid such spurious results on forecasting large crime decreases.

We reject classical approaches to identifying outliers based on tolerance limits for two reasons. First, besides outliers, crime series data exhibit pattern changes such as step jumps. Tolerance limits incorrectly identify step jumps as outliers. Second, we desired a method of identifying outliers that match the decision rules used in this article. We thus decided to use ad hoc decision rules as follows:

*Property crime outlier rule*:

$$\text{if } (A\Delta_{i,\,T+k} \geq 15 \text{ and } A\Delta_{i,\,T} \leq -15) \text{ or}$$
$$(A\Delta_{i,\,T+k} \leq -15 \text{ and } A\Delta_{i,\,T} \geq 15) \text{ then } Y_T \text{ is an outlier}$$

This rule simply states that, if the monthly crime count changes in one direction by more than 15 and immediately reverses with the same change in the opposite direction, it is an outlier. The similar rule for violent crimes is as follows:

*Violent crime outlier rule*:

$$\text{if } (A\Delta_{i,\,T+k} \geq 6 \text{ and } A\Delta_{i,\,T} \leq -6) \text{ or}$$
$$(A\Delta_{i,\,T+k} \leq -6 \text{ and } A\Delta_{i,\,T} \geq 6) \text{ then } Y_T \text{ is an outlier}$$

For property crimes, there are 122 records with $|A\Delta_{i,\,T}| \geq 15$. Of these, there are 26 pairs of consecutive records satisfying the outlier identification criteria. Five of the pairs are in the central business district where crime volume is very high so that these are not outliers. Thus, there are 21 pairs of records that are outlier cases. Sixteen out of the 21 are high outliers. We dropped all 21 outliers from the contingency table analysis. For violent crimes, there are 119 records with $|A\Delta_{i,T}| \geq 6$. There are 42 pairs of consecutive records satisfying the outlier criteria; however, there are three grid cells with high crime volume and five or more pairs identified in each. We designate these as nonoutliers. Thus, in the end, the procedure identifies 24 outliers and of these 17 are high outliers. We dropped all 24 outliers from the contingency table analysis of violent crimes.

The second issue is similar to that of outliers. Our models have greater success in forecasting large crime decreases than large crime increases, as will be seen in the next section of this article. Of course, the latter are more important for prevention of crime and also, we believe, provide the true test of crime leading indicator forecast models. The following scenario describes the issue at hand. Suppose that a time series has been moving along a steady, long-term trend for a large number of months but then has a step jump increase. The majority of such step jumps in crime time series are increases, likely reflecting a new criminal element in a neighborhood. After a period of time, police recognize the increased activity, investigate it, and likely are successful in suppressing it, causing crime to decrease and return to the long-term trend.

The ability to forecast a step jump increase depends directly on the predictive ability of the leading indicators and corresponding model. Only the neural network model for violent part 1 crimes in Tables 7–9 is successful for this case. In contrast, the ability to forecast the return to the long-term trend can be independent of leading indicators. All that is necessary is to have a model that is nonreactive to change and thus persists in estimating and forecasting the long-range trend at every forecast origin. The regression model best fits this description and is also the best forecaster of crime decreases in Tables 4–9. Thus, the good results by regression analysis for forecasting crime decreases may simply be due to its inability to adjust quickly to pattern changes, and such cases may not be true tests of leading indicator models.

## Results

Tables 2 and 3 present sample regression estimates for P1P and P1V leading indicator models for the first 3-year data window (January 1993 through December 1995) and last window (January 1996 through December 1998) out of 36 sets of such regressions used for forecasting. All models displayed have relatively high $R^2$ values, in the range 0.69–0.79, and many significant coefficients. Out of the 42

**Table 2** Estimated Coefficients for Leading Indicator Forecast Model: Part 1 Property Crimes (P1P)

| Variable | Estimated coefficients | |
|---|---|---|
| | 1993–1995 | 1996–1998 |
| Intercept | − 1.10487* | − 0.63315 |
| C_DRUGS | 0.03421 | − 0.08362 |
| NC_DRUGS | 0.06626 | 0.13656 |
| C_TRUAN | 0.21061 | 1.62792*** |
| NC_TRUAN | − 0.34338 | − 0.59283 |
| C_VICE | − 0.99699*** | 0.12639 |
| NC_VICE | − 0.45854 | − 0.46786 |
| CRIMIS | 1.12172*** | 0.74551*** |
| NCRIMIS | 0.50151*** | 0.44472** |
| DISORD | 0.74668* | 1.04442*** |
| NDISORD | − 0.13513 | − 0.25105 |
| LIQUOR | 0.25334 | 0.82613*** |
| NLIQUOR | − 0.01709 | 0.26481 |
| TRESPAS | 1.37588*** | 1.05194*** |
| NTRESPAS | 0.53914 | 0.57485 |
| WEAPO | 0.69690 | 0.94968*** |
| NWEAPO | 0.72023 | 2.18272*** |

Note: $N = 5076$ in each time period. $R^2 = 0.79$ for 1993–1995; $R^2 = 0.76$ for 1996–1998. Two-tailed significance levels using robust estimates of standard error that account for non-independent clustering of observations over time in the same grid cell: $^*P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$. Variable names starting with N and shaded are space and time lags; other variables are simple time lags.

**Table 3** Estimated Coefficients for Leading Indicator Forecast Model: Part 1 Violent Crimes (P1V)

| Variable | Estimated coefficients | |
|---|---|---|
| | 1993–1995 | 1996–1998 |
| Intercept | − 0.24545*** | − 023292*** |
| C_DOMES | 0.000973 | 0.01672 |
| NC_DOMES | 0.00073 | 0.01346 |
| C_DRUGS | 0.08572*** | 0.05347*** |
| NC_DRUGS | − 0.04740 | 0.00549 |
| C_PUBLIC | − 0.00298 | 0.00239 |
| NC_PUBLIC | − 0.07115*** | − 0.03823 |
| C_SHOTS | 0.06798*** | 0.08168* |
| NC_SHOTS | 0.00042 | − 0.01492 |
| C_VICE | − 0.00811 | 0.02605 |
| NC_VICE | 0.09459 | − 0.10392 |
| C_WEAPO | 0.03947 | 0.05960** |
| NC_WEAPO | − 0.04256 | 0.00860 |
| CRIMIS | 0.04894** | 0.04291*** |
| NCRIMIS | 0.02641 | 0.01478 |
| DISORD | 0.04835* | 0.09128** |
| NDISORD | 0.00315 | 0.01797 |
| LIQUOR | 0.00875 | 0.08745 |
| NLIQUOR | − 0.02362 | 0.02496 |
| PROST | 0.10139** | 0.12945*** |
| NPROST | − 0.11611 | 0.01516 |
| PUBDRUN | 0.20295*** | 0.06443 |
| NPUBDRUN | − 0.03453 | 0.08649 |
| SIMPASS | 0.12191*** | 0.07472** |
| NSIMPASS | 0.07719** | 0.01037 |
| TRESPAS | 0.04523 | 0.10623** |
| NTRESPAS | 0.09618 | − 0.02356 |

Note: $N = 5076$ in each time period. $R^2 = 0.76$ for 1993–1995; $R^2 = 0.69$ for 1996–1998. Two-tailed significance levels using robust estimates of standard error that account for non-independent clustering of observations over time in the same grid cell: $^{*}P<0.05$,$^{**}P<0.01$,$^{***}P<0.001$. Variable names starting with N and shaded are space and time lags; other variables are simple time lags.

estimated coefficients for time-lagged leading variables (not the time- and space-lagged variables) in Tables 2 and 3, 25 are significant at traditional levels and only one of those is negative. All but three are significant at the 0.01 level or better. Thus, there is some evidence that the proposed leading indicators do lead serious crimes, although comparisons with extrapolative models have stronger evidence of this.

Figs. 1 and 2 provide time series plots of robust-regression estimated parameters of crime variables that have both space and time lags. These plots have the

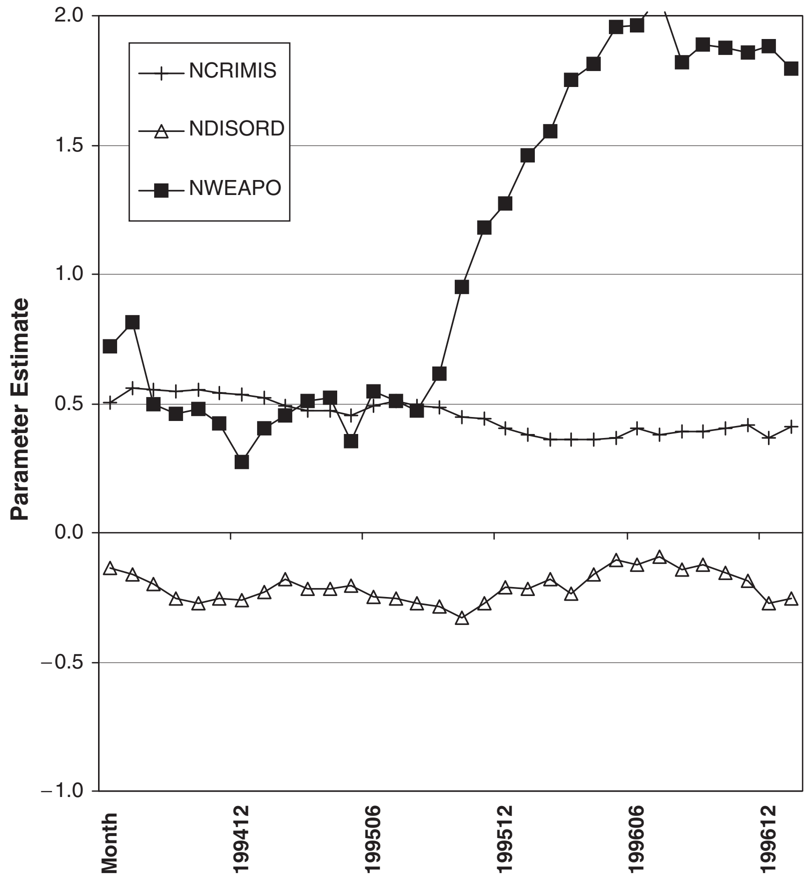


**Figure 1**. Estimated parameter paths from moving 3-year data window: space- and time-lagged variables, part 1 property crime (P1P).

purpose of examining attractor (negative coefficients) versus displacement (positive coefficients) behavior. Figure 1 has parameter paths predictive of property crimes that have at least one significant coefficient in Table 2 at the 0.10 level or better. (Table 2 denotes significance levels down to only the 0.05 level with two space- and time-lagged variable coefficients significant at this level or better. The 0.10 level admits one more coefficient, for NDISORD.) Figure 2 for violent crimes has similarly identified parameter paths for variables that are significant at the 0.10 level in Table 3. We plotted each estimate at the center of its data window, thus providing estimates of conditions at correct times on the horizontal scale.

In Fig. 1, for P1P, coefficients had time parameter paths that remained roughly stable, except for weapons offenses. Criminal mischief and weapons violations have displacement behavior while disorderly conduct has attractor behavior. The weapons coefficient started positive and then rapidly and markedly increased, more than doubling in the latter months of 1995 as a displacement factor. This

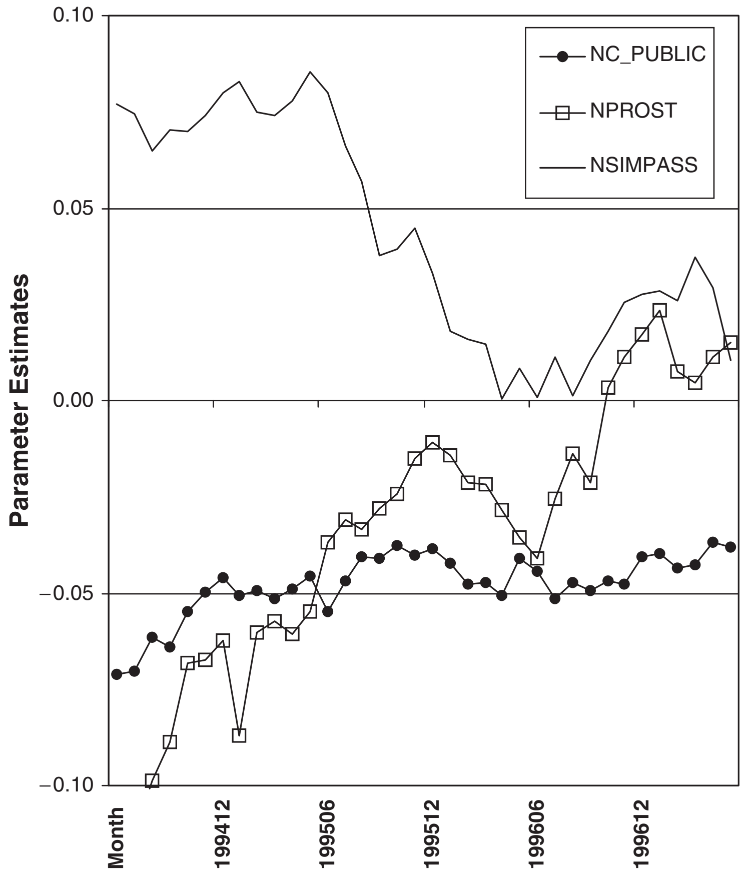


**Figure 2**. Estimated parameter paths from moving 3-year data window: space- and time-lagged variables, part 1 violent crime (P1V).

corresponds to a period during which the Pittsburgh Bureau of Police started enforcing gun laws aggressively and so perhaps explains the increasing displacement effect. Disorderly conduct (NDISORD in Fig. 1) has crime attractor behavior, because of its negative coefficient. This is sensible because disorderly conduct is the most visible of the three crimes, perhaps signaling deteriorating conditions.

The patterns in Fig. 2 for part 1 violent crimes are quite different than those of Fig. 1. The parameter path for 911 calls on public disturbances remained roughly constant, around $-0.05$ as a crime attractor. Estimated parameter paths for simple assaults and prostitution, which were significantly positive or negative in the beginning, deteriorated over time with parameter paths approaching zero. In general, crime levels were declining during this time period, but we have no explanation for the decline of significance in these leading indicators. Prostitution and public disturbances are good candidates as crime attractors because of their visibility in neighborhoods. If simple assaults were associated with gang violence,

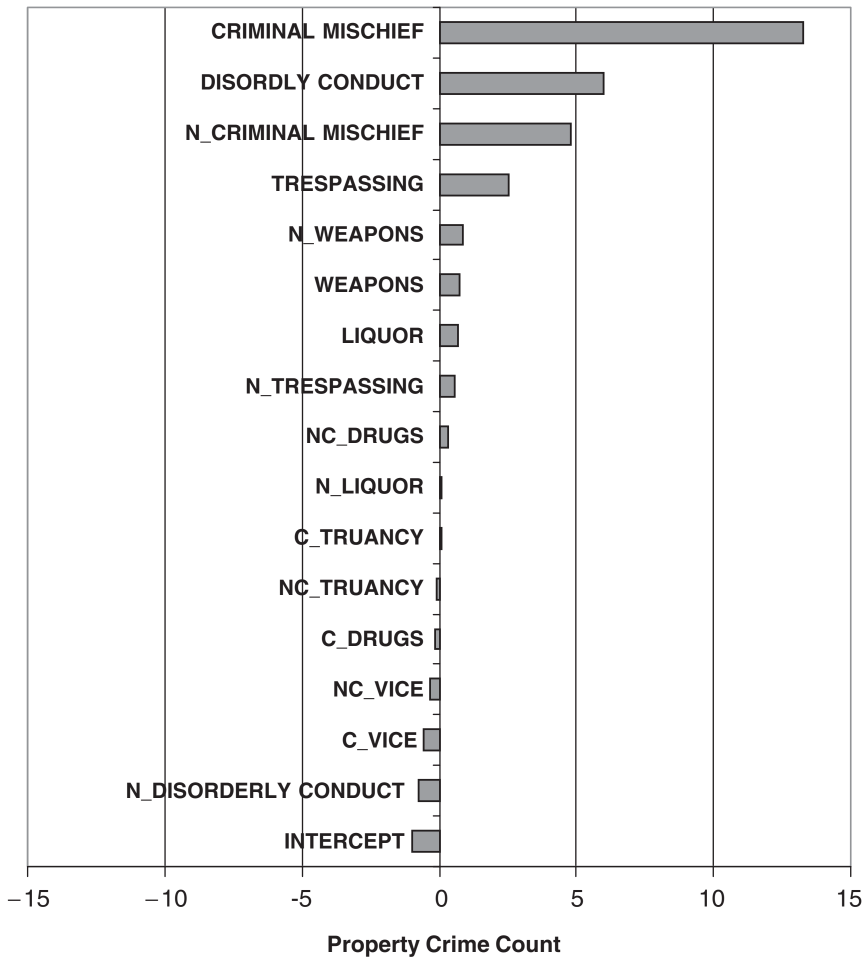


**Figure 3**. Average term contributions: part 1 property crime leading indicator regression model (based on average indicators for grid months with 10 or more property crimes).

then these crimes are good candidates for displacement effects either because of increased police attention or retaliation on the opposing gang's turf.

Figs. 3 and 4 provide another assessment of leading indicators; their practical significance in predicting volume of serious crimes. In this case, we estimated the model by OLS regression over the entire study period for regression analysis of 1993–1998. Each of the bar charts in these figures was obtained by averaging the leading indicators across ''active'' grid cells, defined to be cells with average dependent variable crime counts of 10 or more for property crimes and 6 or more for violent crimes. Then, we multiplied the averaged leading indicators by estimated regression coefficients, with the results displayed as bar charts indicating the average contribution of each term. For example, Fig. 3 shows that criminal

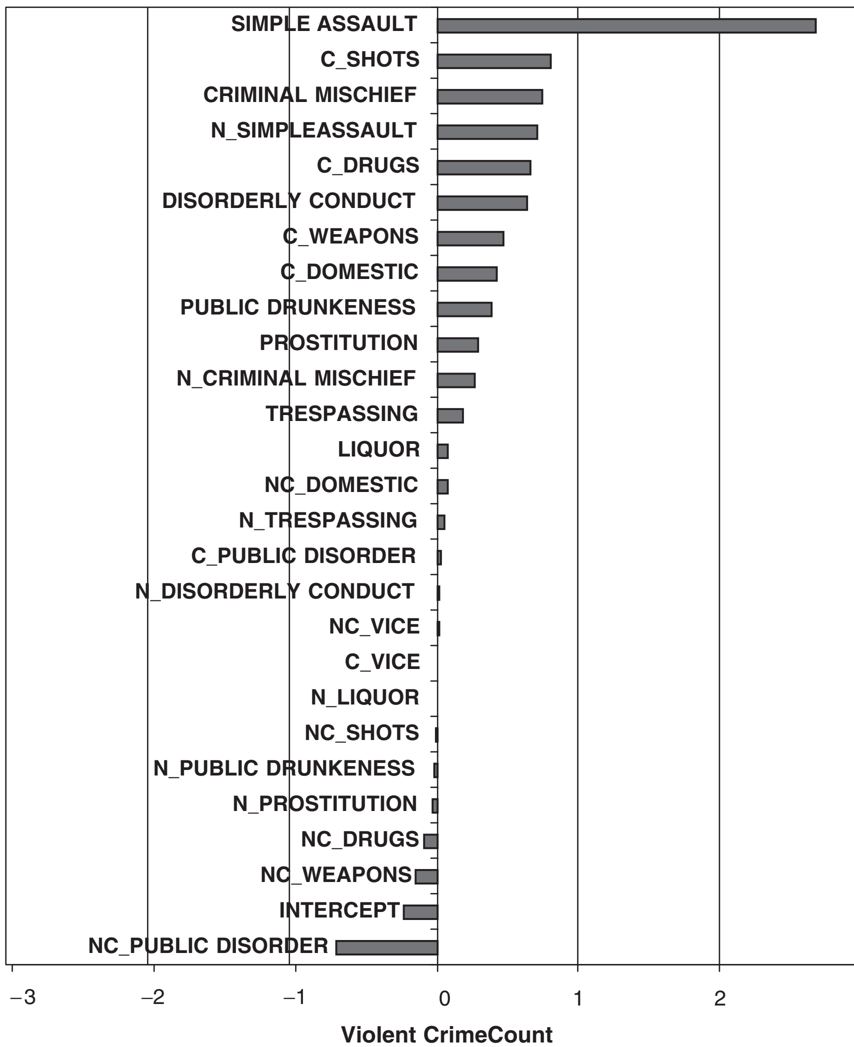


**Figure 4**. Average term contributions: part 1 violent crime leading indicator regression model (based on average indicators for grid months with five or more violent crimes).

mischief typically is correlated with about 13 part 1 property crimes and in Fig. 4 that simple assaults is correlated with nearly three part 1 violent crimes.

There are relatively few practical leading indicators for part 1 property crimes (Fig. 3). Criminal mischief has the largest impact, with disorderly conduct next, followed by criminal mischief in neighboring grid cells, and then trespassing. For part 1 violent crimes (Fig. 4), simple assaults in the same grid cell dominate; however, a number of other leading indicators contribute practically including citizen calls for shots fired, criminal mischief, simple assaults in neighboring grid

**Table 4** Percentage of Positives and False Negatives for Large Change Actuals: Part 1 Property Crimes

| | Forecasted change | Univariate | Regression | Neural network |
|---|---|---|---|---|
| Actual change is 15 or greater decrease (45 cases) | | | | |
| Positive | ≥ 15 decrease | 24 | 53 | 20 |
| False negative | 0–14 decrease | 64 | 36 | 71 |
| False negative | 0–14 increase | 11 | 11 | 9 |
| False negative | ≥ 15 increase | 0 | 0 | 0 |
| | Total | 100 | 100 | 100 |
| Actual change is 15 or greater increase (38 cases) | | | | |
| False negative | ≥ 15 decrease | 0 | 5 | 0 |
| False negative | 0–14 decrease | 26 | 34 | 21 |
| False negative | 0–14 increase | 61 | 58 | 74 |
| Positive | ≥ 15 increase | 13 | 3 | 5 |
| | Total | 100 | 100 | 100 |

cells, citizen drug calls, disorderly conduct, and citizen weapons calls. On the negative, attractor side, public disorder citizen calls have the largest impact.

Tables 4–9 present results comparing the robust regression model, nonlinear neural network model, and univariate forecasting method. The univariate method is the best as determined by Gorr, Olligschlaeger, and Thompson (2003): Holt exponential smoothing including a time trend with seasonality estimated using classical decomposition applied to pooled, city wide data for seasonality.

First are forecast error comparisons for part 1 property crimes, using change and error measures (1) and (2). Table 4 displays relative frequencies for cases in which the true change from the last historical data point to the forecast period was

**Table 5** Part 1 Property Crimes: Pairwise Comparisons Test Results

| Actual change | Crime level | | Mean forecast for time $T+1$ | | | Mean absolute forecast error | | | $N$ |
|---|---|---|---|---|---|---|---|---|---|
| | Time $T$ | Time $T+1$ | Uni-variate | Regression | Neural network | Uni-variate | Regression | Neural network | |
| ≥ 15 decrease | 46.9 | 26.4 | 35.1 | 30.4 | 37.6 | 10.5 | 8.4* | 11.9 | 65 |
| 0–14 decrease | 7.8 | 5.8 | 7.0 | 7.4 | 8.0 | 1.8* | 3.0 | 2.6 | 3011 |
| 0–14 increase | 9.0 | 12.9 | 10.1 | 10.7 | 10.7 | 3.4 | 4.5 | 3.1* | 1671 |
| ≥ 15 increase | 27.6 | 50.0 | 33.6 | 29.4 | 32.5 | 16.5* | 20.6 | 17.4* | 47 |

*Most accurate forecast or not significantly worse than most accurate forecast, 5% or better significance test.

**Table 6** Number of Positives and False Positives for Large Change Forecasts: Part 1 Property Crimes

| | Actual change | Univariate | Regression | Neural network |
|---|---|---|---|---|
| Forecasted change is 15 or greater decrease (45 actual cases) | | | | |
| Positive | ≥ 15 decrease | 11 | 24 | 9 |
| False positive | 0–14 decrease | 7 | 56 | 0 |
| False positive | 0–14 increase | 2 | 27 | 0 |
| False positive | ≥ 15 increase | 0 | 2 | 0 |
| | Total | 20 | 109 | 9 |
| Forecasted change is 15 or greater increase (38 actual cases) | | | | |
| False positive | ≥ 15 decrease | 0 | 0 | 0 |
| False positive | 0–14 decrease | 2 | 15 | 0 |
| False positive | 0–14 increase | 3 | 26 | 1 |
| Positive | ≥ 15 increase | 5 | 1 | 2 |
| | Total | 10 | 42 | 3 |

large: 15 or more decrease in the top panel and 15 or more increase in the lower panel. Rows labeled ''Positive'' are percentages of correct forecasts, for example, the percentage of forecasts of a decrease of 15 or more part 1 property crimes that were actually in that range. The regression model was most accurate for large decreases, finding 53% of actual such cases. The univariate method and neural network ran a distance second and third at 24% and 20%, respectively. For large increases, the more important case, all three models failed, with the highest percentage of such cases found being 13%.

Table 5 has additional information and significance tests for differences by forecasted change category. The first two columns under crime level indicate the

**Table 7** Percentage of Positives and False Negatives for Large Change Actuals: Part 1 Violent Crimes

| | Forecasted change | Univariate | Regression | Neural network |
|---|---|---|---|---|
| Actual change is 6 or greater decrease (41 cases) | | | | |
| Positive | ≥ 6 decrease | 22 | 41 | 12 |
| False negative | 0–5 decrease | 73 | 53 | 81 |
| False negative | 0–5 increase | 5 | 5 | 7 |
| False negative | ≥ 6 increase | 0 | 0 | 0 |
| | Total | 100 | 100 | 100 |
| Actual change is 6 or greater increase (37 cases) | | | | |
| False negative | ≥ 6 decrease | 0 | 0 | 0 |
| False negative | 0–5 decrease | 16 | 24 | 3 |
| False negative | 0–5 increase | 81 | 70 | 51 |
| Positive | ≥ 6 increase | 3 | 5 | 46 |
| | Total | 100 | 100 | 100 |

**Table 8** Part 1 Violent Crimes: Pairwise Comparisons of Forecast Errors

| Actual change | Mean crime level | | Mean forecast for time $T$+1 | | | Mean absolute forecast error | | | $N$ |
|---|---|---|---|---|---|---|---|---|---|
| | Time $T$ | Time $T$+1 | Univariate | Regression | Neural network | Univariate | Regression | Neural network | |
| ≥ 6 decrease | 12.5 | 4.7 | 8.3 | 6.9 | 9.1 | 3.9 | 2.9* | 4.4 | 61 |
| 0–5 decrease | 1.1 | 0.6 | 0.9 | 1.0 | 1.3 | 0.5* | 0.6 | 0.8 | 3601 |
| 0–5 increase | 1.5 | 3.4 | 2.1 | 2.3 | 2.7 | 1.5 | 1.4 | 1.3* | 1095 |
| ≥ 6 increase | 5.7 | 14.3 | 7.5 | 7.3 | 10.3 | 6.8 | 7.0 | 4.3* | 37 |

*Most accurate forecast based on paired difference test that contrasts each forecast method to the most accurate method at $P \leq 0.05$ significance level.

average crime levels in the last month of the estimation data set and in the month in the forecast period, so the reader can see the magnitudes of changes. The third through fifth columns under Mean forecast for time $T$+1 are average forecasts corresponding to the average actual crime levels in $T$+1. All forecast methods are biased to the mean, as expected, with forecasts for decreases too high on the average and the opposite for increases. The remaining columns display the MAPE within each forecasted change category, and include results of a pairwise comparison of the best method with the remaining methods in each row. The tests show that regression is significantly the best model for large decreases, as is the univariate method for smaller decreases, and the neural network for all increases. The last

**Table 9** Number of Positives and False Positives for Large Change Forecasts: Part 1 Violent Crimes

| | Actual change | Univariate | Regression | Neural network |
|---|---|---|---|---|
| Forecasted change is 6 or greater decrease (41 cases) | | | | |
| Positive | ≥ 6 decrease | 9 | 17 | 5 |
| False positive | 0–5 decrease | 3 | 10 | 3 |
| False positive | 0–5 increase | 1 | 4 | 0 |
| False positive | ≥ 6 increase | 0 | 0 | 0 |
| | Total | 12 | 31 | 8 |
| Forecasted change is 6 or greater increase (37 cases) | | | | |
| False positive | ≥ 6 decrease | 0 | 0 | 0 |
| False positive | 0–5 decrease | 1 | 0 | 6 |
| False positive | 0–5 increase | 6 | 3 | 24 |
| Positive | ≥ 6 increase | 1 | 2 | 17 |
| | Total | 8 | 5 | 47 |

column has the number of data points in each category. These sample sizes reflect the loss of two data points in each of the 141 grid cells from differencing data.

Finally, for part 1 property crimes, Table 6 provides information on the number of positives and false positives. The top panel has information for cases in which the forecast was for a decrease of 15 or more part 1 property crimes. The best method, the regression model, signaled a total of 109 possible such decreases, of which 24 (22%) were correct and 56 more were actually smaller decreases. So the total number of decreases forecasted was 80 (73%). The neural network was never wrong; all nine of the cases that it predicted were actually large decreases. The univariate method had a low number of cases triggered also, with 11 (55%) out of 20 correct and only two (10%) being increased. The positives in the lower panel are unremarkable.

Next are forecast error comparisons for part 1 violent crimes. Table 7 displays the relative frequencies for cases in which the true change from the last historical data point to the 1-month-ahead forecast was large: 6 or more decrease in the top panel and 6 or more increase in the lower panel. For large decreases, the regression model was again the best, finding 41% of the positives compared with 24% for the univariate method and 12% for the neural network. In the second panel, for the important case of large increases of 6 or more, the neural network is by far the best, finding 46% of such actual cases. The univariate and regression methods found only 3% and 5%, respectively.

This latter result for the neural network is remarkable. Apparently, the neural network's ability to self-recognize and estimate nonlinear patterns paid off in yielding an effective crime-forecasting model for large P1V increases. One explanation is that the model for violent crimes is more successful than that of property crimes; it has more leading indicators and more significant ones. Thus, the neural network had a rich set of independent variables for estimating violent crimes. In examining individual forecasts made by our models, it is apparent that the neural network is more capable of making large change forecasts, whereas the regression model and univariate model have forecasts closer to the mean crime level. In work being carried out at the time of writing this article, preliminary results using additional data from Pittsburgh and data from Rochester, New York, are yielding a similar pattern. The neural network is more accurate for forecasting P1V increases than a regression model or a univariate method.

In Table 8, skipping to the significance tests, the regression model was significantly the best for large decreases, the univariate test so for smaller decreases, and the neural network for all increases. Lastly, in Table 9 for numbers of positives and false positives, the regression model signals 31 large decrease cases, with 17 correct and another 10 being smaller decreases and only four small increases. Again, the neural network signals a few cases, only eight. The univariate method is similar. In the lower panel, the neural network now has the most forecasted large changes, 47, of which 17 are correct and another 24 are smaller increases. So the neural network signals have 41 increases (87%).

In summary of these results, the leading indicator models are significantly better than the univariate method in three out of four large change cases: large decreases for P1P and P1V and large increases for P1V. The neural network and univariate methods tie as best for P1P large increases, but performance in identifying large increases is unacceptably low in that case.

## Discussion and conclusion

This article has yielded theoretical, empirical, and practical results on leading indicator forecast models for serious crimes. At the theoretical level, we drew on environmental crime theories to determine leading indicators for serious crimes and interpret signs of estimated coefficients. If coefficients for space- and time-lagged independent variables (selected lesser crimes and incivilities) are negative, the variables correspond to attractors, drawing crimes away from an observation area. Otherwise, positive coefficients correspond to crime displacement from nearby areas to the observation area. Coefficients of time-lagged independent variables are all expected to be positive, reflecting crime attraction and leading behaviors.

The design of the leading indicator forecast model and its empirical tests are intended to provide information needed by police for deploying resources to prevent crime increases (or to retract resources from areas forecasted to have large crime decreases). The empirical tests use a state-of-the-art, rolling-horizon forecast experimental design with the innovation that forecasts are evaluated in the context of threshold decision rules. Corresponding tests use a contingency table analysis where multiple ranges of forecasted changes are examined and tested using paired comparisons.

The results are promising. Estimated models have coefficients with expected positive signs for time-lagged independent variables and a mixture of positive and negative coefficients for time- and space-lagged independent variables, reflecting crime attractor and displacement crime theories. The end result of forecast validations is that the leading indicator models produce useful forecasts that are significantly better than the extrapolative method in three out of four cases and for the fourth there is a tie but overall poor performance. Hence, the leading indicator model passes Granger causality testing. The regression model is best for forecasting large crime decreases of both crime types, but the neural network is best at forecasting large increases in violent crimes, all by wide margins.

In summary of Tables 4–9, the leading indicators yield a total of 58 positives out of 187 forecasts of large changes in the three successful cases over 36 months. This corresponds to a workload averaging only 5.2 forecasts of large changes per month, of which 1.6 are positives and 3.6 are false positives, spread over 141 grid cells in Pittsburgh. Of the 3.6 false positives per month, however, 2.5 have actual changes in the forecasted direction (increase or decrease) but not in the large crime change range, and thus are not totally without merit. Thus, the workload seems acceptable and potential net benefit positive.

Our current work on crime leading indicators is focusing on a replication study in a second city, Rochester, New York, in which we are reexamining the assumptions and results of the current article. We are modifying our neural network estimation procedure to increase its ability to estimate and forecast large crime changes.

## Acknowledgements

This project was supported by Grants No. 98-IJ-CX-K005 and 2001-CX-0018 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are ours and do not necessarily represent the U.S. Department of Justice. We are grateful for the assistance of Commander Kathleen McNeely of the Pittsburgh Police Bureau and John Shultie of Pittsburgh City Information Systems for providing crime data. We are also indebted to the anonymous referees of this article for their valuable suggestions.

## Note

1 Standard crime reporting by police to the FBI's Uniform Crime Reporting program includes robbery as a violent offense because of the risk of injury it poses for victims. We include robbery with property offenses because it shares many features (e.g., offender attributes, crime location, and time of day) with other offenses involving theft.

## References

Brantingham, P. J., and P. L. Brantingham, (Eds.). (1981). *Environmental Criminology*. Beverly Hills, CA: Sage.

Brantingham, P. J., and P. L. Brantingham. (1984). *Patterns in Crime*. New York: Macmillan.

Capone, D. L., and W. W. Nichols Jr. (1975). ''Crime and Distance: An Analysis of Offender Behavior in Space.'' *Proceedings of the Association of American Geographers* 7, 45–49.

Cohen, L., and M. Felson. (1977). ''Social Change and Crime Rate Trends: A Routine Activities Approach.'' *American Sociological Review* 44, 588–608.

Cornish., D. B., and R. V. Clarke, (Eds.). (1986). *The Reasoning Criminal: Rational Choice Perspectives on Offending*. New York: Springer-Verlag.

Deadman, D. (2003). ''Forecasting Residential Burglary.'' *International Journal of Forecasting (Special Section on Crime Forecasting)* 19, 567–78.

Eck, J. E. (1993). ''The Threat of Displacement, Problem Solving Quarterly.'' *Police Executive Research Forum* 6, 1–2.

Eck, J. E., and D. Weisburd, (Eds.). (1995). *Crime and Place. Crime and Prevention Studies,*, Vol. 4. Monsey, NY: Criminal Justice Press.

Gorr, W. L., and R. Harries. (2003). ''Introduction to Crime Forecasting.'' *International Journal of Forecasting (Special Section on Crime Forecasting)* 19, 551–55.

Gorr, W. L., and S. A. McKay. (2005). ''Application of Tracking Signals to Detect Time Series Pattern Changes in Crime Mapping Systems,.'' In *Geographic Information Systems and Crime Analysis*, edited by F. Wang. Hershey, PA: Idea Group Publishing.

Gorr, W. L., A. Olligschlaeger, and Y. Thompson. (2003). ''Short-Term Forecasting of Crime.'' *International Journal of Forecasting (Special Section on Crime Forecasting)* 19, 579–94.

Granger, E. S. (1969). ''Investigating Causal Relationships by Econometric Models and Cross-Spectral Models.'' *Econometrica* 37, 424–38.

Harries, K. (1999). *Mapping Crime Principle and Practice*. Washington, DC: U.S. Department of Justice Office of Justice Programs.

Harries, R. (2003). ''Modelling and Predicting Recorded Property Crime Trends in England and Wales—A Retrospective.'' *International Journal of Forecasting (Special Section on Crime Forecasting)* 19, 557–66.

Hesseling, R. (1994). ''Displacement: A Review of the Empirical Literature.'' In *Crime Prevention Studies*, Vol. 3: 197–230, edited by R. Clarke. Monsey, NY: Criminal Justice Press.

Henry, V. E., and W. J. Bratton. (2002). *The CompStat Paradigm: Management Accountability in Policing, Business and The Public Sector*. Flushing, NY: Looseleaf Law Publications.

Kelling, G. L., and C. M. Coles. (1996). *Fixing Broken Windows: Restoring Order and Reducing Crime in Our Communities*. New York: Free Press.

Klein, A. K., and G. H. Moore. (1983). ''The Leading Indicator Approach to Economic Forecasting—Retrospective and Prospect.'' *Journal of Forecasting* 2, 119–35.

Olligschlaeger, A. M. (1997). ''Spatial Analysis of Crime Using GIS-Based Data: Weighted Spatial Adaptive Filtering and Chaotic Cellular Forecasting with Applications to Street Level Drug Markets.'' Unpublished dissertation, Carnegie Mellon University, Pittsburgh, PA.

Olligschlaeger, A. M. (1998). ''Artificial Neural Networks and Crime Mapping.'' In *Crime Mapping Crime Prevention*, 313–47, edited by D. Weisburd and T. McEwen. Monsey, NY: Criminal Justice Press.

Ratcliffe, J. (2002). ''Burglary Reduction and the Myth of Displacement.'' Trends and Issues in Crime and Justice, No. 232, Australian Institute of Criminology.

Rossmo, K. (2000). *Geographic Profiling*. Boca Raton, FL: CRC Press.

Swanson, N. R., and H. White. (1997). ''Forecasting Economic Time Series Using Flexible Versus Fixed Specification and Linear Versus Nonlinear Econometric Models.'' *International Journal of Forecasting* 13, 439–61.

Wilson, J. Q., and G. L. Kelling. (1982). ''Broken Windows: The Police and Neighborhood Safety.'' *Atlantic Monthly* 249, 29–38.