Assessment of Crime Forecasting Accuracy for Deployment of Police[1]

By

Wilpen Gorr
Andreas Olligschlaeger
Yvonne Thompson

June 13, 2000

H. John Heinz III School of Public Policy and Management
Carnegie Mellon University

[1] Work on this paper was supported by Grant No. 98-IJ-CX-K005 awarded by the National Institute of Justice, Office of Justice programs, U.S. Department of Justice. Points of view in this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice.

Abstract

Crime forecasting is a new area of research, following upon the success of crime mapping for support of tactical deployment of police resources. The major question investigated in this paper is whether it is possible to accurately forecast crime one month ahead at a "small-scale" aggregation, i.e., at the precinct level. In a case study of Pittsburgh, Pennsylvania, we contrast the forecast accuracy of standard, univariate time series models with non-modeling practices commonly used by police. Included is a comparison of seasonality estimates made by precinct versus the city as a whole. As suspected for the small-scale data of this problem, average crime count by precinct and crime type is the major determinant of forecast accuracy. A fixed effects regression model of absolute percent forecast errors shows that such counts need to be on the order of 30 or more to achieve accuracy of 20 percent error or less. A second major result is that practically any model-based forecasting approach is vastly more accurate than current police practices. Thirdly, this is the first empirical paper to investigate crime seasonality at the sub-city level. Our seasonality estimates provide evidence supporting the routine activities theory of crime, but not earlier theories.

1.0 Introduction

With the emergence of geographic information systems (GIS) technology, and its successful use as a tool for crime analysis, the study and modeling of historical or current crime data to identify spatial crime patterns has emerged as a new research area in recent years (see Anselin 1988; Glaeser, Sacerdote and Scheinkman 1996, Zahn and Jamieson 1997, Engberg 1999, Cohen, Cork, Engberg and Tita 1998). *Crime forecasting*, however, is a new field, just now being investigated. Key questions are 1) using univariate methods, can crime patterns be extrapolated in the short-run in small areas such as precincts accurately enough for use in deployment, scheduling, and as counterfactuals for evaluating police effectiveness, and 2) are there leading indicators of crime for use in multivariate forecast models that can predict turning points and other new patterns, such as the start of a new crime "hotspot" (Wilson and Kelling 1982, Sherman, Gartin and Buerger 1989, Wilson and Coles 1996, Jeffries et al. 1998, Le Beau 1998). In this paper we address the first question, and will deal with the second question in future research. Our current work, besides having practical applications for police, also establishes a benchmark against which to measure performance of leading indicator and other more sophisticated forecast methods.

Seasonality is thought to be a major component of crime series data. We have reviewed 111 papers from the criminology and social psychology literature on the seasonality of crime. This paper is, however, the first to examine postulated models on seasonal patterns at the sub-city level and to test them based on forecast accuracy. Past empirical research has been limited due to spatial heterogeneity (aggregation bias) because the geographic unit of analysis used has been at the city or multi-city level. For some crime types, aggregation may mask important seasonal variations in smaller ecological areas. For example, university areas may have seasonal patterns in property crimes influenced by the comings and goings of the student body. Also, shopping areas may have peaks at holiday times.

We study monthly crime data by precinct in Pittsburgh, Pennsylvania over the period 1991 through 1998. Using holdout samples and a rolling horizon experimental design, we compare forecast accuracy of current police practices (e.g., use of last June's data point as the coming June's forecast) to standard univariate methods. We analyze results using significance testing of differences in forecast accuracy, and a regression model for testing the determinants of forecast error.

The major limitation on crime forecast accuracy in our study is that we must estimate forecast models using small-scale data series. For example, monthly robbery counts by precinct are only in the order of 10 to 30, and subject to much randomness. Corresponding model estimates are thus imprecise, leading to poor forecast accuracy. Moreover, seasonal factors are especially difficult to estimate accurately. The seasonal effect; for example, of June is observed only once per year. To combat this problem we compare forecast models using seasonal factors estimated for each precinct versus pooled estimates using city-level data. Duncan, Gorr and Szczypula (1993, 1999) and Bunn (1999) found that for small aggregations mostly effected by randomness, pooling results in more reliable parameter estimates and more accurate forecasts.

2.0 Forecasting Requirements of Police

Forecasting and decision problems are often classified by the length of the planning horizon: short-term (tactical deployment), medium-term (resource allocation), and long-term (strategic planning). Our research is on short-term forecasting and tactical decision-making. There are several reasons for this choice: 1) our research objective is to extend crime mapping through analytical tools, and most crime mapping applications are at the tactical level; 2) current policing methods such as problem-oriented policing, COMPSTAT, hot spot enforcement, etc. are tactical; 3) police budgets are for the most part devoted to resources that can easily be redeployed in the short-term (personnel and vehicles) so that there is little need for medium and long-term forecasts; and 4) the largest and most active forecasting literature is on short-term methods so that there is much to draw on for this class of problem. Consequently, we forecast monthly crime count data one month ahead by police administrative area (precincts and car beats).

Seasonality estimates are especially useful for scheduling work during peak and trough times. For example, training and vacations can be scheduled in low crime months, while focused interventions may be targeted in high crime months. Police may also use short-term seasonal forecasts to direct patrols. For example, robberies of persons may increase in August in university areas with the increase of student populations and thus easy targets for robbers. Armed with this knowledge, police could study crime maps of past hot spot areas for robbery and intensify patrols in those areas, proactively.

Police use counterfactuals as the basis of comparison for judging whether or not there has been a change in crime. Two important kinds of changes considered by police are 1) changes in crime levels and 2) changes in crime patterns. A common police practice for evaluating changes in crime levels in small areas is questionable, given the volatility of small data aggregates. To estimate changes in crime levels, while roughly controlling for seasonality, police commonly calculate 12-month differences using the 12-month lagged data value as the counterfactual. If t is the most recent month for which we have crime data, and A_t is the actual crime level for month t, then the level change is:

$$\Delta_t^{12} = A_t - A_{t-12}$$

Δ_t^{12} includes a whole year's accumulation of level trend changes not accounted for in this measure. Moreover the variance of A_t is generally high because of small crime counts, thus Δ_t^{12} has a high variance and likely little information value. A better approach would be to replace A_{t-12} (and perhaps A_t) with time series model estimates in calculations of Δ_t^{12}.

Changes in crime patterns address questions such as "Have we stopped the increases in robberies that occurred this past year or past month?" The appropriate measure is $P_t^{\delta} = A_t - F_t^{\delta}$, where F_t^{δ} is the crime forecast made with data available up through t-δ. By definition, F_t^{δ} extrapolates or carries the status quo pattern forward in time, adjusting for known trends and seasonality. For example, serious crime was on an increasing trend in Pittsburgh, Pennsylvania during the early 1990s due to the crack epidemic and rise in youth gangs. In Pittsburgh's Precinct 2 there were 20 robberies of persons in January 1991. In January 1992, the number had risen steadily by approximately one robbery per month to 32 robberies. There was no change of crime pattern (P_t^{δ} = 0), although the robberies of persons had risen dramatically, $\Delta_t^{12} = 12$. Our results in Section 5.0 provide a comparison of alternative counterfactual estimates for one-month changes. If we assume that Δ_t^{12} is intended to provide a one-month change estimate, with A_{t-12} serving as a naïve forecast for A_t, then we can compare Δ_t^{12} directly with P_t^1 in terms of one-month ahead forecast accuracy.

3.0 Criminology Literature on Seasonality

There is general agreement that crime is to some extent a seasonal phenomenon. The most common findings at the city level are that 1) crimes against property (burglary, robbery, and theft) are high in fall and winter, and 2) crimes of aggression such as assaults, homicides and rape peak in mid-summer and are lowest in January (Cohen 1941, Tennenbaum et al. 1994). The oldest and for many years the most determined theory in the environmental criminology literature was the “temperature aggression hypothesis,” stating that weather increases violent crime by means of ambient temperature and anger arousal (Guerry 1833, Ferri 1882, Baron 1972, Rotton and Frey 1985, Anderson 1987, 1989, Cohn 1990, Bell 1992, Anderson and De Neve 1992). Another area of examination, predicated on a needs-based view of property crime, originating with Alexander Von Ottingen in 1882, and Falk in 1952, has suggested that seasonal unemployment and increased living expenses influence levels of criminal activity at different times of year.

Within the context of the new ‘opportunity’ theories of crime, which include routine activities, crime pattern theory, and the rational choice perspective (see Felson and Clarke 1998, for an overview of opportunity theories), seasonality can be viewed in a different light. Routine activity theory (Cohen and Felson 1979) holds that crime opportunities are concentrated in time and place, with spatial-temporal differences affecting the probability of convergence of three conditions: 1) motivated offenders, 2) suitable targets and 3) the absence of a capable guardian (see also Boggs 1966, Newman 1972, Repetto 1974, Felson 1974, Brantingham and Brantingham 1975, 1984).

Seasonality may give rise to changes in one or more of these three components in a number of ways. Hylleberg (1992) groups exogenous causes of crime into classes of calendar events such as the number of weekend days in a month, daylight savings, public holidays, and timing decisions such as school vacations, industry vacations, and tax year. Weather also influences the propensity to pursue discretionary activities (Chapin 1974, Cohen and Felson 1979, Landau 1993, Harries and Stadler 1998 pg.198). Schools are closed during the summer, some months have more long-weekends, and holidays such as Halloween, St. Patrick’s Day, July 4th, and New Years eve, may have differential effects in neighborhoods with different urban ecologies. Time of year can effect opportunities to commit crime in a number of ways. For example, during peak Christmas shopping season small high-value items such as electronics goods, which may be on open display in stores or in customers’ shopping bags. Cash registers contain more cash, and we expect an increase of crime in commercial areas. Similarly, individuals in the downtown shopping area may be carrying larger amounts of cash than usual at Christmas. In college neighborhoods students are moving possessions into dorms and apartments in August. An area with many bars may attract drunken revelers and hence propitious targets on weekends or holidays such as St. Patrick's Day and New Years Eve.

4.0 Experimental Design

Case study of precincts in Pittsburgh Pennsylvania

The spatial limit of our case study is the city boundary of Pittsburgh and its six police precincts[2] (see Figure 1 for a map and Table 1 for a profile of precincts). We selected precincts as the unit for analysis because they are large enough to provide adequate base rates of crime incidents, are reasonably similar in size, and conveniently organized based on neighborhoods and census tracts. Precincts also have obvious administrative advantageous in terms of usefulness to police.

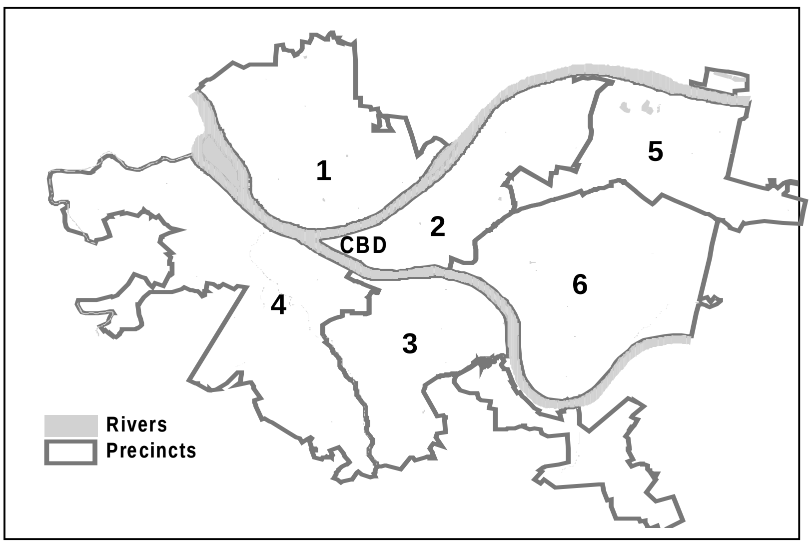

Figure 1. Pittsburgh Police Precincts

Most crime in Pittsburgh occurs in adjacent Precincts 2 and 3, both containing large run-down commercial centers in proximity to high-density, poor neighborhoods (see Table 1). Precinct 2 includes the Central Business District (CBD), and Precinct 5 contains the second largest commercial

---

[2] Pittsburgh Police refer to Precincts and Beats as 'Zones' and 'Sectors'.

center, East Liberty. Both of these precincts have large black populations, concentrations of public housing, and characteristics associated with low socio-economic status, including low median household income, a high percent of subsidized families, many female-headed households, and low levels of education. Precinct 6, which includes some of the wealthiest neighborhoods in the city, has low levels of violent crime, but the third highest levels of property crime after Precincts 2 and 5. Much of the crime in Precinct 6 is concentrated in poorer areas near the boundaries with Precincts 2 and 5.

Precinct 6 also contains a large number of university students and a younger population, providing a rich source of targets. Precinct 1, to the north (across the river) contains public housing projects, with sporadic flare ups of drug activity, some concentrations of bars and the third highest levels of simple assaults of the six precincts. Overall, Precincts 3 and 4 have lower population densities, fewer areas with commercial and residential concentrations, and fewer serious crimes in most categories.

Table 1
Pittsburgh Statistics by Police Precinct: 1990 Census

	1 North Side	2 CBD & Hill District	3 South Side	4 West End	5 East Liberty	6 Squirrel Hill
Land Use						
Population (1,000's)	53	28.7	50.5	73.2	84.5	88.2
Square Miles	9.2	4.1	9.5	13.4	9.0	10.7
Percent Built	17	18	12	10	19	18
Average Commercial Assessment ($1,000's)	34.1	116.5	15.8	36.4	29.7	53
Social						
Median Household Income ($1,000's)	22.5	14.2	19.8	24.4	18.3	28.1
Percent Subsidized Families	7	15	9	3	7	1
Percent Black	31	59	20	10	47	14
Percent Female Headed Households	18	24	21	15	21	11
Percent Unemployed	11	17	13	7	12	7
Education						
Percent College Graduates	10	12	8	13	17	46
Percent over age 18 No High School	32	36	33	25	28	15

Data Sample

The primary data consist of eight years (approximately one million records) of 911 CAD (Computer Aided Dispatch) and offense data for all individual events for 1991 through 1998 obtained from the Pittsburgh Bureau of Police. We chose to focus on Part 1 (serious) crimes, selecting two property-crimes (robbery and burglary) and two violent-crimes (simple and aggravated assault). The most often studied crime, homicide, had to be excluded, as monthly crime counts were too small. Because drug activity has been so highly associated with serious crime, we included drug calls to police. We use CAD drug data rather than police offense reports because calls made by the public are a better reflection of actual drug market levels. Offense reports for drugs are constrained by police selectivity bias and the relatively few narcotics police making arrests. We located all offense records from 1991 to 1998 by address, using a geographic information system (GIS), yielding points on a map. Point data were then spatially aggregated into monthly time series of crime counts by precinct, compiling 35 univariate time series, one for each of seven locations (six precincts and the city), and for five crime types. Table 2 provides summary statistics for these series.

As Table 2 shows, the crime types selected provide considerable variety in terms of the crime counts, varying from a low of 85 for aggravated assaults to a high of 896 for simple assaults. When crimes are disaggregated to the Precinct level, aggravated assaults range from an average of only nine per month in Precincts 3 and 6 to 23 per month in Precinct 5; burglary averages 41 to 76 monthly crimes per precinct; drugs average 27 to 102 (drug activity tends to be more concentrated in a few areas); robbery has an average of 11 to 43 monthly crimes and simple assault is the most frequent and uniformly distributed crime, with an average of 105 to 201 per precinct per month.

Table 2
Pittsburgh Crime Statistics 1990-1998 (N=36)

Crime	Statistic	Precinct 1	2	3	4	5	6	City
Aggravated Assault	**Mean Monthly Crime Count**	**14**	**20**	**9**	**10**	**23**	**9**	**85**
	Standard Deviation	6	8	5	4	8	4	26
	Coefficient Of Variation	0.48	0.38	0.53	0.40	0.33	0.45	0.31
	Minimum	3	6	1	2	9	2	46
	Maximum	44	49	28	22	43	25	182
Burglary	**Mean Monthly Crime Count**	**53**	**58**	**33**	**41**	**76**	**74**	**335**
	Standard Deviation	17	17	12	12	25	28	92
	Coefficient Of Variation	0.32	0.29	0.34	0.28	0.33	0.38	0.28
	Minimum	26	28	12	21	27	25	166
	Maximum	116	101	70	75	142	143	571
Drugs	**Mean Monthly Crimes**	**60**	**103**	**42**	**27**	**102**	**30**	**364**
	Standard Deviation	22	54	29	10	76	10	174
	Coefficient Of Variation	0.37	0.52	0.69	0.37	0.75	0.33	0.48
	Minimum	19	35	9	2	15	4	102
	Maximum	119	279	160	57	368	68	852
Robbery	**Mean Monthly Crime Count**	**23**	**43**	**14**	**11**	**38**	**27**	**155**
	Standard Deviation	7	13	6	4	14	9	42
	Coefficient Of Variation	0.32	0.30	0.44	0.36	0.36	0.35	0.27
	Minimum	6	22	2	4	10	8	88
	Maximum	51	76	34	26	88	49	263
Simple Assault	**Mean Monthly Crime Count**	**153**	**201**	**105**	**125**	**191**	**121**	**896**
	Standard Deviation	35	44	24	27	41	25	165
	Coefficient Of Variation	0.23	0.22	0.23	0.21	0.22	0.20	0.18
	Minimum	64	112	63	70	114	68	553
	Maximum	249	336	165	199	295	199	1331

Table 2 also shows that the crime-space-time series data are highly variable, with precinct coefficients of variation in the range of 0.20 for Simple assaults in Precinct Six, to 0.75 for drug offenses in Precinct Five. The coefficient of variation is greatest for drug offenses (0.51 on average), due to the dramatic declines in drug activity during the study period. If we exclude drug offenses, on average the coefficients of variation are highest for low-frequency crimes of aggravated assaults and robbery, followed by burglary, and are also highest for all crime types in Precinct 3, making these series the most challenging to smooth and forecast. For example, the average number of monthly aggravated assaults in Precinct 3 was only nine, with a range of one to 28, and robberies ranged from two to 34, with an average of 14 per month. We will see in Section 6 that this series produced the highest forecast errors, 47% on average. Overall, offense patterns in Pittsburgh exhibit the type of geographic differentiation, affected by the demographic and economic structure and by aspects of the physical environment (such as land use) commonly found in cities in numerous studies elsewhere (see Shaw 1929, Wilks 1976, Reiss 1976, Blau 1977, Sampson 1993, Hirschfield and Bowers 1997, for example).

Figure 2 shows annual time-series plots of each crime type. Burglaries, robberies, and drug calls experienced particularly strong downward trends during the study period. For drugs, Precincts 1 and 3 had increases in the first year of the estimation period, followed by a reversal and downward trend thereafter. In contrast, in Precincts 2 and 5 the decline in drugs began earlier and was far more precipitous. For robberies there were substantial declines (as much as 50%) in the high-crime Precincts 2, 5, and 6, beginning in 1993 in Precinct 5 and 1994 in Precinct 2. In Precincts 2 and 5 patterns in violent crimes (aggravated and simple assaults) closely track those of robbery. Burglaries were in a steady decline through until 1997, when we see increases in Precincts 3 and 2.

Treatments

We used classical decomposition (Makridakis et al, 1978) to calculate monthly seasonal indices at both the city and precinct level with multiplicative ratio-to-moving averages [3]. A seasonal index indicates the seasonality for a given month in relation to the other months. The base value of one indicates no seasonal influence. A seasonal index of 1.2 indicates that the month averages a 20 percent increase over the non-seasonal time line. Conversely, an index of 0.8 indicates that the effect of seasonality is 20 percent below average. For each the 30 precinct series (five crime-types and six precincts) we generated two deseasonalized series: one series calculated using the individual precinct's seasonal indices (SI's), and another series using the pooled city seasonal indices's (CSI's), calculated using the entire data sample for Pittsburgh. We forecast a total of 100 series (30 unadjusted precinct series, 30 precinct series deseasonalized using SI's, 30 precinct series deseasonalized using CSI's, five unadjusted city series, and five city series deseasonalized with CSI's).

[3] There are two forms of seasonality; additive or multiplicative adjustments to a time trend model. Additive adjustments, in units of numbers of crimes, reflect the scale of crime in different areas. Hence such seasonal factors are not readily transferable or comparable across areal units. We use multiplicative adjustments because they have the advantage of being dimensionless and thus are well suited for application across areal units of differing crime scales.

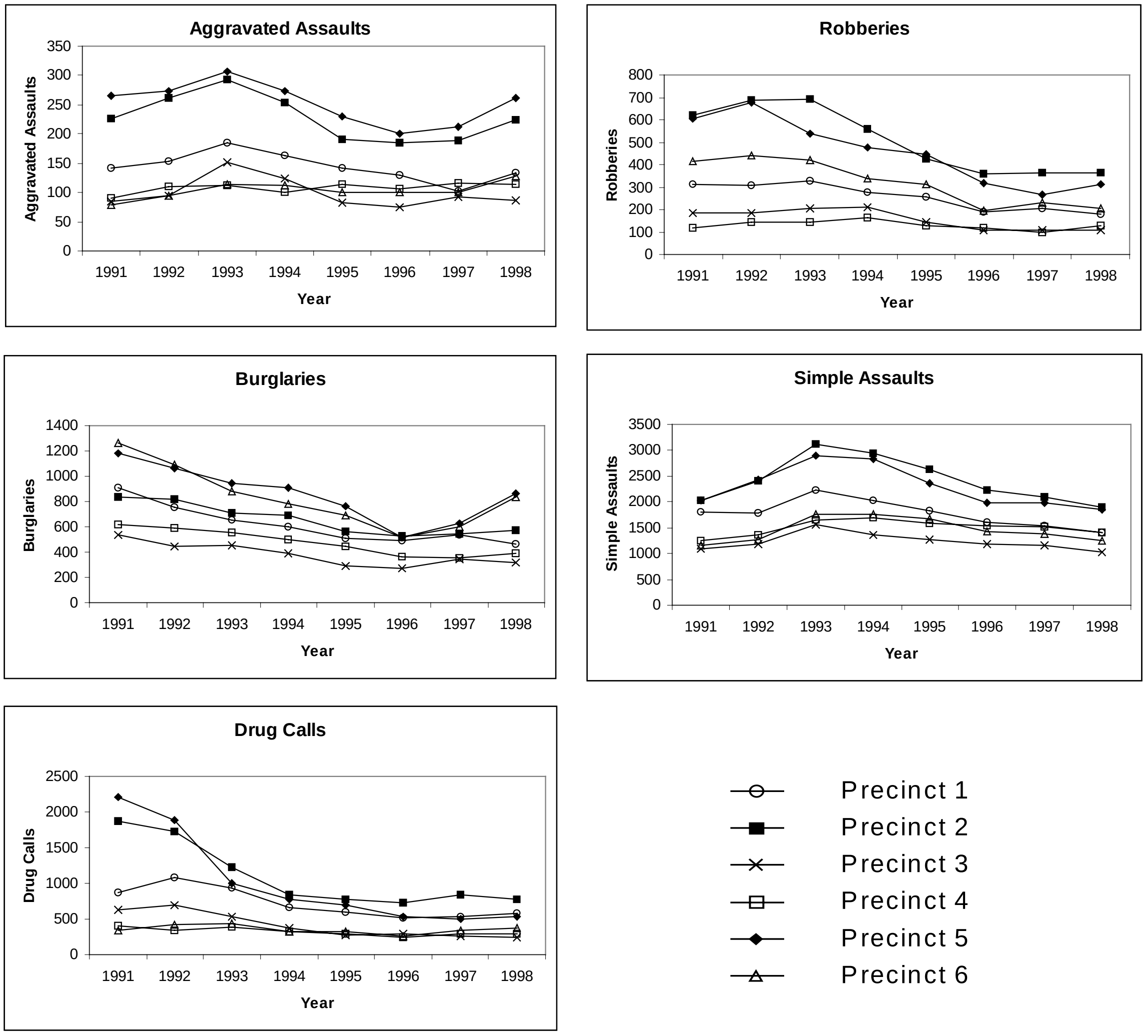

Figure 2. Crime Trends in Pittsburgh Police Precincts by Crime Type, 1991- 1998

Only data before the forecast origin, in the preceding five-year period, were used for the decomposition. The seasonal indices were recalculated annually for each series (three times, once for each forecast year), using a rolling-window of five-year periods (1991 through 1995; 1992 through 1996; and 1993 through 1997). We then forecast each of the 100 series using three forecasting methods: Random Walk, Brown's Simple Exponential Smoothing (Brown 1963), and Holt's Two-Parameter Linear Exponential Smoothing (Holt 1957, Makridakis and Wheelwright 1978). Exponential smoothing is commonly used in forecasting for short-term. Generally smoothing methods must trade off responsiveness to structural changes versus precision of estimates. We used an optimization procedure to estimate the smoothing parameters.

The ten methods employed are:

1) *Naïve 1* – also known as the ‘random-walk’, is the ‘straw-man’ in our model set, the simplest model. The number of assaults in February 1996 is forecasted to be the same as the number in January 1996, or in general $X^F{}_t = X_{t-1}$. Naïve 1 is a poor predictor when there is high randomness (i.e. a large coefficients of variation) or strong trend and seasonality. It is a good predictor for series with frequent pattern changes because it adapts immediately (Duncan et al., 1995).

2) *Naïve 1, Lag 12* – has been included because it is a method commonly used by many police departments to estimate crime level while controlling for seasonality. It provides a rough measure of seasonality. If t is the most recent month for which we have crime data then the forecast for the next period will be $X^F{}_t = X_{t-12}$. Simply, the number of assaults in February 1996 is forecasted to be the same as the number in February 1995. Problems with this method are that the forecast does not include the 12 months trend changes and individual crime counts are very noisy.

3) *Naive 2* – is the same as Naïve 1, but uses data deseasonalized with individual precinct seasonal indices, $X^F{}_t = X^{SI}_{t-1}$.

4) *Naïve 2 with Pooling* – $X^F{}_t = X^{CSI}_{t-1}$, is the random-walk using deseasonalized data with the pooled citywide City Seasonal Indices.

5) *Simple exponential smoothing* – (Brown 1963), is based on averaging past values with exponentially decreasing weights. The more the level of the process is changing, the more a newly observed time series value should influence smoothed estimates, and thus the larger the smoothing constant should be set. The noisier the data the lower should be the weight given to the most recent observation. The forecast is the most recently smoothed value:

$$X^F{}_{t+1} = \alpha X_t + (1-\alpha) X^F_t$$

where α = smoothing constant $0 \le \alpha \le 1$ estimated optimally via a grid search.

6) *Exponential smoothing on deseasonalized data* – uses Brown’s method on data that has first been deseasonalized using individual precinct seasonal indices.

7) *Exponential smoothing pooled* – is Brown’s method on data that has been deseasonalized using the pooled city seasonal indices.

8) *Holt's two parameter exponential smoothing method* – estimates a smoothed time trend model (see, e.g., Makridakis et al. 1983, p.97), also with optimal smoothing constants via a grid search.

Holt's method is given by three recursive equations:

$$L_t = \alpha X_t + (1-\alpha)(L_{t-1} + S_{t-1}) \quad (1)$$
$$S_t = \beta(L_t - L_{t-1}) + (1-\beta)S_{(t-1)} \quad (2)$$
$$X^F_{t+m} = L_t + mS_t \quad (3)$$

where

L_t = the current smoothed level of the time series

X_t = the current data series

S_t = the current smoothed slope of the time series unit

X^F_{t+m} = the m step ahead forecast

T = the forecast origin (last historical data point)

m = the number of periods ahead to be forecast and

$0 \leq \alpha \leq 1$ and $0 \leq \beta \leq 1$ are the smoothing parameters.

9) *Holt's exponential smoothing on deseasonalized data* – using individual precinct seasonal indices.

10) *Holt's exponential smoothing pooled* – is Holt's method on data that has been deseasonalized using he pooled city seasonal indices.

We selected parameters for exponential smoothing through a grid search procedure, with values ranging from 0 to 1 with a step of 0.1, selecting the values for the smoothing parameters (α for the level smoothing parameter in simple exponential smoothing; α and β for level and trend parameters in Holt's two parameter exponential smoothing) that minimize the sum of squared one-step-ahead forecast errors. (Makridakis et al. 1983, pp. 89-90) The grid search optimization routine was repeated annually to re-optimize the parameters.

When using the exponential smoothing method, irrespective of deseasonalization or pooling, the average level parameter for precincts was between 0.2 and 0.3 for all crimes except drugs. For Holt's method the average level parameter for all crimes is 0.3 without seasonality, 0.25 with deseasonalization, and 0.24 with pooling. Drug offenses, however, behaved quite differently. For drugs, the average precinct level parameter for exponential smoothing was 0.5. After the data were deseasonalized, with or without pooling, the level estimate fell to an average 0.3. This was mainly due to Precincts 1 and 3, where there were pattern changes: in these precincts deseasonalization and pooling reduced the optimal level parameter by half. In Precinct 1 the optimal level parameter estimates on the unadjusted series were extremely high, 0.8 on average, whereas after deseasonalization or pooling the average estimate was 0.4. Trend parameter estimates average 0.16 to 0.17 for all crimes for all precincts regardless of seasonality or pooling.

Error Measures and Rolling Horizon Design

Forecast model validation is based on splitting data samples into both estimation and holdout samples to measure out-of-sample forecast accuracy. We used the rolling horizon design (Makridakis et al. 1982, 1983) to produce 36 one-month-ahead forecasts for each precinct over the period 1996 through 1998. For each monthly forecast, the immediately preceding 60-month period was used to estimate the forecast model, and the forecast was then compared to the actual holdout month to calculate the resultant error. Thus, the forecast for February 1996 is based on data from the period February 1, 1991 through January 31, 1996. Rolling the window one month forward, the forecast for March 1996 is then based on March 1, 1991 through February 31, 1996, and so on. This rolling horizon approach gives a sample size of 36 forecast errors per time series. A rolling horizon increases the number of degrees of freedom for comparisons and experience over many contexts and times that is required for generalization of results.

Since no single metric is overall best for comparisons (Armstrong and Collopy 1992) three different methods were used for forecast accuracy measurement: mean absolute error (MAE), mean squared errors (MSE), and mean absolute percent error (MAPE). Our results, nevertheless, are similar for all measures. For ease of interpretation we present results on the MAPE, which eliminates scale differences between precincts.

$$MAPE_{T+k} = (100/N) \sum_{i=1}^{N} \left| Y_{i,T+k} - Y^{F}_{i,T+k} \right| / Y_{i,T+k}$$

with rolling horizon, where

T	= forecast origin
K	= forecast horizon
$Y_{i,T+k}$	= data value for origin T, horizon k, and unit i

$Y^F_{i,T+k}$ = forecast for origin T, horizon k, and unit i

N = number of observational units in the reference group

5.0 Results

Seasonality

In general, we find the classical seasonal patterns of increased property crime levels late in the year (crime opportunities afforded by holidays) and increased aggression crimes in summer due to increased social interactions. There is little support in our findings that criminological cold weather might increase burglary and robbery crimes due to seasonal economic pressures or unemployment.

The heavy line in each of the graphs in Figure 3 is the city seasonality index (CSI), and the bars in each month are seasonal indices (SI's) for precincts one to six, ordered left to right. There is evidence of seasonal variations in precincts around the mean city level. For aggravated assaults, robbery, and burglary, seasonal influences vary from place to place within the city of Pittsburgh. In contrast to the heterogeneity of aggravated assaults, robbery, and burglary, the SI's for drugs and simple assaults are relatively homogenous. Table 3 provides the seasonal indices based on data for 1992 to 1996. In the case of robberies, at the city level there is no discernible seasonal effect for January, March, June, July and November, whereas there are strong differences at the precinct level that cancel out at the city level.

As would be expected, the seasonal indices for precincts are more extreme than are those for the city as a whole. For infrequent crimes of robbery and aggravated assault the extremes are more pronounced, because precincts contain smaller counts and are therefore more sensitive to the impacts of special events such as a police crackdown on gangs or prison release of a high offender. However, even in higher count crimes of simple assault and burglary there are precincts with opposite seasonal indices. Burglary has notable variation in seasonality between precincts in some months. Seasonality often moves in opposite directions in different precincts, perhaps depending on the commercial/residential mix. Seasonal results are discussed below by crime type.

1. Simple assault, the most frequent crime, has the least seasonal variation, with a range of 0.08 to 1.2. Simple assaults are highest in the warmer months from May to October, and lowest from November through March. Block and Block (1984, 1988) generally did not find evidence of seasonality for assaults, and have suggested that the monthly variability lies in the surveillance, not the criminal activity.

2. Aggravated assault is strongly seasonal with indices ranging between 0.7 and 1.6. However, extreme high or low seasonal indices may be the effect of small monthly crime counts in some precincts. The seasonal index of 1.6 in August in Precinct 3, for example, is based on an average frequency of only nine aggravated assaults for August. Despite some suggestion of underlying volatility in aggravated assaults, due to small crime counts, the seasonal indices are generally consistent with Cohen (1941), who found peaks in midsummer with January being the lowest month. In a regression analysis, using dummy variables for months we found that Pittsburgh had significant and large parameter estimates for July, August and September. Running the same regression for precincts, we found that most precincts had significant and positive coefficients for summer months. The seasonal indices appear to be largely spatially homogenous. It is unclear to what degree seasonality is a result of variability in surveillance rather than time of year differences in frequency of social interaction or temperature.

3. We find seasonality for **robbery** is interesting, as it behaves neither entirely like a violent crime nor like a property crime, showing more month-to-month oscillation in seasonality. The October peak (in five precincts) is consistent with the same peak found in burglary, and the December extremes (1.2 or above) in all precincts is somewhat consistent but more extreme than for burglary. Negative seasonality in May found in robbery appears in no other crime type, but could suggest some mechanism of reduced opportunity in May.

Robbery crimes are supposed to increase in the fall and winter. Deutch (1978), for example, finds peaks in December or January, and Cheatwood (1988) found that robberies are most likely to be high in October, November, December and January. Although Cheatwood’s review of the literature concludes that robbery tends to be more prevalent in the colder months, our results differ. With the important exception of December, winter months are never positive for robbery. For precincts we find most seasonal indices in excess of 1.1 in October to December, but none in January or February. This gives cause to question the validity of several explanatory theories in the literature. It has been suggested, for example, that robberies may be higher in the winter due to more hours of darkness (providing cover) and fewer potential witnesses.

Another theory is that the increased cost of living in winter and seasonal unemployment stimulate increases in property crime motivated by need (Sutherland 1947; Haran and Martin 1984, Landau and Pfeffermann, 1988 Cohn 1990,). This does not appear to hold true for robbery. An argument in the opposite direction is that there are fewer robbery targets in most on-street locations in winter months, with the notable exception of the December Christmas shopping period.[4]

There is spatial heterogeneity of seasonal effects for robbery (for example, contiguous precincts 2 and 5, both containing with major commercial centers, have opposite robbery effects for October and November). Some of the monthly oscillation in seasonal indices between precincts might be explained if alternative locations are substitutable target destinations for perpetrators, and moving around to different areas is a way of avoiding detection. Robbery may allow an offender greater geographic mobility than burglary perhaps requiring less specific local knowledge than burglary, such as which properties are likely to be unattended, and greater portability of the items stolen – mostly cash and jewelry.

4. Burglaries have seasonal peaks in August in all precincts, and peaks in October in three precincts. Seasonality is uniformly negative in February and March. Overall, seasonality is very moderate, with most indices close to one. There is some spatial heterogeneity, some of which may be due to conflicting temporal patterns in commercial versus residential robberies. For example Precinct 2, containing the CBD, is more dominated by commercial activity, possibly explaining its positive seasonal indices in December and January, during the Christmas season, as compared to Precinct 5 which has little seasonality during these months. There is an interesting positive SI in October of 1.3 in Precinct 5, in contrast to either very low or else negative seasonality in other precinct. This suggests some underlying phenomenon the effect of which is specific to Precinct 5. As we

[4] The regression Robbery against Time with dummy variables for months found no significant months in any police precincts, but there was significant effect of individual months detectible at the city level, with December having the strongest positive effect on robberies.

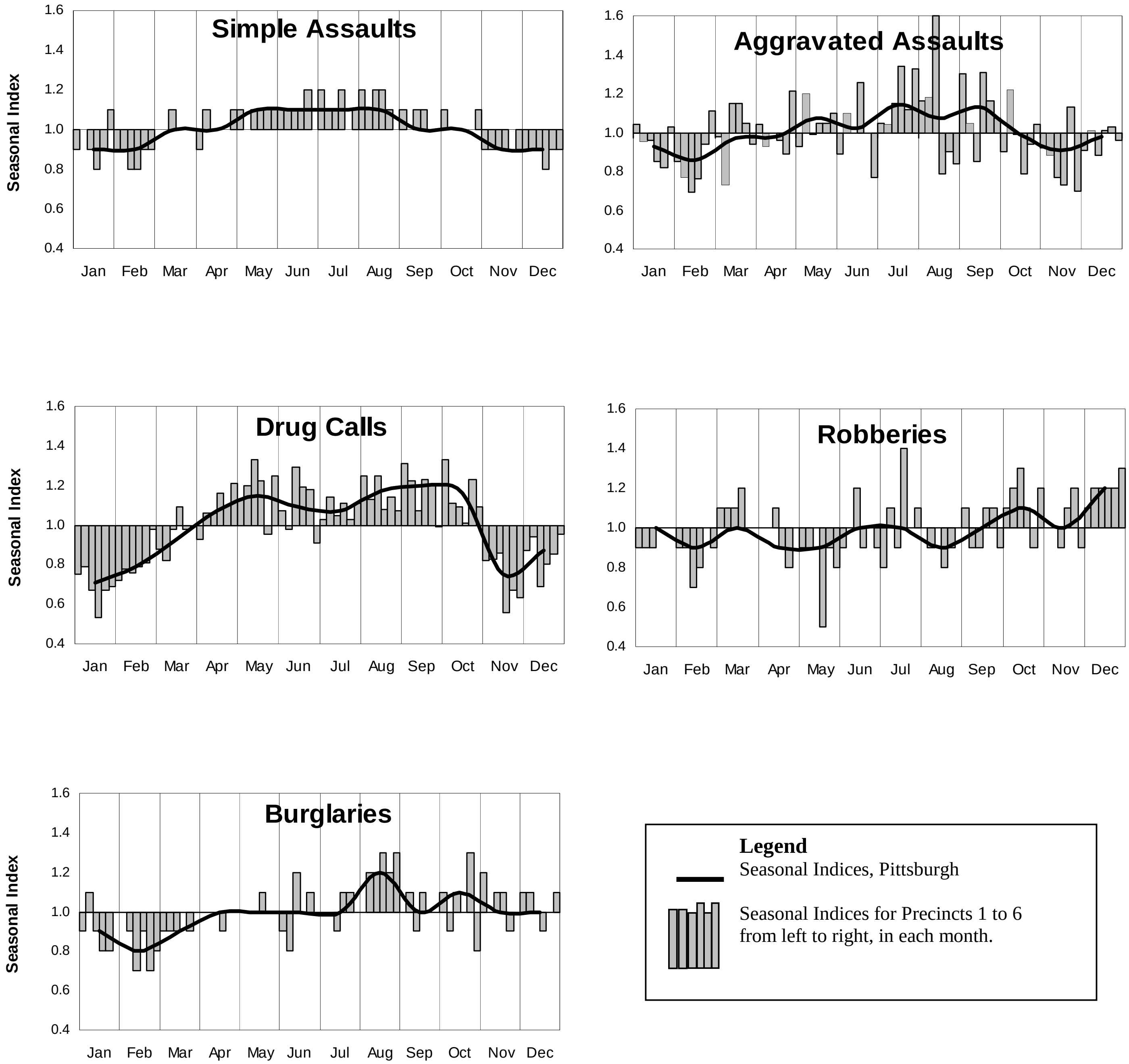

Figure 3 Seasonal factors for Pittsburgh and Precincts 1991-1995

Table 3

Seasonal Indices Pittsburgh for 1992-1996

Simple Assaults

Precinct	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	0.9	0.9	1.0	0.9	1.1	1.1	1.2	1.1	1.1	1.1	0.9	0.9
2	0.9	0.9	1.0	1.1	1.0	1.1	1.1	1.1	1.0	1.0	0.9	0.9
3	0.9	0.8	1.1	1.0	1.1	1.1	1.1	1.2	1.1	1.0	0.9	0.9
4	0.8	0.8	1.0	1.0	1.1	1.1	1.2	1.2	1.1	1.0	0.9	0.8
5	0.9	0.9	1.0	1.0	1.1	1.2	1.0	1.1	1.0	1.0	1.0	0.9
6	1.1	0.9	1.0	1.1	1.1	0.9	1.1	1.0	1.0	1.1	0.9	0.9
City	0.9	0.9	1.0	1.0	1.1	1.1	1.1	1.1	1.0	1.0	0.9	0.9

Aggravated Assaults

Precinct	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	1.0	0.9	1.0	1.0	0.9	0.9	1.1	1.2	1.3	0.9	0.9	0.9
2	1.0	0.8	0.7	0.9	1.2	1.1	1.0	1.2	1.1	1.2	0.9	1.0
3	1.0	0.7	1.2	1.0	1.0	1.0	1.2	1.6	0.9	1.0	0.8	0.9
4	0.9	0.8	1.2	1.0	1.1	1.3	1.3	0.8	1.3	0.8	0.7	1.0
5	0.8	0.9	1.1	0.9	1.1	1.0	1.1	0.9	1.2	0.9	1.1	1.0
6	1.0	1.1	0.9	1.2	1.1	0.8	1.3	0.8	1.1	1.0	0.7	1.0
City	0.9	0.9	1.0	1.0	1.1	1.0	1.1	1.1	1.1	1.0	0.9	1.0

Robberies

Precinct	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	0.9	0.9	1.1	1.0	0.9	0.9	0.8	1.0	1.1	1.1	1.0	1.1
2	0.9	0.9	1.1	1.0	0.9	1.0	1.1	0.9	0.9	1.2	1.0	1.2
3	0.9	0.7	1.1	1.1	1.0	1.2	0.9	1.0	0.9	1.3	0.9	1.2
4	1.0	0.8	1.2	0.9	0.5	0.9	1.4	0.8	1.1	1.1	1.1	1.2
5	1.0	1.0	1.0	0.8	0.9	1.0	1.0	0.9	1.1	0.9	1.2	1.2
6	1.0	0.9	1.0	1.0	0.8	0.9	1.1	1.0	0.9	1.2	0.9	1.3
City	1.0	0.9	1.0	0.9	0.9	1.0	1.0	0.9	1.0	1.1	1.0	1.2

Burglaries

Precinct	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	0.9	1.0	0.9	1.0	1.0	0.9	1.0	1.0	1.0	1.1	1.2	1.1
2	1.1	0.9	0.9	1.0	1.0	0.8	1.0	1.2	1.1	0.9	1.0	1.1
3	0.9	0.7	0.9	1.0	1.0	1.2	0.9	1.2	0.9	1.1	1.1	1.0
4	0.8	0.9	1.0	0.9	1.1	1.0	1.1	1.3	1.1	1.0	1.1	0.9
5	0.8	0.7	0.9	1.0	1.0	1.1	1.1	1.2	1.0	1.3	0.9	1.0
6	1.0	0.8	1.0	1.0	1.0	1.0	1.0	1.3	1.0	0.8	1.0	1.1
City	0.9	0.8	0.9	1.0	1.0	1.0	1.0	1.2	1.0	1.1	1.0	1.0

Drug Calls

Precinct	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	0.8	0.7	0.9	0.9	1.1	1.1	1.0	1.3	1.3	1.3	0.8	0.9
2	0.8	0.8	0.8	1.1	1.2	1.0	1.1	1.1	1.2	1.1	0.8	0.9
3	0.7	0.8	1.0	1.1	1.3	1.3	1.1	1.3	1.1	1.1	0.9	0.7
4	0.5	0.8	1.1	1.2	1.2	1.2	1.1	1.1	1.2	1.0	0.6	0.8
5	0.7	0.8	1.0	1.1	1.0	1.2	1.0	1.1	1.2	1.2	0.7	0.9
6	0.7	1.0	1.0	1.2	1.3	0.9	1.1	1.1	1.0	1.1	0.6	1.0
City	0.7	0.8	0.9	1.1	1.2	1.1	1.1	1.2	1.2	1.2	0.8	0.9

will see later (in Table 8) deseasonalizing the crime series does not usually reduce forecast errors. Seasonal indices are likely not capturing commercial versus residential differences, and these should probably be forecast as separate crime types.

5. Drug calls parallel the pattern for aggravated assaults, except that the strength of the seasonal indices is more pronounced. Seasonal indices suggest a weather influence to weather, with most indices greater than 1.0 for the warmer and dryer months of April through October, and less than 1.0 for November to March. Drug SI's are also the most homogeneous of all crimes tested. Since these offenses are based on 911 calls this pattern may be more a function of visibility than actual drug activity. It may be that drug dealing moves indoors in the winter, or else those most troubled by the street activity are likely either to call in a complaint to the police or have themselves retreated indoors and are less likely to observe the activity. Nonetheless, from the forecast results in Table 9 it appears that seasonality is being well captured at the city level.

Forecasting Results and Significance Testing

Tables 4 through 9 provide the relative performance of the alternative forecasting methods from the rolling horizon experiments. In order to facilitate interpretation, the mean absolute percent error (MAPE) for each method was divided by the MAPE of the "best" method, providing a relative comparison. The method with the lowest MAPE thus has a value of one, but, because there are often several competing methods with very similar results, the method with the lowest value beyond one decimal place for each Precinct has been highlighted. A value of 1.2 indicates that the MAPE for that method was 20 percent higher than the best method for that crime type and precinct. The actual MAPE for the highlighted best method is given at the bottom of each column. This allows the original MAPE's to be easily re-calculated, by multiplying the value in the cell by the minimum MAPE at the bottom of the column.

We used the Wilcoxon two-sample rank sum test (also called the Mann-Whitney test) for matched pairs, to test whether the difference between the best method and any alternative method was statistically significant. We paired the 36 APE's from each of the nine competing methods with the 36 APE's of the best method, calculating the paired differences, and in each case testing the hypothesis that the median difference is 0 versus median greater than 0. The methods that test as significantly worse than the best method are identified by asterisks. For example, for Simple

Assaults in Precinct 2, the best method is Holt with Pooling, with a MAPE of 8.4 percent. Under the Wilcoxon test all other methods have median APE's that are significantly higher than Holt's with Pooling. In Precinct 6 forecast errors there are eight statistically equivalent methods, and two (Naïve 1 and Naïve 1 Lag12) that are significantly worse than Holt's.

Forecasting results for Simple Assaults in Table 4 are excellent, with a MAPE for Pittsburgh of 5.2 percent and between 8.4 and 10.5 percent for precincts. Simple assaults stand out as the only crime type where both seasonality and time trend are important in forecasting, as is borne out by the overall superior performance of Holt Deseasonalized. Simple assault offenses exhibited strong time trends, increasing rapidly between 1992 and 1994, declining at a lesser rate between 1994 and 1998 (see Figure 2). Holt Deseasonalized Pooled improved the forecasts in four out of six precincts, although the improvement was statistically significant only in the highest crime district, Precinct 2, where pooling resulted in a 30 percent reduction in the MAPE versus using the precinct seasonal indices. Seasonality is critical when forecasting simple assaults, and extrapolative methods consistently outperform Naïve methods. All Naïve methods, in the first four rows in Table 4, are significantly worse than deseasonalized extrapolative method, often more than 50 percent worse. The method commonly used by police, Naïve 1, Lag 12, is overall the worst method. In four out of six precincts the Naïve 1 Lag 12 method results in errors 50 to 90 percent higher on average than Holt Pooled, with an average precinct MAPE of 17 percent, compared to 11 percent for Holt Pooled.

Table 1
Mean Absolute Percentage Error Forecast Accuracy for
One-Step-Ahead Monthly Forecasts:
Simple Assaults in Pittsburgh for 1996, 1997, 1998

Method	Precinct 1	2	3	4	5	6	City
Naïve 1	1.4*	1.4***	1.6***	1.4***	1.4*	1.0	1.5***
Naïve 1, Lag 12	1.2***	1.9***	1.5***	1.2**	1.7***	1.6**	1.9***
Naïve 2	1.3**	1.4***	1.5***	1.4***	1.2**	1.4**	1.1*
Naïve 2 Pooled	1.3**	1.3***	1.4***	1.1*	1.2*	1.0	1.1*
EXPO No Seasonality	1.2**	1.4***	1.6***	1.3**	1.3**	1.0	1.8***
EXPO Deseasonalized	1.0*	1.1**	1.1**	1.0	1.1*	1.1	1.0
EXPO Deseasonalized Pooled	1.0	1.1*	1.0	1.0	1.1**	1.1	1.0
HOLT No Seasonality	1.1*	1.5***	1.6***	1.4**	1.4***	1.0	2.0***
HOLT Deseasonalized	1.0	1.3***	1.0	1.0	1.0	1.0	1.0
HOLT Deseasonalized Pooled	1.0	1.0	1.0	1.0	1.0	1.1	1.0
Minimum MAPE	12.3%	8.4%	10.5%	13.3%	9.6%	10.7%	5.2%
Average Monthly Simple Assaults	153	201	105	125	191	121	896

Significance (Wilcox Rank Pairs) * ** p< .001 ** p< .01 * p< 0.1

Aggravated assaults (see Table 5) sometimes have very high forecast errors of over 30 percent in half of the precincts. The precincts with higher levels of aggravated assaults (Precincts 2 and 5) forecast with the most precision. Precinct 3, with an average of only nine offenses per month, has the highest MAPE of 47 percent. Despite the fact that all Precincts, with the exception of Precinct 4, exhibited modest trends, with minor trend reversals in aggravated assaults during both the estimation and the forecasting periods, Holt's method did not out-perform simple exponential smoothing. Simple exponential smoothing always performs substantially and significantly better, on average 24 percent better, than Naive methods. Simple exponential smoothing for Pittsburgh reduced the MAPE from 18 percent for the Naïve 1 method to 13 percent. Deseasonalizing the data, pooling, or including a trend parameter (Holt's) did not significantly improve the forecasts. Pooling had no statistically significant effect.

Table 5
Mean Absolute Percentage Error Forecast Accuracy for
One-Step-Ahead Monthly Forecasts :
Aggravated Assaults in Pittsburgh for 1996, 1997, 1998

Method	Precinct						
	1	2	3	4	5	6	City
Naïve 1	1.3***	1.1	1.4***	1.4***	1.0	1.3***	1.3***
Naïve 1 Lag 12	1.6***	1.1*	1.1	1.1*	1.3*	1.3**	1.1
Naïve 2	1.3***	1.2**	1.5**	1.3**	1.1	1.3**	1.2**
Naïve 2 Pooled	1.2**	1.2**	1.3***	1.3**	1.0	1.3**	1.2**
EXPO No Seasonality	1.0	1.1	1.1	1.0	1.0	1.0	1.1
EXPO Deseasonalized	1.1***	1.0***	1.0*	1.2	1.0	1.3*	1.0
EXPO Deseasonalized Pooled	1.1***	1.0	1.0	1.2	1.1	1.3**	1.0
HOLT No Seasonality	1.0	1.2**	1.1	1.1**	1.0	1.0	1.1
HOLT Deseasonalized	1.2**	1.1**	1.1***	1.2	1.2*	1.2**	1.0
HOLT Deseasonalized Pooled	1.1*	1.0	1.1**	1.1	1.1	1.3**	1.0
Minimum MAPE	39%	29%	47%	28%	23%	34%	14%
Average monthly aggravated	14	20	9	10	23	9	85

Significance (Wilcox Rank Pairs) * ** p< .001 ** p< .01 * p< 0.1

Robbery incidence exhibited both upward and downward trends in Pittsburgh, closely mirroring national trends.[5] As a result, precincts had on average 29 percent fewer serious property crimes during the forecasting period than during the estimation period. In most areas of Pittsburgh crime counts for robbery are low for forecasting purposes, averaging between 11 and 43 per month by precinct. On average, the overall trends were declining in all precincts in both the estimation and forecasting periods (see Table 6). Thus the forecast results for robberies are similar to those for aggravated assaults, with relatively large forecast MAPE's corresponding in large part to monthly average robbery counts per precinct (see Table 7). The precincts with commercial areas and drug markets, Precincts 2 and 5, have the highest data frequencies and consequently the lowest forecast MAPE's (14 and 22 percent, respectively).

[5] Robbery, a fast, easy, cash crime (defined as theft using force or the threat of physical injury), became favored over burglary in drug areas in many US cities in the late1980's. (Baumer et al 1999)

Table 2
Percent Change in Number of Robberies and Burglaries from Estimation Period (N=60) to the Forecast Period (N=12)

Precinct	Crime Type		
	Part 1 Property	Robbery	Burglary
1	-29	-19	-23
2	-32	-16	-21
3	-37	-17	-20
4	-15	-24	-19
5	-37	-19	-24
6	-38	-16	-30
Pittsburgh	-33	-19	-24

Table 3
Mean Absolute Percentage Error Forecast Accuracy for
One-Step-Ahead Monthly Forecasts:
Robberies in Pittsburgh for 1996, 1997, 1998

Method	Precinct 1	2	3	4	5	6	City
Naïve 1	1.2**	1.5***	1.4**	1.3***	1.3***	1.3**	1.2**
Naïve 1 Lag 12	1.5***	1.2**	1.8**	1.1	1.5**	1.7***	1.5**
Naïve 2	1.3***	1.4**	1.6**	1.4**	1.4**	1.1*	1.3*
Naïve 2 Pooled	1.3**	1.4**	1.5***	1.3*	1.5***	1.2*	1.3*
EXPO No Seasonality	1.0	1.1*	1.2**	1.2	1.0	1.2*	1.0
EXPO Deseasonalized	1.0*	1.1	1.4**	1.1	1.1	1.0	1.1
EXPO Deseasonalized Pooled	1.0	1.1	1.3**	1.0	1.0	1.0	1.1
HOLT No Seasonality	1.0	1.0	1.0	1.0	1.0	1.8***	1.0
HOLT Deseasonalized	1.0	1.1	1.2*	1.0	1.1	1.3***	1.0
HOLT Deseasonalized Pooled	1.0	1.0	1.2	1.0	1.1	1.3***	1.0
Minimum MAPE	34%	14%	46%	35%	22%	26%	12%
Average Monthly Robberies	23	43	14	11	38	27	155

Significance (Wilcox Rank Pairs) * ** p< .001 ** p< .01 * p< 0.1

As is the case with aggravated assaults, smoothing methods are overall best for forecasting robberies, with Naive methods performing on average 27 percent worse (see Table 7). In Precincts 3 and 5, the highest frequency locations, as well as the city as a whole, there is no statistically significant difference between any of the smoothing methods, with or without seasonality or pooling. Results are mixed for the lower frequency precincts. All precincts experienced a downward trend in burglaries during the forecast period. Precincts had 23 percent fewer monthly burglaries on average during the forecasting period than during in the estimation period (see Table 5). The greatest declines were in Precincts 5 and 6. As seen in Table 8, seasonal adjustment of burglaries did not reduce forecast errors. This is likely attributable to conflicting temporal patterns in commercial versus residential robberies, as is suggested by differences in seasonality between Precinct 2 and 5. Precinct 2, containing the central business district, is more dominated by commercial activity, explaining its positive seasonal indices in December and January, during the Christmas season, as compared to Precinct 5 which has negative or zero seasonality during these months.

Overall, the forecast errors for robbery are in the medium range, between 17 and 24 percent. Exponential smoothing, without the presence of seasonality, is uniformly the best method, with significantly lower MAPE's than any Naive method. The 12 percent MAPE at the city level, using exponential smoothing, is a 60 percent improvement over Naïve 1, Lag 12, the method currently being used by police for counterfactual forecast. In Precincts 3 and 4, which have the lowest data frequencies, smoothing with pooling has the lowest MAPE, but the difference from exponential smoothing is not statistically significant.

Table 4

Mean Absolute Percentage Error Forecast Accuracy for
One-Step-Ahead Monthly Forecasts :
Burglaries in Pittsburgh for 1996, 1997, 1998

Method	Precinct						
	1	2	3	4	5	6	City
Naïve 1	1.2**	1.1	1.3**	1.4***	1.0	1.2***	1.1
Naïve 1 Lag 12	1.6***	1.3**	1.5***	1.7***	1.4***	2.2***	1.6***
Naïve 2	1.5***	1.2	1.4***	1.4***	1.2**	1.3**	1.2**
Naïve 2 Pooled	1.2*	1.2*	1.3**	1.5***	1.1	1.2*	1.2**
EXPO No Seasonality	1.0	1.0	1.1	1.2**	1.0	1.0	1.0
EXPO Deseasonalized	1.2**	1.1	1.0*	1.1	1.1	1.1	1.0
EXPO Deseasonalized Pooled	1.2***	1.0	1.0	1.0	1.0	1.1	1.0
HOLT No Seasonality	1.1*	1.0	1.2**	1.1	1.0	1.1	1.1**
HOLT Deseasonalized	1.2**	1.1*	1.1*	1.0	1.0	1.2	1.1
HOLT Deseasonalized Pooled	1.2***	1.1	1.0	1.0	1.2**	1.1	1.1
Minimum MAPE	17%	18%	22%	18%	24%	19%	12%
Average Monthly Burglaries	47	50	29	37	63	58	284

Significance (Wilcox Rank Pairs)   * ** p< .001   ** p< .01   * p< 0.1

Table 9 shows that the forecasts for 911 Drug Calls at the city and precinct level are most often improved by deseasonalizing the data using classical decomposition and by exponential smoothing. The Naïve methods perform poorly. The MAPE at the city level when using exponential smoothing is 11 percent, which is four percentage points lower than the MAPE for the Naïve 1 Lag 12 method (15 percent). Gains from pooling are modest, due to low seasonal indices and homogeneous seasonality, as seen in Figure 3. When smoothing is used, pooling improves the forecast by an average of eight percent. The gains from smoothing are more important: Smoothing methods have MAPE's which are on average 17 percent lower than Naïve methods. At the city level, classical decomposition results in errors on average 17 percent lower than . There is no significant difference between using simple exponential smoothing or Holt's method

Table 9

Mean Absolute Percentage Error Forecast Accuracy for
One-Step-Ahead Monthly Forecasts :
911 Drug Calls in Pittsburgh for 1996, 1997, 1998

Method	Precinct						
	1	2	3	4	5	6	City
Naïve 1	1.2**	1.1*	1.3**	1.3**	1.2*	1.4**	1.2**
Naïve 1 Lag 12	1.3**	1.2**	1.4**	1.8***	1.6**	1.4**	1.4***
Naïve 2	1.1	1.3***	1.3***	1.4**	1.8***	1.3**	1.2**
Naïve 2 Pooled	1.0	1.2*	1.2**	1.3*	1.4***	1.1	1.2**
EXPO No Seasonality	1.2**	1.0	1.2*	1.2*	1.0	1.2*	1.2*
EXPO Deseasonalized	1.0.	1.0	1.0	1.1	1.2*	1.1	1.0
EXPO Deseasonalized Pooled	1.1	1.0	1.0	1.0	1.2	1.0	1.0
HOLT No Seasonality	1.2**	1.1*	1.2*	1.3*	1.1*	1.1*	1.3**
HOLT Deseasonalized	1.0	1.1*	1.1*	1.2*	1.2*	1.1	1.0
HOLT Deseasonalized Pooled	1.1	1.0	1.0	1.0	1.1*	1.0	1.0
Minimum MAPE	18%	18%	24%	23%	18%	21%	11%
Average Monthly 911 Drug Calls	60	103	42	27	102	30	364

Significance (Wilcox Rank Pairs) * ** p< .001 ** p< .01 * p< 0.1

Results summary

The seasonality of crimes where data frequencies are low is difficult to calculate accurately at the precinct level. Unlike other small data aggregations (e.g., products) that have variation dominated by pattern changes, variation in small-area crime counts seems to be dominated simply by randomness. It is apparent that Classical Decomposition does not always provide accurate estimates of seasonality in smaller data aggregations (i.e., in low crime precincts). Simple Assaults in Precinct 6 exemplifies the type of problem that can emerge in the calculation of seasonal indices in small data series, and thus increase the error in forecasting using the deseasonalized series. In Precinct 6 the effect of one extreme and unusual low frequency month, June 1995, disturbed the results of seasonal forecasting methods: this was the only Precinct where deseasonalization of the simple assault data did not help the forecast. This problem could be addressed by either trimming outliers prior to calculating the seasonal indices or else dropping the smallest and largest values in the Classical Decomposition calculations to eliminate the possible effects of unusual events. However, the results from using city level pooling suggest that it may be more effective to increase scale using city-level estimates than it is to account for spatial heterogeneity.

In Table 10 we compare the results of the ten models for all 30 individual time series, eliminating methods that were significantly worse than the "best" method. The remaining methods have been sorted from best to worst (top to bottom), based on a rank count from left to right. Comparing methods using classical decomposition based on the precinct's data, with the pooled city models (CSI), we find that total-sample pooling reduced the forecast errors more often than any other method. There are exceptions, such as simple assaults in Precincts 1 and 6, and burglaries in Precinct 1. However, Table 10 shows that exponential smoothing pooled was the best method in 24 out of 30 precinct forecast series (80 percent of the cases, closely followed by Holts with Pooling).

The results for smoothing versus Naïve methods provide the most interesting insights for practical implementation, for use in the New York City approach to enforcement (Computer Statistics or COMPSTAT). What stands out is that all of the Naive methods, irrespective of adjustment for seasonality or pooling, perform significantly worse than smoothing methods for all crimes. Our results show common police practices to perform poorly, and simple exponential forecasting models to perform relatively well. In Table 10 we see that Naïve 1 Lag12, the method currently used by police to provide a counterfactual, was the worst method. In all but three out of 30 cases (10 percent), Naïve 1 Lag 12 resulted in significantly higher errors than other methods. There is much to be gained by smoothing the series. In particular, the absolute percent errors from the Naïve 1 Lag 12 method were on average 37% worse than the best method for drugs, 62% worse for burglary, 53% worse for simple assault, 47% worse for robberies, and 40% worse for aggravated assaults.

Table 10

Summary of Best, or Equivalently Best Methods (Lowest APE's)
for Forecasting Crimes in Pittsburgh PA, Based on 36 One-Month Ahead Forecasts, 1996 to 1998

	Drug Calls						Simple Assaults						Robberies						Burglaries						Aggravated Assaults						
Precinct	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	**Ranked**
EXPO Deseasonalized Pooled (EP)	EP	EP	EP	EP	EP	EP	EP		EP	EP		EP	EP	EP		EP	EP	EP		EP	EP	EP	EP	EP		EP	EP	EP	EP		24 **Best**
HOLT Deseasonalized Pooled (HP)	HP	HP	HP	HP		HP	HP	HP	HP	HP	HP	HP	HP	HP	HP	HP	HP			HP	HP	HP		HP		HP		HP	HP		23
EXPO No Seasonality (E)		E	*		E							E	E	E		E	E		E	E	E		E	E	E	E	E	E	E	E	18
EXPO Deseasonalized (ED)	ED	ED	ED	ED		ED				ED		ED		ED		ED	ED	ED		ED		ED	ED	ED				ED	ED		17
HOLT Deseasonalized (HD)	HD		HD			HD	HD		HD	HD	HD	HD	HD	HD		HD	HD					HD	HD	HD				HD			16
HOLT No Seasonality (H)												H	H	H	H	H	H			H		H	H	H	H		H		H	H	15
Naïve 1 (N1)												N1								N1			N1			N1			N1		5
Naïve 2 Pooled (NP)	NP					NP						NP											NP						NP		5
Naïve 2 (N2)	N2																			N2									N2		3
Naïve 1 Lag 12 (NL)																NL											NL				2 **Worst**
Minimum MAPE	18	18	24	23	18	21	12	8	11	13	10	11	34	14	46	35	22	26	17	18	22	18	24	19	39	29	47	28	23	34	
Average Monthly Crimes	60	103	42	27	102	30	153	201	105	125	191	121	23	43	14	11	38	27	47	50	29	37	63	58	14	20	9	10	23	9	

Color Key:

* Omitted cells represent cases where the APE's were significantly higher for the corresponding method, at > 90% confidence, using Wilcox Ranked Pairs Test.

(white) Method Uses No Classical Decomposition

(light grey) Method Uses Classical Decomposition, Seasonal Indices (SI's)

(dark grey) Method Uses Classical Decomposition, City Level Pooling of Seasonal Indices (CSI's)

Note: Forecasts for all exponential smoothing methods (E, H, ED, HD, EP, HP) were based on a 60-month (rolling horizon) estimation period, beginning with January 1991 to December 1995. Absolute Percent Errors (APE's) were calculated by comparing the forecast with the holdout sample.

Forecast Error Analysis

To investigate determinants of crime forecast errors, we assembled the 36 forecast APE's from the exponential smoothing pooled method for each of the five individual crime types and for each precinct. This provided a total sample of 1,080 absolute percent errors. We selected exponential smoothing pooled errors because this method produced very good results overall. Figure 4 is a scatter plot of these 1,080 APE's versus the average monthly crime count of each estimation series. The plot suggests an inverse relationship between these two variables.

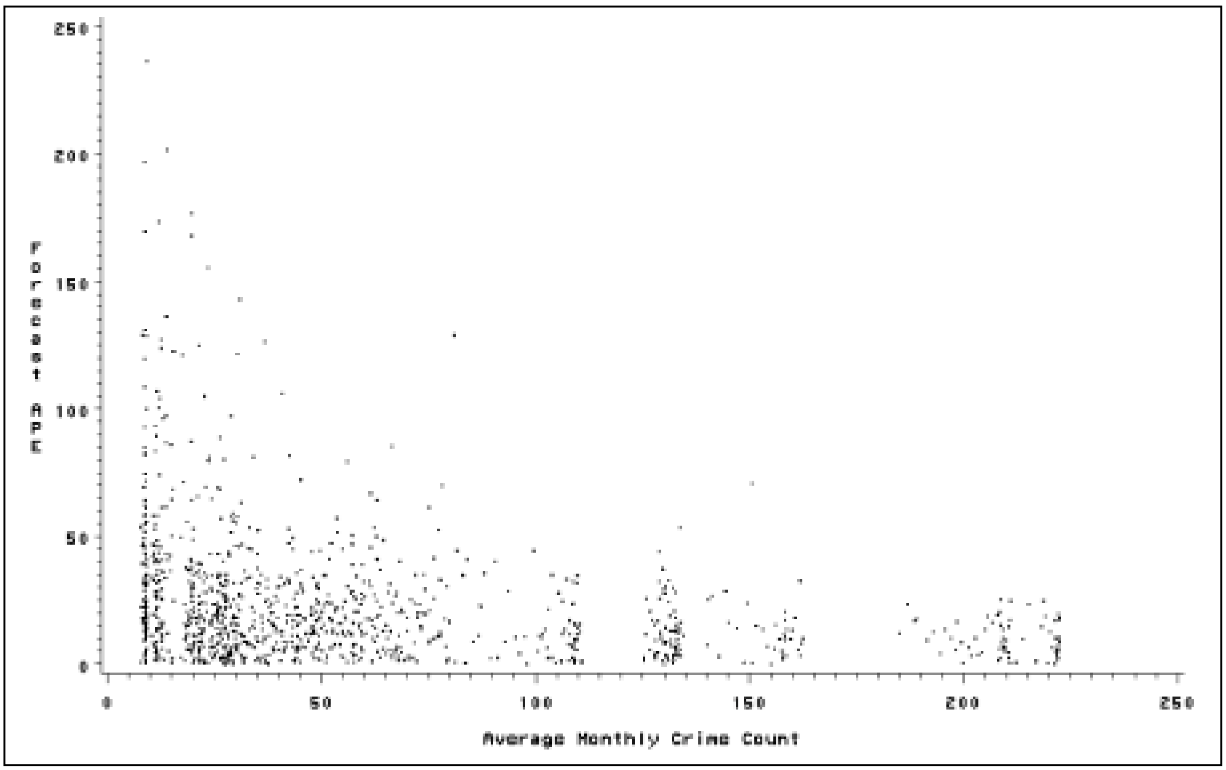

Figure 4. Absolute Percent Errors (APE's) from Monthly Exponential Smoothing Pooled Forecasts (1996-1998) of Pittsburgh Precinct Crimes, plotted against Average Crimes per Month in the Five Year Estimation Period.

In seeking gain some insight into what is driving the forecast errors, we ran a fixed-effects regression of the 1080 APEs against a number of explanatory variables. The independent variables included:

1/n	= inverse of average crime count, calculated as the average number of monthly crimes in each of the estimation series.
Seasonality	= amount of seasonality present: expressed as the average variance of each set of seasonal indices.
Trend	= absolute value of the linear time trend slope in each estimation period.

Also, we included fixed effects dummy variables for location (precinct) with Precinct 6 suppressed, and fixed effects dummy variables for crime-type with burglaries suppressed. Table 11 summarizes the results. The model has a significant F-Value of 11.4, but large unexplained variance and a low R^2 of 0.11. Only the constant, 1/n, and the dummy variable for simple assaults are significant, and all have the expected signs.

The 1/n and trend variables are correlated at 0.73; level and seasonality are correlated at 0.57; and trend and seasonality are correlated at 0.40. In formal testing for multicollinearity[6] we found the highest variance inflation index for 1/n was 9.6. In stepwise regression to add variables one at a time, 1/n has the largest independent t-stat and is chosen first (followed by trend). For all practical purposes the additional variables add very little to the model, and most of the explanatory power is in 1/n. In Table 12 the model has been re-estimated dropping all independent variables except the constant and 1/n.

[6] A variance inflation factor (VIF) measures how much the variance of an estimated regression coefficient increases if your predictors are multicollinear. Montgomery and Peck (1980) suggest that a VIF > 5-10 might warrant concern.

Table 11

Estimates on Absolute Percent Errors from Exponential Smoothing with Pooling Method – Simple Regression

Source	Sum of Squared Errors	Degrees of Freedom	Mean Squared Error
Model	117707	12	9809
Residual	914880	1067	857
Total	1032586	1079	
Adj R^2 0.11	**F-Value 11.44*** **	**Root MSE 29.28**	**# of Obs. 1080**
Variable	**Coefficient**	**Standard Error**	**t-Value**
Constant	14.1	4.8	2.9***
1/N	202.5	83.1	2.4***
|Trend|	27.8	32.2	0.86
Seasonality	2.4	9.7	0.25
Dummy Variables For Location			
Precinct 1	3.2	3.3	0.99
Precinct 2	-2.1	3.4	-0.60
Precinct 3	5.2	3.3	1.60
Precinct 4	-2.8	3.3	-0.85
Precinct 5	0.8	3.5	0.22
Dummy Variables For Offense Type			
Drugs	-0.5	3.3	-0.15
Robberies	4.6	3.6	1.26
Aggravated Assaults	0.4	5.6	0.07
Simple Assaults	-5.2	3.0	-1.74*

Significance Levels: * =0.01, **=0.05, ***=0.01

Table 12

Estimates on Absolute Percent Errors from Exponential Smoothing with Pooling Method – Simple Regression

Source	Sum of Squared Errors	Degrees of Freedom	Mean Squared Error
Model	98279	1	98279
Residual	934306	1078	866
Total	1032586	1079	
Adj R^2 **0.10*** **	***F-Value*** **113.4*** **	***Root MSE*** **29.4**	***# of Obs*** **1080**
Variable	**Coefficient**	**Standard Error**	**t-Value**
Constant	13.5***	1.36	9.95
1/N	287***	29.99	10.65

Significance Levels: * = 0.1, ** =0.05, *** =0.01

In Figure 5 we show the graph of the regression in Table 12. In general, roughly 30 monthly offenses are needed to forecast with any reliability. Crime-type and level of the data series is critical in obtaining accurate forecasts. The curve shows large gains in forecast accuracy (declines in APE) when going, for example, from a monthly crime count of ten to 30 crimes. At 10 crimes per month the expected forecast absolute percent error (APE) is 42 percent. Moving to 20 crimes the expected forecast APE has decreased by 14 percentage points, to 28 percent, and at 25 crimes per month the expected forecast APE is 25 percent. After 30 crimes per month we have rounded the knee of the curve and the gains from increasing the sample size move more gradually along the asymptote of 13.5 percent error.

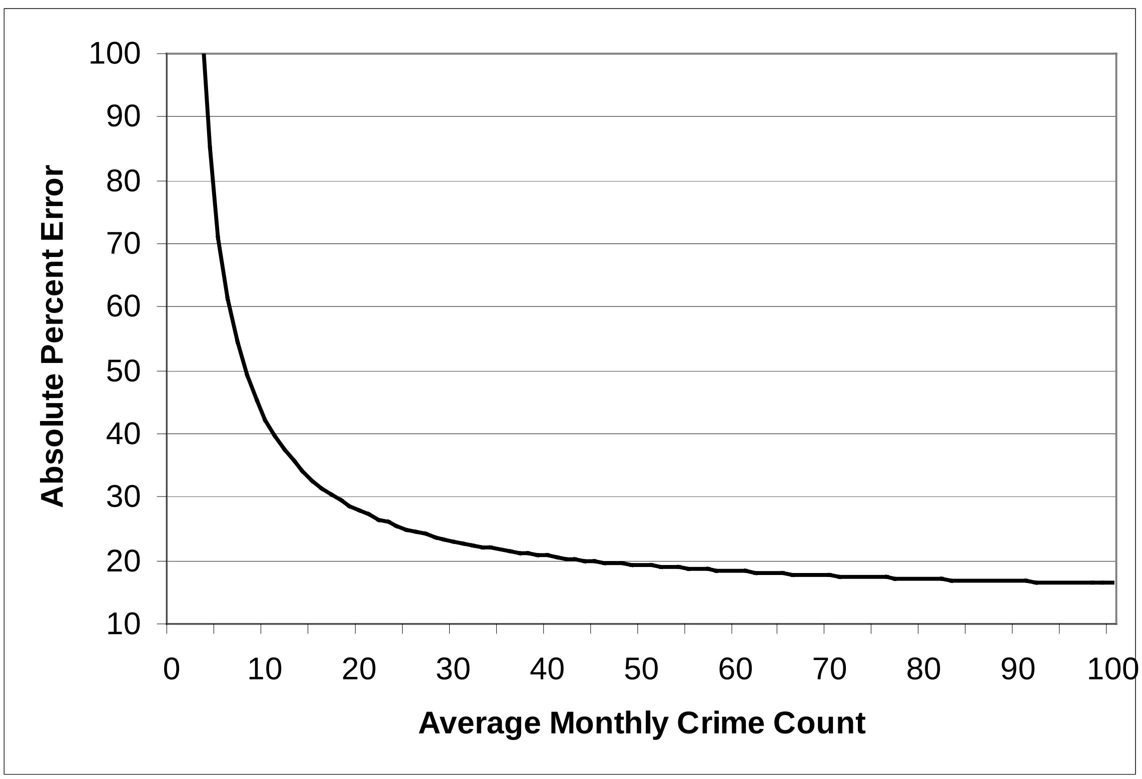

Figure 5. Relationship between Number of Crimes and Forecast Error (APE=13.7+216/n)

6. Summary and Future Work

The major limitation remains that small-scale data aggregations are subject to randomness and thus forecast accuracy is largely a function of the crime counts in the series. The small-scale characteristic of crime forecasting is critical. It is necessary to have monthly crime counts of 30 or higher by spatial unit (precinct or car beat) to approach acceptable accuracy (20 percent absolute forecast errors), and this result is independent of crime type or location in our case study. Some crime types, such as simple assaults, burglaries and drugs, have large enough counts that they can be extrapolated successfully in the short-run in precincts using univariate methods. These crimes have errors in the range of eight percent to 24 percent, accurate enough for use in deployment, scheduling. Police use counterfactuals as the basis of comparison for judging whether or not there has been a change in crime. We find that a prevalent police practice for evaluating changes in crime levels in small areas gives very poor results, given the volatility of small data aggregates.

Unlike non-crime forecasting experience (Withyscombe 1989, Bunn and Vassilopolous 1999) we did not find evidence of large pattern changes, but find that much of the variation in data is simply randomness, perhaps due to the large number of individual actors involved. Seasonality may be too small in many areas to identify in relationship to seemingly random historical disturbances or "shocks". Our case study suggests that data pooling, combining data at the city level to estimate seasonality at the precinct level, is a partial solution to the problem of small crime counts. Smoothing methods with pooled estimates for seasonal factors are most often the best univariate methods. Given the promising overall performance of pooling in the univariate methods tested, it appears that for forecast accuracy it is more important to increase scale (use city-level seasonality estimates) than it is to account for spatial heterogeneity (i.e., use precinct level seasonality estimates) when estimating seasonality. In high crime precincts, local seasonality does provide useful insights for monthly scheduling officers for peak and tough times, particularly for observable crimes such as drug offenses, assaults and robberies. For burglaries, insight into localized seasonality might more effectively be addressed by localized public information campaigns prior to peak months.

There is potentially large room for improvement using multivariate approaches in less frequent crimes such as aggravated assault and robbery, and for forecasting in smaller units such as beats. Our research has established benchmark values of expected absolute percent errors, based on the exponential smoothing with pooling method, of 12 percent for simple assaults, 21 percent for burglaries and drugs, 36 percent for aggravated assaults, and 32 percent for robberies. We have also established that below an average of 30 crimes per month in the unit of aggregation the forecast error becomes unacceptably high. In future research we will use the same data and geographic units to measure and compare performance of leading indicator and other sophisticated methods. We propose to continue to expand this work and replicate the study to test more complex forecasting techniques. We will use the results in this paper as a baseline, and using the same time periods and dependent variables, to compare Bayesian vector autoregression (BVAR) models (Litterman 1980, Doan, Litterman and Sims 1984, Le Sage 1989, LeSage and Pan 1990, 1995) and Olligschlaeger's Neural network model, Chaotic Cellular Forecasting (CCF) (Olligschlaeger 1997, 1998).

References

Anderson, C. A. 1987. Temperature and aggression: Effects on quarterly, yearly and city rates of violent and non-violent crimes. *J. Person. Soc. Psychol.* 52, no. 6: 1161-73.

———. 1989. Temperature and aggression: Ubiquitous effects of heat on occurrence of human violence. *Psychol. Bulletin* 106, no. 1: 74-96.

Armstrong, J. S., and F. Collopy. 1993. Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting* 8: 69-80.

Baron, Robert A. 1972. Aggression as a function of ambient temperature and prior anger arousal. 183-89.

Blau, P. M. 1977. *Inequality and heterogeneity: A primitive theory of social structure*. NY: Free Press.

Block, C. R. 1983. *How to Handle Seasonality*. Chicago: Illinois Criminal Justice Authority.

Block, C. R., and R. L. Block. 1988. Violent crime and seasonality: Surveillance and police notification . *Illinois Criminal Justice Information Authority.*

Boggs, Sarah L. 1966. Urban Crime patterns. *American Sociological Review* 30: 899-908.

Brantingham, P., and P. Brantingham. 1984. *Patterns in Crime*. Wadsworth.

———. 1975. Spatial patterns of burglary. *Howard Journal of Penology and Crime Prevention* 14.

Brown, R. G. 1963. *Smoothing Forecasting and Prediction of Discrete Time Series* . Englewood Clifffs, NJ: Prentice Hall .

Bunn, D. W., and A.I. Vassilopoulos. 1999. Comparison of seasonal estimation methods in multi-item short-term forecasting. *International J. of Forecasting* 15: 431-43.

Chapin, F. S. Jr. 1974. *Human Activity Patterns in the City*. NY: John Wiley & Sons.

Cheatwood, D. 1988. Is there a season for homicide? *Criminology* 26, no. 2.

Cohen, J. 1941. The geography of crime. *Annals*: 217.

Cohen, J, D. Cork, J. Enberg, and G. Tita. 1998. The role of drug markets and gangs in local homicide rates. *Homicide Studies* 2, no. 3: 241-62.

Cohen, L. E., and M. Felson. 1979. Social change and crime rate trends: A routine activity approach. *American Sociological Review* 44: 588-607.

Deutch, S. J. 1978. Statistic models of crime rates. *International J. of Comparative and Applied Criminal Justice* 2: 127-51.

Doan, T. R., B. Litterman, and C. Simms. 1984. Forecasting and conditional projections using realistic prior distributions. *Econometric Reviews* 3: 1-100.

Duncan, G., W. Gorr , and J. Szczypula . 1993. Bayesian forecasting for seemingly unrelated time series . *Application to Local Government Revenue Forecasting Management Science* 39: 275-93.

Duncan, G., W.L. Gorr, and J. Szczypula. 1999. *Forecasting Principles*. to appear in: Kluwer Academic Publishers.

Engberg, John. 1999. The spatial dynamics of urban violence and employment. *Heinz School of Public Policy and Management Working Paper*.

Falk, J. J. 1952. The influence of the seasons on the crime rate. *J. of Crime and Law Criminology* 43: 199-213.

Felson, M. 1994. *Crime and Everyday Life*. Thousand Oaks: Pine Forest Press.

Felson, Marcus, and Ronald V. Clarke. 1998. *Opportunity makes the thief: Practical theory for crime prevention*. Policy Seris Paper 98 ed. London: Policing and Reducing Crime Unit, Research Development and Statistics Directorate.

Ferri, Enrico. 1882. Das Verbrechen in seiner Abhaangigkeit von dem jahrlichen tempuraturwechsel.

Guerry, A. M. 1833. *Essai sur la statistique moral de la France*. Cologne: Bohlanverlag. Quote in M. Boehme, Die Moral Statistik, 1971.

Harries, K. D., and S.J. Stadler. 1986. Aggravated assault and the urban system: Dallas. *J. Environmental Systems* 15, no. 3: 243-53.

———. 1983. Determinism revisited: Assault and heat stress in Dallas, 1980. *Environment and Behavior* 15, no. 2: 235-56.

———. 1993. Determinism revisited: Assault and heat stress in Dallas, 1980. *Environment and Behavior* 15: 235-56.

———. 1988. Heat and violence: New findings from Dallas field data. *J Applied Social Psychology* 18: 128-38.

———. 1984. Seasonality and assault: Explorations in inter-neighborhood variation, Dallas 1980. *Annals of the Association of American Geographers* 74: 590-604.

Hirschfield A., and J. Bowers. 1997. The effect of social cohesion on levels of recorded crime in disadvantaged areas. *Urban Studies* 34, no. 8: 1275-95.

Holt, C. C. 1957. *Forecasting Seasonality and Trends by Exponentially Weighted Moving Averages*. Pittsburgh: Carnegie Institute of Technology.

Hylleberg, S. ed. 1995. *Modelling Seasonality*. Oxford: Oxford University Press.

Jeffeies, E., J. LaVigne, J. Szakas, C. Nahabedian, L.Mazerolle, T. Conover, K. Harries, D.

Williamson, N. Levine, R. Langworthy, J.DeVoe, L. Groff, and P. Canter. 1998. A multi-method exploration of crime 'Hot Spots'. *Crime Mapping Research Center* (project in process).

Landau, S. F., and D. Pfeffermann. 1988. A time series analysis of violent crime and its relation to prolonged states of warfare: The Israeli case. *Criminology* 26: 489-504.

Landau, Sima F., and Daniel Fridman. 1993. The seasonality of violent crime: The case of robbery and homicide in Israel. *J. of Research in Crime and Delinquency* 30, No. 2: 163-91.

LeBeau, J. L. 1988. Comment - weather and crime. *Justice Quarterly* 5: 301-9.

———, Principle Investigator, Grant #97-LB-VX-K010. 1998. *Demonstrating the analytical utility of GIS for policing: Moving beyond the descriptive*. Carbondale: Southern Illinois University at Carbondale.

LeSage, J. P. 1989. Incorporating regional wage relations in local forecasting models with a bayesian prior. *International Journal of Forecasting* 5: 37-47.

LeSage, J. P., and Z. Pan. 1995. "Using spatial contiguity as bayesian prior information in regional forecasting models." .

Lewandowski, R. 1982. Sales forecasting by FORSYS. *Journal of Forecasting* 1: 205-14.

Litterman, R. B. 1986. Forecasting with Bayesian vector autoregressions - Five years of experience. *Journal of Business & Economic Statistics* 4:1.

Makridakis, S., A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, and R. Winkler. 1982. The accuracy of extrapolation (time series) methods: Results of forecasting competition. *International Journal of Forecasting* 1: 111-53.

Makridakis, S., and R.L. Winkler. 1983. Averages of forecasts: Some empirical results. *Management Science* 29: 987-96.

Makridakis, S., and S.C. Wheelright. 1978. *Interactive Forecasting: Univariate and Multivariate Methods*. SanFrancisco: Holden-Day.
2nd Ed.

Makridakis, S., S.C. Wheelright, and V.E. McGee. 1983. *Forecasting: Methods and Applications*. Chichester: Wiley. 2nd ed.

Miller, Tan. 1993. Seasonal exponential smoothing with damped trends: An application for production planning. *International J. of Forecasting* 9: 509-15.

National Institute of Justice. 1998. "Solicitations for policing research and evaluation: Fiscal year 1998." *Section VI: The Impact of Technology on Policing; Part B: Developing Predictive Models*.

Newman, O. 1972. *Defensible Space: Crime prevention through urban design*. NY: Macmillan.

Olligschlaeger, A. M. 1997. Artificial neural networks and crime mapping. *Crime Mapping, Crime Prevention.* D. Weisburd, and T. McEwen (eds)Money, NY: Criminal Justice Press.

———. 1997. "Spatial analysis of crime using GIS based data: Weighted spatial adoptive filtering and chaotic forecasting with applications to street level drug markets." Carnegie-Mellon University.

Pokorny, A. D. 1965. Human violence: A comparison of homicide, aggravated assault, suicide, and attempted suicide. *J. of Criminal Law, Criminology, and Police Science* 56: 488-97.

Reiss, A. 1986. Why are communities important in understanding crime? *Communities and Crime.* A. Reiss, and M. Tonry (Eds.)Chicago: University of Chicago Press.

Repetto, T. A. 1974. *Residential Crime*. Cambridge, MA: Ballinger.

Rotten, J., and J. Frey. 1985. Air pollution, weather, and violent crimes: Concomitant time-series analysis of archival data. *J. of Personality and Social Psychology* 49: 1207-20.

Sampson, R. J. 1993. The community context of violent crime. *Sociology and the public agenda.* Ed. W.J.Wilson, 259-86. Newbury Park: Sage.

Sherman, L. W., P.R. Gartin, and M.E. Buerger. 1989. Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology* 27: 27-55.

1994. *J. of Quantitative Criminology* 10, no. 4: 317-42.

Wilson, J. Q., and G.L. Kelling. 1982. Broken windows: The police and neighborhood safety. *Atlantic Monthly* 249: 29-38.

Withycombe, R. 1989. Forecasting with combined seasonal indices. *International J. of Forecasting* 5: 547-52.

Zahn, Margaret A., and Katherine Jamieson. 1977. Changing patterns of homicide and social policy. *Homicide Studies* 1, no. 2