CRIMINAL CAREERS OF PUBLIC PLACES

by

William Spelman
University of Texas at Austin

Abstract: *Most police* calls *for service come from especially dangerous locations, or "hot spots." If risks at these locations are stable—hot spots stay hot—community problem-solving techniques may reduce crimes and disorders substantially. If locations run high risks only temporarily or sporadically, location-based strategies may not work. Analysis of calls for service at high schools, housing projects, subway stations and parks in Boston shows that risks remain fairly constant over time; most apparent changes may be attributed to random processes. Autoregression and displacement in space and call type are statistically significant but unimportant indicators of call rates. In addition to verifying the effectiveness of community problem-solving strategies, these results have practical implications for problem-solving techniques.*

CRIMINAL CAREERS OF PUBLIC PLACES

Over the past decade, research has confirmed what many criminal justice practitioners always knew: a few, particularly frequent, offenders are responsible for a disproportionate amount of crime (Chaiken and Chaiken, 1982; Horney and Marshall, 1991: Mande and English, 1988). Such findings have led to calls for "selective incapacitation," aimed at imprisoning the frequent few and—presumably—preventing the crimes in which they would have participated, had they been free (Blumstein et al., 1986: Zedlewski, 1987).

Nevertheless, three elements are required before a crime can be committed. Not only must someone be motivated to commit it, but a suitable target must be available, in a (relatively) unguarded location, providing the offender with an opportunity to commit the crime (Cohen and Felson, 1979). Recent research suggests that, like offenders, some victims and places are particularly crime-prone. Some individuals run

Address correspondence to: William Spelman, Lyndon B. Johnson School of Public Affairs. University of Texas. P.O. Box Y. Austin, *TX* 78713-7450.

especially high risks of becoming victims—people in dangerous jobs, for example, or women in abusive personal relationships (Nelson, 1984; Reiss, 1980). And some locations run especially high risks of being the site of victimizations (Pierce et al.. 1986; Sherman et al., 1989).

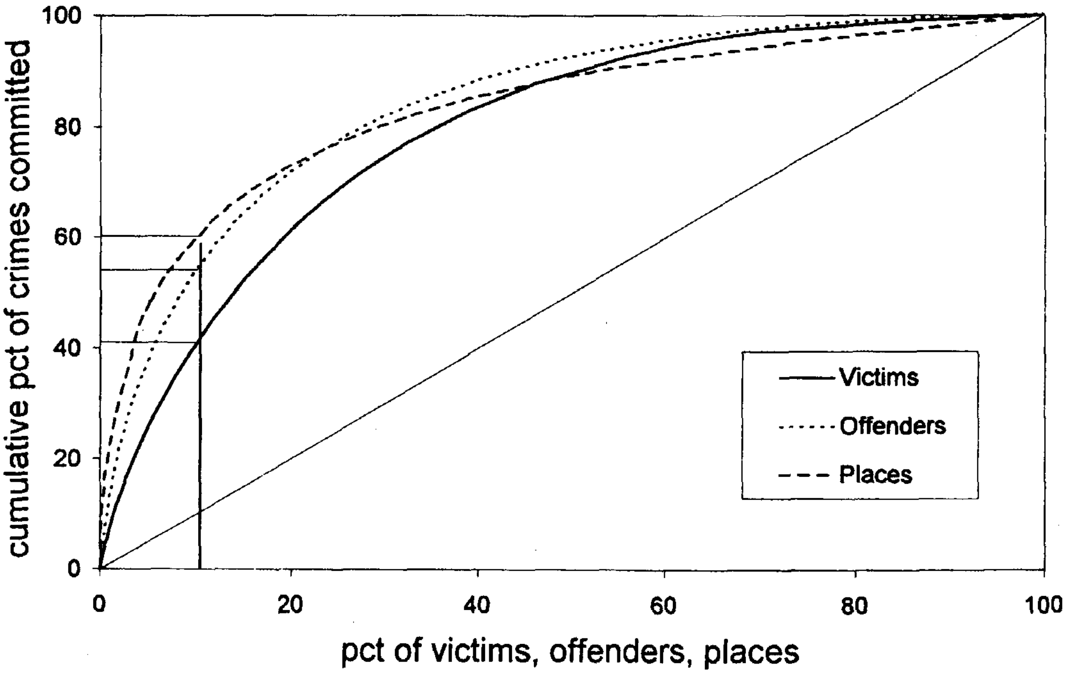
To measure the importance of these high-risk cases, suppose we took all the criminals active in a community and lined them up in order of the frequency with which they committed crimes. Those who committed crimes most often would go to the head of the line; those who committed crimes only occasionally would go to the end. If all offenders were alike, then it would not matter much where we would line each one of them up; the offenders at the front of the line would commit about as many crimes as those at the end. For example, the "worst" 10% of criminals would account for about 10% of all crimes. But if there were significant differences among offenders, those at the head of the line would account for far more than their share of all crimes committed; the worst 10% would account for much more than 10% of all crimes.

Figure 1 shows what happens if we conduct such an experiment with available data on repeat offending, repeat victimization and repeat calls for service. Surveys of jail and prison inmates suggest that the most frequent 10% of offenders are responsible for about 55% of all street crimes committed (Blumstein et al., 1986; Chaiken and Chaiken, 1982; Clarke and Weisburd, 1992). According to the National Crime Survey, the 10% of potential victims at highest risk account for over 40% of all victimizations (Nelson, 1984; see also Fienberg, 1980; Reiss, 1980). Calls-for-service data in Boston and Minneapolis show that the addresses producing the most repeat calls account for over 60% of all calls (Pierce et al., 1986; Sherman et al., 1989).

These "sitting ducks" and "dens of iniquity" not only form a symmetry with the "ravenous wolves" of frequent offending; they also have policy implications of their own. Many innovative police departments have adopted community- and problem-oriented approaches, in large part to deal with persistent victimization problems (Goldstein, 1990; Greene and Mastrofski, 1988). With the support of the U.S. National Institute of Justice, the Minneapolis Police Department even formed a special squad aimed at solving problems at high-risk locations (Sherman et al., 1989).

Like selective incapacitation, however, such "repeat-call" strategies work only if the crime- and disturbance-prone addresses would have remained crime- and disturbance-prone in the absence of police action. If the typical high-risk location only remains vulnerable for a few weeks or months, for example, then time-consuming police efforts to solve the problems that caused the vulnerability may be unnecessary: by the time the police have identified a solution, the problems would have solved

Figure 1: Ducks, Wolves, and Dens. Crime is Concentrated.



themselves. More generally, some locations may be permanently and predictably vulnerable, while others are vulnerable only temporarily. If police could distinguish between the two. this would help them to allocate their efforts effectively.

Three questions are particularly important for the development of effective policies toward high-risk locations:

- Do high-rate locations tend to remain high rate for long periods, or does the typical "high-crime criminal career" only last a few weeks or months?
- How big are the differences in long-run risks among locations? Are they sufficiently large that it makes sense to focus on the worst locations?
- Are the causes of these risks likely to be within or beyond public control?

Before addressing these questions, let us first consider alternative explanations for differences among locations and over time.

WHAT MAKES HOT SPOTS HOT?

Consider a typical police spot map. Each day, the previous day's crimes are marked on the map. Over the course of a month, some locations will have racked up many more spots than others—that month's "hot spots." Since the police department has limited resources, it might reasonably decide to assign only the hottest locations to community problem solvers. But depending on what makes each hot spot hot, such a strategy may be completely ineffective. To see why, consider four reasons why hot spots may get hot.

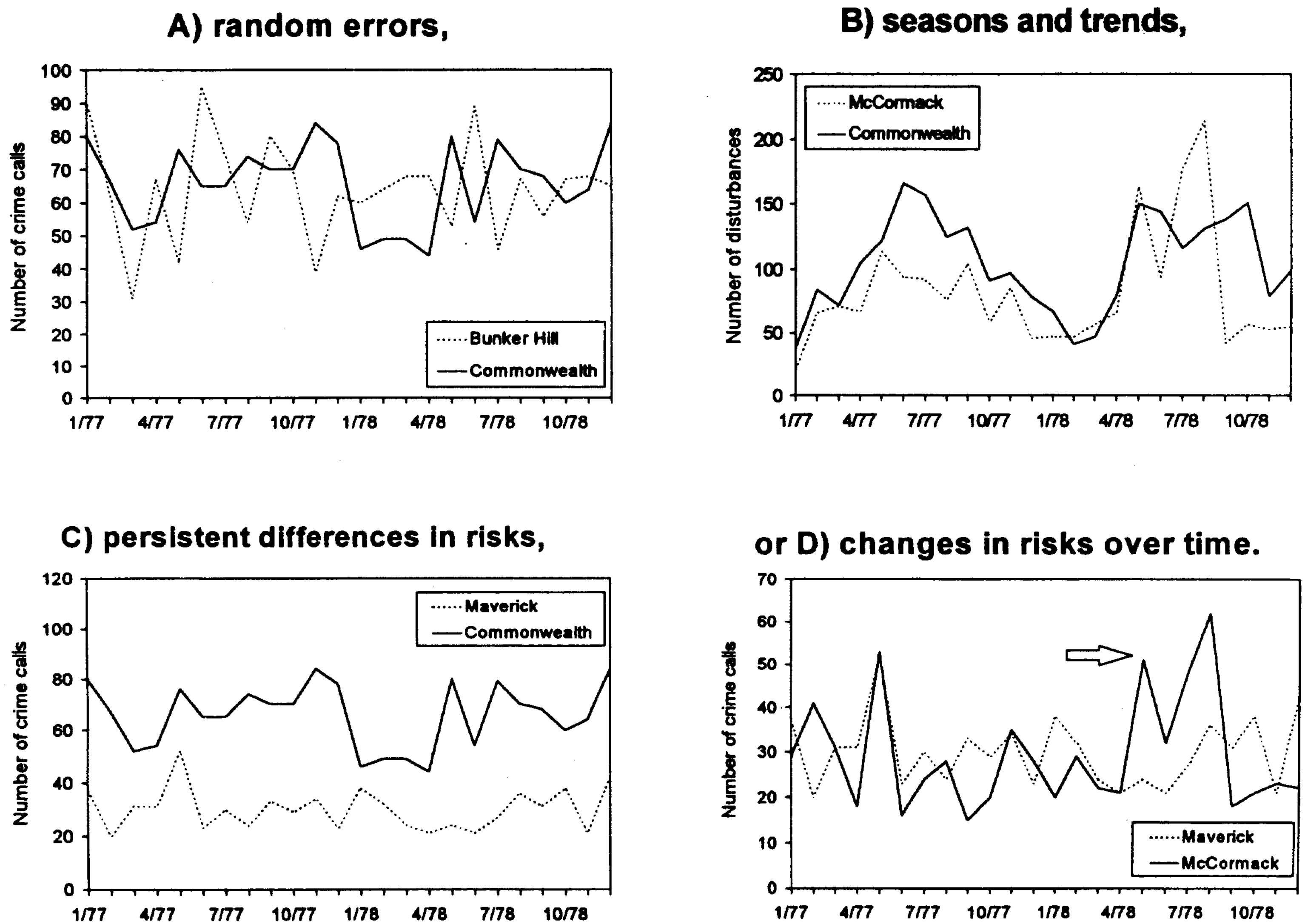
Random Error. Assembly lines are reliable; General Motors' Saturn plant cranks out a completed automobile every eight minutes. Not so the social processes that produce crimes, disturbances and other calls for service. As any patrol officer knows, there is a chance of a call coming in from anywhere, at any time. Even if some places and times are more likely to produce calls than others, the unpredictability remains.

Figure 2A shows how chance can make a location look hot. Over the long run, the same number of crimes were reported each month during 1977 and 1978 at Commonwealth as at Bunker Hill, two large Boston housing projects. Further, the average number of incidents reported each month did not change over time. Yet because incidents are random, the actual number reported each month *did* change over time; during some months, Commonwealth looked like the hot spot, while other months Bunker Hill looked hotter. But this is an illusion, caused by the random nature of crime.

Citywide seasons and trends can also make a location look hot for short periods. Particularly in a northern city such as Boston, we might expect more calls anywhere during the summer than the winter, and more calls during an especially dry and mild spring. In many cities the overall trend is toward more calls over time. So another reason a location may look especially hot at some time is that *all* locations are hot. In this case, it may be necessary to devote more than the usual resources to handle all of the incidents, but any attempt to solve the underlying problems would be doomed to imaginary success. The "problem" is nice weather or some other citywide condition that will not respond to local changes in the physical or social environment.

Figure 2B shows the number of disturbance calls reported for the Commonwealth and M.E. McCormack housing projects in Boston. Both have the expected pattern, peaking in the summer and bottoming out in

Figure 2: Hot Spots May be Due to:



January and February. At least for these projects, August's hot spots cool down quickly as winter approaches.

Persistent differences between locations. Some locations may repeatedly attract active offenders or potential victims, or allow them to come together in such a way that the offender cannot be controlled and the victim cannot be guarded. If the differences in risk are large and persistent enough, it makes sense to spend substantial resources to identify and eliminate the causes.

Figure 2C shows two time-series plots associated with consistently different locations. Here, Commonwealth is compared to Maverick Square, a smaller housing project in another section of town. Note that the number of crime calls at each location remains unpredictable from month to month, but that Commonwealth consistently reports more crime than Maverick.

Changes in locations over time. This is not necessarily so if hot spots do not stay hot for long. Consider the problem of rowdy youths in a park. The kids may make trouble for a month or two, then become bored and simply quit; thus the problem may solve itself without outside intervention. The police or neighborhood residents may persuade the kids to leave the park, displacing them to another location and making them someone else's problem. Or the kids may stay at the park but quietly begin selling drugs; thus the police and the public might trade a disturbance problem for a more serious crime problem.

Figure 2D compares crime calls in the Maverick Square and McCormack projects. During the average month, about the same number of crimes are reported in each; but the risks appear to increase in McCormack during the summer of 1978. Between May and August, the McCormack crime rate is nearly double that of Maverick Square, although it returns to normal in September. A police agency might well have identified McCormack as a problem project during one of the bad months, and any time spent solving problems at the project would have appeared successful by the end of the summer. But the success would be illusory, since the problem would have solved itself without police assistance.

Potential implications. As Figures 2A through 2D show, all of these four explanations are partially correct. But some are more correct, or correct more often, than others. Identifying which of the four are the most important would heip policymakers determine the importance of problem solving at hot spots.

If differences among locations and over time are mostly due to *random errors,* the whole notion of focusing police and citizen attention on the worst locations or times is wrongheaded. There are no worst and best places and times, no "problems" to be "solved" and nothing to be gained by focusing our resources.

If many of the differences can be accounted for by *seasons and trends,* citywide efforts are called for. We need more police officers, better community guardianship and more cautious potential victims at some times than others. It may help to focus our efforts on some locations, since this may be a more efficient means of getting the message across to the public. But all locations are alike, and changing the characteristics of these locations is itself unlikely to prove useful.

The opposite is true if most of the variation is due to *persistent differences among locations.* In this case, there is something about the individual place that produces more or fewer crimes, disturbances and persons in need of service. The key to call reduction lies in the identification and solution of very localized problems. Further, because risks are more or less fixed we can have some faith that once solved, problems will stay solved.

Finally, we must take a mixed approach if differences among locations are large but *changing over time.* Problems come and go; even if we do not take great pains to solve them, they will (eventually) solve themselves. This is especially true if spatial displacement appears to be an important cause of changes in risks. It is less true if high-risks locations are liable to remain higher than average for several months before returning to normal. In general, the implications depend greatly on the nature of the temporal changes, and more analysis is likely to be needed if this explanation is correct.

Given these general policy guidelines, let us now turn our attention to the empirical evidence.

DATA

Boston represents a particularly good test of these four alternatives. Until the early 1980s, the Boston Police Department (BPD) took an almost entirely incident-driven approach to policing. Swamped by 911 calls, BPD officers—like those in most jurisdictions—had little time to identify, analyze and solve problems on their beats. Although the focus has shifted over the last decade, the pattern during the late 1970s is probably much like that facing incident-driven police agencies nationwide.

The data set consists of calls for police service reported between January 1977 and December 1980 and recorded by the BPD's computer-aided dispatch system (Pierce et al., 1986). All calls for service were examined for:

- 35 public and private high schools,
- 35 public housing projects,
- 53 subway stations, and
- 135 parks and playgrounds.

Public places were chosen for several reasons. They are easy to identify. Each type is relatively homogeneous; for example, all high schools are similar enough that we can expect their hot-spot characteristics to be about the same. Because it is hard to restrict access to any of them, the environment should be especially important to problem solvers. Finally, over the four-year period each of these four types of public places produced far more than their share of calls for service.

Calls for service were chosen for study because they include a wider variety of problems than reported crime data, and many of the problems involving crime have serious costs for the police and the public. For example, nonviolent domestic disputes sometimes escalate into dangerous situations. In addition, juvenile disturbances may cause little tangible harm, but they increase fear of crime by creating an impression that things are out of control.

The nature of each call was recoded to fit one of three categories: crime, disturbance and service. This means that each category includes a wide variety of incidents; for example, assault, auto theft and drug dealing are all classified as crimes. If each of these crimes behaves differently—hot spots for auto theft are permanent while hot spots for drug dealing move around, for instance—our results may be misleading. Nevertheless, large categories were a practical necessity to keep the analysis from becoming too complicated. Table 1 shows the nature of the calls produced by each location type.

Finally, the time and date the call was received was recoded. Since many of the locations of each type generated relatively few calls, relatively long time periods were chosen for study—28 days. Each period thus includes the same number of weekdays and weekends, making them easier to compare.

Before continuing, it is important to recognize the limitations of this data set. Boston is a fairly dense, northern city with a good public transit network. Therefore, we can expect more evidence of seasonality and spatial displacement here than in other places. Simply because they are open to the public, we might expect the criminal careers of public places to be different from the careers of, say, single-family houses or office buildings.

Table 1: Nature of Calls Received by Location Type

	high schools	housing projects	MBTA stations	parks
Crime calls	2,822	33,109	2,352	11,200
	56.5%	26.6%	42.3%	43.1%
violent crimes	453	4,267	332	926
property crimes	550	7,903	389	3,172
vice and drug offenses	40	1,680	110	123
alarms and investigations	1,779	19,259	1,521	6,979
Disturbance calls	840	50,035	1,611	5,927
	16.8%	40.1%	28.9%	22.8%
Gang and juvenile disturbances	622	26,953	868	5,102
Other disturbances	217	19,542	707	785
Interpersonal disputes	1	3,540	36	40
Service calls	1,310	32,398	1,573	8,380
	26.2%	26.0%	28.3%	32.2%
Traffic accidents and violations	56	3,084	264	977
Injuries and illnesses	63	7,809	250	361
Fires	68	1,985	38	231
Other services to public	1,119	13,630	499	6,710
Internal police services	4	5,890	522	101
Unknown call type	26	9,125	30	498
	0.5%	7.3%	0.5%	1.9%
Total	4,998	124,667	5,566	26,005

We might also expect them to be different from retail stores or bars; even though open to the public, the private owners of these locations may respond more quickly and effectively than the police and the public to changes in risks. So we may extend these results to other places and problems only at some risk.

Method

The principal analysis method used here is often used in pooled time-series cross-sectional data sets, and is sometimes called "least squares with dummy variables" (Berk et al., 1979; Stimson 1985). Briefly, 12 separate analyses were conducted, corresponding to three crime types for each of four location types. For each combination of crime and location type, let X_{it} be the number of calls for service recorded at individual location i at time t. Let D_t refer to a vector of t dummy variables, each with a value of one for time t and zero otherwise; let D_i be a vector of i dummy variables, each with value one for location i and zero otherwise. Then we use ordinary least squares to estimate

Equation [1] $X_{it} = \Sigma_{it} [\beta_t D_t + \beta_i D_i] + \epsilon_{it}$

In equation [1], β_i and β_t are sets of coefficients that measure the differential risks associated with each i and t, and ϵ_{it} is an error term. Thus the β_t measure the citywide effects of seasons and trends, and the β_i measure the average long-run value of X_i, controlling for seasonality and trends. We may measure the relative importance of seasons and trends and of persistent differences as the R^2 obtained by entering each of the two sets of dummy variables into the equation separately. (Since the two sets are orthogonal, the order of entry does not matter.) We can measure the statistical significance of each of these differences through F tests.

Least squares with dummy variables is the simplest means of dealing with a pooled data set. Though other methods are statistically more efficient, they are computationally more difficult and require the assumption that the collective time and location effects—the β_is and β_ts—are Normal-distributed. As shown below, this assumption is simply wrong. As it happens, the gains in efficiency are unimportant for data sets of this size, even if the assumption is merited (Judge et al., 1985).

This yields a measure of the importance of two of the four explanations for hot-spot behavior. The contribution of a third, random processes, can also be obtained easily. Suppose for the moment that the risks of a call being produced are the same for all times and places, and denote the expected number of calls per month as λ. Then the number of calls produced each month will be Poisson distributed, with

Equation [2] $p(x) = \lambda^{x} e^{-\lambda} / x!$

for $x = 0, 1, 2$, and so on. Thus, even though all locations run the same risks over time, higher values of x will be obtained at some times and places simply through the luck of the draw.

The mean and variance of the Poisson distribution are both equal to λ. The variance is especially important for our purposes. Specifically,

Equation [3] $\lambda = \sigma^2_x = \Sigma_{it} (X_{it} - \lambda)^2 / it$,

and the total sum of squares is equal to

Equation [4] $SS_T = \sigma^2_x it = \lambda it$.

Thus, we may compare the variance observed in some data set and compare it to our Poisson expectation from equation [4]. If the empirical $SS_T > \lambda it$, call risks are more variable over time and among locations than expected. Either risks are changing over time, they differ among locations, or both.

The dummy variables of equation [1] will tell us which, if either, is true, but equation [4] is helpful for another reason: If we substitute the mean of X_{it} for λ, then $X_{it}it$ measures the expected variability due to random Poisson processes, even if every X_{it} is different. Thus, we may simply divide the empirical sum of squares by $X_{it}it$ to estimate the proportion of all variability that is due to random processes.

If these three explanations were sufficient to explain all the variation among locations and periods, there would be no variability left to explain. That is, the proportions of variance explained by seasons and trends, persistent differences among locations and random processes would sum to one. In addition, ϵ_{it}, the error term of equation [1], should have the following properties:

Poisson-distributed. It should be Poisson (or approximately Normal) distributed, with mean 0 and variance equal to X_{it}, the average risk over all units and periods.

No serial correlation. For each location *i*, the ϵ_ts should be uncorrelated with one another. Knowing that the number of incidents was especially high for some unit during one period would tell us nothing about that unit during ensuing periods.

No spatial correlation. For each period *t*, the ϵ_is should be uncorrelated with one another. For example, knowing that the number of incidents was especially high for one unit during some period should tell us nothing about any other unit during that period.

Finally, since we are considering a variety of incident types separately, the ϵ_{its} should also have the following property:

> *No intertype correlation.* For all incident types *j*, the ϵ_{its} should be uncorrelated with one another. For example, knowing that many crimes were committed at some location during a given month should tell us nothing about the number of disturbance and service calls received.

Unless the residuals have all these properties, unit risks must be changing over time.

Changes in risks complicate the policy implications considerably. Further, different kinds of short-run changes are liable to have different policy implications. The error-term properties described above suggest three systematic ways in which short-run risks can change, yielding different policy implications.

If risks are *serially correlated*, there is a reliable pattern of changes over time. For example, if a group of rowdy youths begins hanging out at a park, the rate of disturbance calls may jump; if the youths scare residents of the surrounding neighborhood, the disturbance rate may remain high, or increase further. In such a situation, high call rates tend to remain high, and the correlation between rates at times t and $t + 1$ will be positive. On the other hand, if the neighbors or the police chase the kids from the park within a few weeks of their arrival, high call rates at time t will lead to lower rates at $t + 1$, and the correlation will be negative.

We can measure the pattern of serial correlation by applying standard Box-Jenkins techniques to the ϵ_{ts} for each location. By looking for patterns in the autocorrelation and partial autocorrelation functions, it is not difficult to determine whether the residuals are white noise (no serial correlation), or whether they fit an autoregressive or moving average model best (Box and Jenkins, 1976; McCleary and Hay, 1980). If one such model is appropriate for most locations of a given type, we may add the necessary terms to equation [1] and measure the importance of these effects through the increase in R^2.

If risks are *spatially correlated*, there are contemporaneous shifts in call rates among two or more locations. The most familiar spatial pattern is displacement. For example, if the police persuade some rowdy youths to leave a neighborhood park, the youths may simply reconvene at another park nearby. If this happens only occasionally, there will be no discernible pattern between the call rates at the two parks; but if it happens frequently, the two parks' call rates will be negatively correlated. The implication is that when calls are down at park A, police and citizens should be looking for trouble at park B. Positive correlations suggest offender movement of a different kind: one group of offenders may be shuttling back and forth more quickly than the periodicity of the data, or two groups may be stimulating one another to action. For example, a feud between two street gangs may increase the risk of disturbances or violent crimes in the turf of each. In general, we would expect higher (absolute value) correlations between locations that are close together than those that are far apart. It seems unreasonable to expect much offender movement between parks that are located ten or 15 miles apart.

In practice, we may estimate the correlation between the residuals at each pair of locations and examine the relationship between these spatial correlations and distance. For example, we may estimate

Equation [5] $r_{it} = \alpha D_{it}^{\beta}$,

where D_{it} represents the distance between locations *i* and *t*, and λ and β are coefficients to be estimated. If the relationship is negative, as expected, we may take displacement into account by calculating for each location *I* and period *T* a weighted sum of the call rates at each of the other locations. If S_{IT} represents this spatial displacement measure, then

Equation [6] $S_{IT} = \Sigma_{it} \alpha D_{it}^{\beta} X_{it} - X_{IT}$.

Like the serial correlation adjustment, this measure could be added to equation [1]. The change in R^2 would be an estimate of the importance of spatial displacement. Since S_{it} is contemporaneous with the dependent variable, this would not be appropriate if our aim were to forecast call rates. Because we are only searching for patterns in a complex system, the direct approach will serve our needs just as well.

Correlations among locations may be due to factors other than displacement. They may also indicate unmeasured temporal effects that only affect some classes of locations. For example, the public pools all open and close on the same day each summer; calls at these pools are obviously highest during months when the pools are open, and perhaps during months when the weather is unseasonably warm. This would affect all pools equally, no matter how far apart, resulting in a non-trivial, reliable connection. But if we did not know about the opening and closing date regularity, and interpreted the high correlation among call rates at pools as evidence of displacement, we would be wrong. Although such connections may be important, the information needed to interpret the results (for example, information about which parks had public pools and when they opened) was not available. This may be a fruitful area for further research.

Finally, *intertype* correlations may be important. Negative associations among call types suggest that the users of the location are changing, or that their activities are changing. If old people replace teenagers at a park, we can expect fewer gang disturbance calls but more calls for medical assistance. As the temperature drops and snow begins to fall on a housing project, we might expect fewer public intoxication and disorderly conduct calls and more burglaries. This fluidity in call patterns suggests that even successful problem-solving efforts may not reduce the total calls for service workload, though they may shift it to more benign incident types. Positive correlations can be accounted for by overall changes in the number of users. For example, when school starts in September, there will be more users of the Boston University Massachusetts Bay Transit Authority (MBTA) station and more potential for all types of calls. Zero or positive intertype correlations provide fewer pitfalls for problem solvers than negative correlations.

Although intertype correlations could be accounted for with seemingly unrelated regression techniques (Zellner, 1962; Srivistava and Giles, 1987), a more direct approach is simply to add to equation [1] the number of calls received for each of the other two call types at each location during the same period. As with the spatial displacement correction, note that this would not be appropriate if our aim were to predict call rates.

RESULTS

The principal results are shown in Table 2. For each location, similar results were obtained for crimes, disturbances and services. Thus the table only shows the average percentage of variation explained over all call types. Summarizing the table, we find that

Random variation explains between 4 and 37% of the total variation among times and locations. It is especially important for locations such as subway stations and high schools that report fewer calls.

Seasons and trends are relatively unimportant, accounting for only 2 to 4% of the total variation.

Long-run differences among locations account for the largest source of variation in each case—34 to 83% of the total.

Short-run changes account for 11 to 28% of the total variation. Except among high schools, these changes are not nearly as important as long-run differences. Relatively little of this variation can be attributed to displacement in time, space and type of call.

Although the random nature of calls for service makes it difficult to identify the worst locations, most of the variation among times and locations is due to real differences in risks.

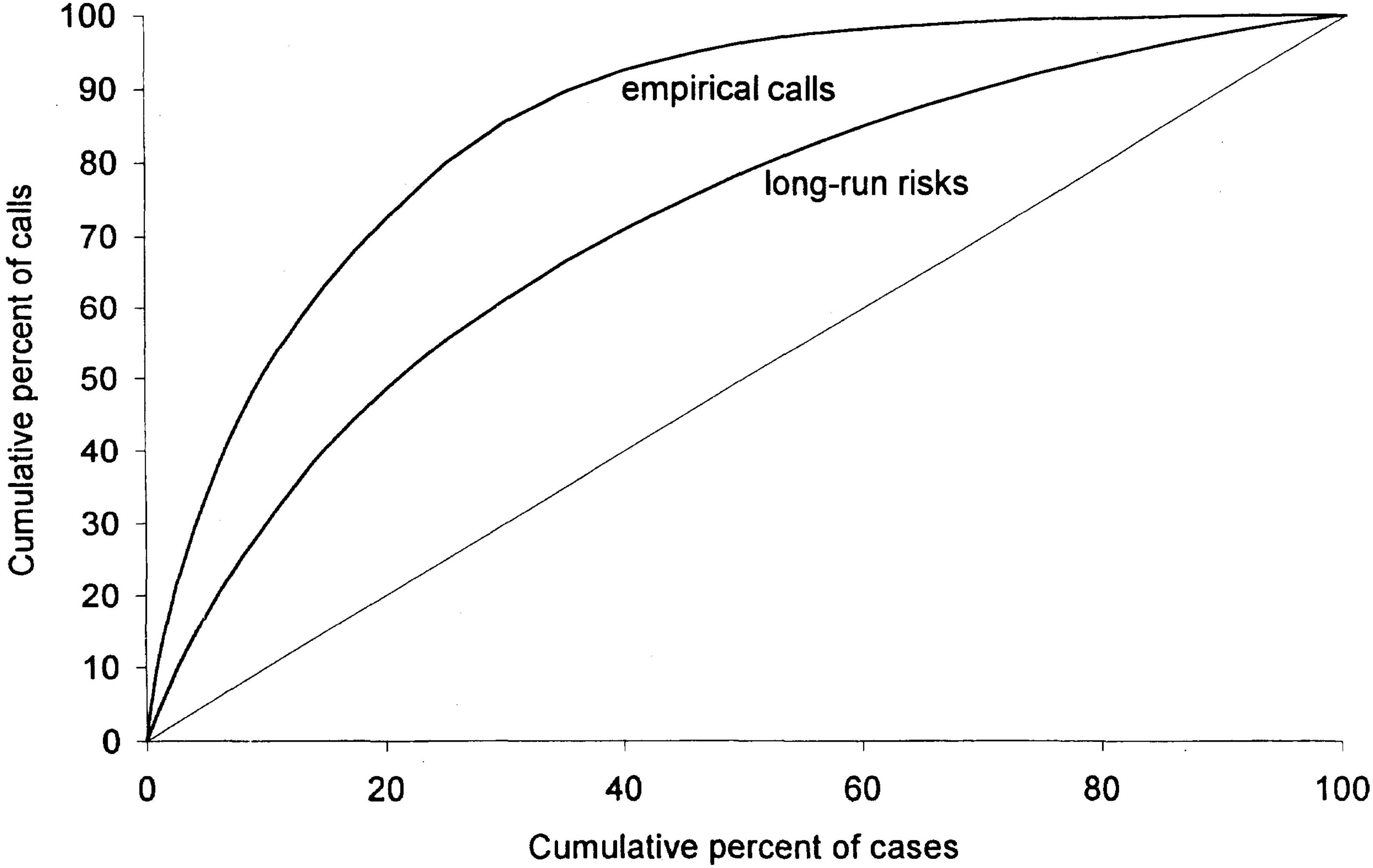
Table 2: Proportion of Variance Explained in Call Risks. Averaged Overall Call Types.

	high schools	housing projects	MBTA stations	parks and play-grounds	average
Random processes	.347	.038	.370	.150	.226
Seasons and trends	.040	.018	.017	.021	.024
Persistent differences among locations	.335	.832	.489	.671	.582
Short-run changes within locations	.278	.112	.124	.159	.168
autoregression	.003	.001	.001	.020	.006
spatial displacement	.028	.009	.010	.014	.015
intertype correlation	.028	.020	.005	.000	.013
all others	.219	.082	.108	.125	.134
Total	1.000	1.000	1.000	1.000	1.000

If we take random variation as a given and consider only non-random variation, the importance of each explanation is even clearer. Of the non-random explanations, long-run differences among locations are by far the most important, dwarfing the other components for three of the four location types considered. Although Table 2 also makes it clear that high schools are different from the others, long-run risk differences still account for over half the variation in these locations over time.

The practical importance of these findings is captured more directly in Figure 3, which compares the Pareto curves for the calls actually received to the Pareto curves for long-run differences, as measured by the J3is. Averaging over all location and call types, the worst 10% of locations and times account for about 50% of all calls for service (upper solid curve). Once random variation, seasons and trends, and changes in each location's risk over time have been filtered out, the worst 10% of locations

Figure 3: Concentration of Calls and Risks
All Location and Call Types



account for only about 30% of all calls (lower solid curve). If we equate "long-runs risks" with "problem-solving potential," we find that the potential is about half what it would appear at first glance. On the other hand, there is still substantial room for reduction of crime and incivilities.

Overall, then, it makes sense for the people who live and work in high-risk locations, and the police officers and other government officials who serve them, to spend the time they need to identify, analyze and solve their recurring problems. A few weeks' or even months' work is a prudent investment, since these problems are unlikely to go away by themselves in the foreseeable future.

Although they were not an important source of variation, correlations in time and space, and between types, are sufficiently important from a theoretical viewpoint to warrant describing them in more detail. Table 3 summarizes the results of tests for temporal displacement. (Results were significantly different from one location type to the next, but the patterns and general implications were quite similar.) The vast majority of time-series were indistinguishable from white noise; that is, there appeared to be no temporal correlation pattern at all. Most of the rest were clearly first-order models (autoregressive or moving-average models with a maximum lag of one period). Many of these were hard to classify—the autoregressive and moving-average models worked about equally well. Of those that were not ambiguous, autoregressive models almost always fit better than moving-average models. Note, incidentally, the number of series that were best fit by models involving three or more lags. In all of these cases, a third or higher-order autocorrelation was different from zero, but the first autocorrelation was not. Since such a result is inconsistent with any reasonable time-series model, we can reject these as random fluctuations.

In sum, there is no support whatever in these data for choosing a moving-average model, very little for choosing a second-order model and none for choosing a model of higher order. We may tentatively conclude that considering each of these series to be first-order autoregressive may not help much, but it is at least a reasonable possibility. As shown in Table 2, the addition of a one-period lag term (a simple, if inelegant, way of accounting for first-order autoregression effects) had a minimal effect on our predictive capacity.

Table 4 compares the empirical distribution of spatial correlations to the distribution one would expect if there were no spatial effects and all such correlations were due to random deviations. The larger the spatial effects, the larger the variance of the empirical distribution relative to the expected distribution. We can measure the difference by dividing the interquartile range of the empirical distribution by that of the expected distribution—the larger the ratio, the bigger the spatial effect. In addition,

Table 3: Summary of Temporal Displacement Analysis Results

	Crime	Disturbance	Service	Total
White noise	72.9%	62.4%	88.8%	74.7%
	188	161	229	578
First-order	14.7%	26.7%	4.3%	15.2%
	38	69	11	118
AR(1)	22	31	2	55
MA(1)	0	0	1	1
Ambiguous	16	38	8	62
Second-order	5.0%	3.9%	0.8%	3.2%
	13	10	2	25
AR(2)	5	2	1	8
MA(2)	0	0	0	0
Ambiguous	8	8	1	17
Other	7.4%	7.0%	6.2%	6.8%
	19	18	16	53
Total	100.0%	100.0%	100.0%	100.0%
	258	258	258	774

Note: Figures shown are number of locations for which ARMA models of the type shown are most appropriate.

a χ^2 test shows the significance of these deviations (Breusch and Pagan, 1980). Spatial effects are evident in every case, but especially among housing projects and parks and for disturbance calls. This makes sense; it is probably easier to imagine movement among parks by gang members than any other form of displacement.

Although we can be sure that spatial effects exist, they are not closely associated with distance. When equation (5) was estimated for all combinations of call and location types, $\beta < 0$ for every case—a strong indication that the extent of displacement drops as the distance increases. Neverthe-

Table 4: Deviations from Random Expectations in Spatial Correlation Distribution

	IQR ratio	χ^2	df(χ^2)	p(χ^2)
High schools				
crime	1.149	805.1	595	.0000
disturbance	1.175	862.2	595	.0000
service	1.069	694.4	595	.0029
Housing projects				
crime	1.320	1,384.6	630	.0000
disturbance	1.357	1,407.7	630	.0000
service	1.259	1,017.79	630	.0000
MBTA stations				
crime	1.034	1,497.6	1,378	.0123
disturbance	1.123	2,749.7	1,378	.0000
service	1.112	2,047.8	1,378	.0000
Parks				
crime	1.211	20,118.0	9,453	.0000
disturbance	2.078	51,741.2	9,453	.0000
service	1.099	18,697.8	9,453	.0000

Note: IQR ratio is the empirical interquartile range divided by the IQR that would be obtained if all correlations were random errors. χ^2, df(χ^2), and p(χ^2) refer to parameters and results of the Breusch-Pagan (1980) χ^2 test.

less, only two of the regressions were significant at the .05 level, and the highest R^2 obtained was .005. Little wonder that spatial displacement, as measured by S_{it} from equation [6], was not an important predictor of variability among locations over time.

Finally, Table 5 shows the correlation of residuals of equation [1] among call types for each of the four location types. All are positive and statistically significant, and the interrelationships among crime, disturbance and service calls all seem to be about the same size. This suggests that intertype effects are not measures of displacement so much as they

Table 5: Correlation of Residuals among Location Types

	Crime and Disturbance	Crime and Service	Disturbance and Service	Cronbach's Alpha
High schools	.062	.178	.108	.282
Housing projects	.218	.150	.105	.360
MBTA stations	.063	.071	.050	.164
Parks	.154	.143	.105	.337

Figure shown is correlation coefficient between residuals of each pair of regressions. All correlations significant at $p < .05$. Cronbach's alpha is a measure of the consistency of results among pairs of call types.

are a different way to estimate the number of users of a location. Nevertheless, the largest correlation between types for any location type is just over .20. As Table 2 shows, intertype effects appear to be relatively minor factors in the production of temporary changes in risks.

In summary, correlations in time and space and between types accounted for little of the total variation: 2 to 6% among all types of locations. This represents about 20% of the total variance attributable to short-term changes in risks. Thus,the vast majority of these short-term changes are unpredictable, at least given these means of operationalizing current theory. Better methods or more complete theories may be needed to explain these changes. Nevertheless, it is important to emphasize that short-run changes are consistently less important than persistent differences among locations, so there is probably more to be gained by focusing attention on the causes of long-term differences than of short-term changes.

IMPLICATIONS AND EXTENSIONS

Community problem-solving remains an inexact science, and the full potential of problem-solving methods will not be tapped until they are viewed as reliable and procedural. Although the results described above are but one step toward this objective, they can help us to make problem-solving more systematic. Three policy-relevant extensions of this analysis are considered below.

Identifying the Worst Locations. Even if long-run risks are the most important source of variation, for some locations random errors and changes in risks are also important. So if we identify "high-rate" locations on the basis of only one month of data, we will certainly be fooled some of the time.

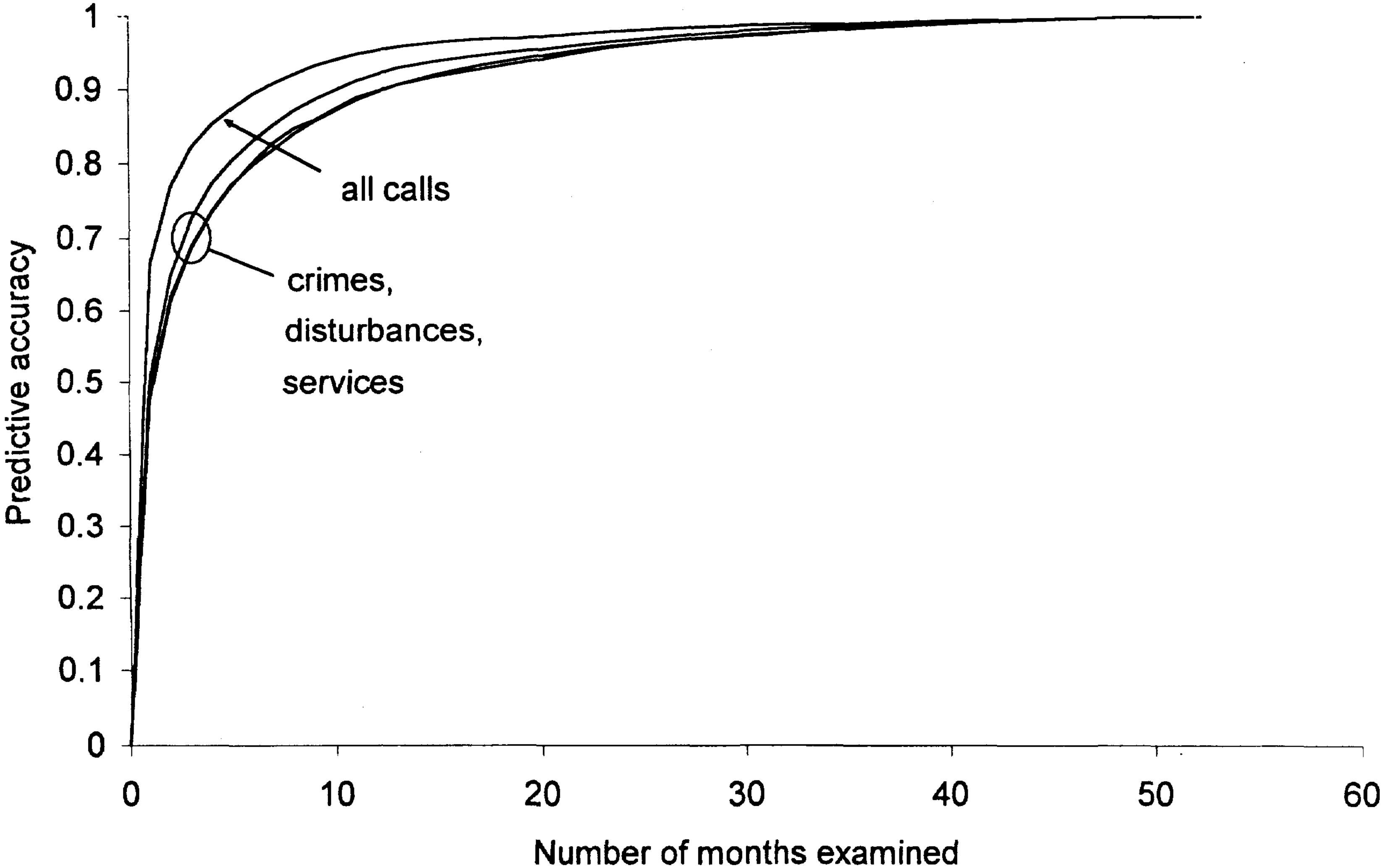
The simplest solution is to look at calls for service over a longer period—three months, six months or a year—and only identify a location as high rate if it produces a lot of calls throughout the period. As the length of the observation period increases, predictive accuracy will improve.[1] If the period is long enough, our predictions will be perfect and accuracy will be 100%. The problem is that longer time periods are more difficult to program into the computer. In some police agencies, computer-aided dispatch system data more than a few months old are simply not available. Thus we must know how much data we need to collect before we can accurately identify recurring high-risk locations.

The MBTA stations sample provides an appropriate data set for this kind of thought experiment. Some 37% of the total variation among MBTA stations over time is due to random error—more than any of the other locations—so we can be sure that the predictive accuracy for any given time period will be greater for the others than for this sample. If, for example, six months is adequate for predicting risks at MBTA stations, it will be more than adequate for high schools, housing projects and parks.

Figure 4 shows results for three call types and for total calls. For each call type, the results are similar. One month of data is sufficient to predict long-run risks with about 50% accuracy; with two months, accuracy rises to between 60 and 65%; after six months, accuracy is about 80%; and at one year (thirteen 28-day periods), the accuracy is a quite respectable 90%. If our only aim is to predict long-run risks for *all* calls, rather than for crimes, disturbances and service calls separately, then we can achieve 95% accuracy with only one years' worth of data. Although additional data may nail down some close calls, and help problem-solvers fine tune their resource allocations, the differences are unlikely to be critical.

Setting Expectations for Community Problem-solving. If police, neighborhood residents, merchants and other users of a location use community problem-solving techniques, how well can they be expected to work? This question is of enormous policy and operational importance. Policymakers are interested in knowing how many eggs we should put in the community problem-solving basket; operational personnel need specific objectives that can be reasonably achieved and at least a rough idea of when to quit. For example, if problem-solving can realistically reduce crime by, say, 40%

Figure 4: Predictability of Long-run Risks
MBTA Stations



in some location, then line officers and neighborhood organizations err if they quit after a 10% reduction—there are many gains left on the table. They also err if they persist after a 38% reduction—there is little left to accomplish, and they could probably achieve more if they took on a different problem.

The only real answer to this question can be obtained through experimentation—trying out a wide variety of responses in a wide variety of locations, and seeing how well they work. Unfortunately, this will take years of effort. But we can obtain a rough-cut answer to this question by rephrasing it slightly: What percentage of the long-run differences among locations can be attributed to factors beyond the control of problem-solvers? If this figure is very high—say, 90%—there is little to be done. On the other hand, if the figure is low, the potential gains are enormous.

The operational problem with answering this question is that it is almost impossible to get a complete list of the factors that affect risks that are beyond the control of problem-solvers. Some are well-known. We can collect data on the number of people who use a location, since more users mean more opportunities for crime and disorder. We can examine the demographic characteristics of the users, identifying how many are in high-risk groups for offending or victimization, such as young males, single mothers and the poor. In practice, however, we cannot be sure we have all the relevant variables. This means that any results will probably underestimate the variation due to uncontrollable factors, and overestimate the importance of the factors controllable by problem-solvers. Keeping this significant caveat in mind, let us apply the method to the MBTA station data.

MBTA stations provide a particularly convenient sample for such an analysis. The number of people using the stations can be reliably measured from fare-collection data. And, for most locations, we can expect that the people who use the station are much like those who live in the surrounding neighborhood. Thus census data can be used to measure the number of potential offenders and victims.

When all available data are considered, the best predictors of MBTA station calls for service turn out to be the total number of station users, the average income of neighborhood residents, whether the station is located in a commercial area and whether it is on the Orange line, which runs through the most crime-ridden sections of the city. Other characteristics of the station and its neighborhood added little to our ability to explain subway crime. These four factors accounted for 55% of the variation among stations, suggesting that roughly half that variation may be subject to attack by problem-solvers.

Another way of describing the same findings is more direct. The worst five MBTA stations produce 28% of all calls to subway stations. When factors beyond the control of problem-solvers are accounted for, they *should* have produced only 15%. Thus nearly half of the calls to the worst stations could, in theory, be reduced by changing conditions that create criminal opportunities. If each violent crime costs the victim about $14,000, and each property crime costs about $565 (Cavanagh, 1991; Cohen, 1988), the annual benefits to victims in crime reduction alone could be as much as $236,000 and $357,000 per year. This does not take into account indirect costs of crime, such as avoidance of the subway system, anxiety and fear among those who use it, and the costs of responding to these calls.

Although highly speculative, these figures do not seem unreasonable. The more general result certainly seems defensible: factors unique to each location appear to be important determinants of long-run risks; they are possibly as important as the factors common to all locations. The plausibility of the result suggests that the method may help problem-solvers guide their resource allocation efforts in the future.

Measuring Effectiveness. If the worst locations cannot be identified with only one or two months' data, it is clear that changes in long-run risks at these locations cannot be identified this quickly, either. Random error is liable to swamp any true changes unless rates are measured for several months.

This has two implications for problem-solvers. First, it hampers a strategy of incremental responses. Unless the effect of any given intervention is enormous, it will not be possible to make a fast, accurate assessment of its effectiveness before trying out another intervention. More generally, this means that many problems will require a long-term commitment from the police. Even if officers always play it safe and implement every response they can think of, they will not succeed. But their failure—when at last it becomes apparent after several months—may suggest other avenues of approach not obvious from the beginning. Unless the causes of and solutions to a problem are obvious, then, the police will need to count on an incremental response strategy.

The lag between response and assessment will be even longer if police find the methods of assessment difficult. The methods most often used for this purpose—interrupted time-series analysis—are complex and difficult to implement because few computer programs are available.[3] If seasonal effects and sporadic risk changes are relatively unimportant, however, we can rely on the *control chart*, a simple means of accounting for random variations in time-series data.

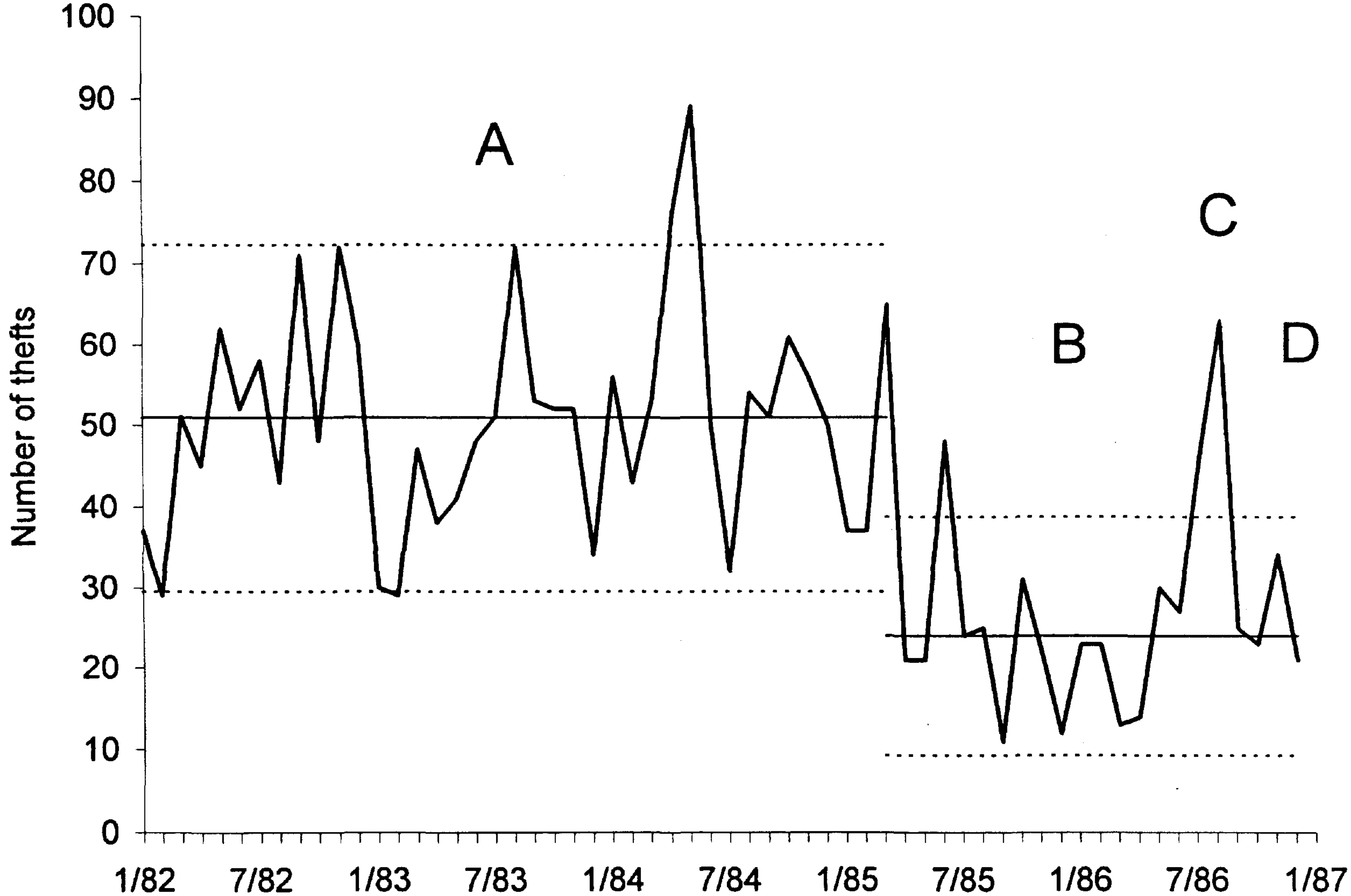
Control charts are often used to monitor the incidence of diseases and quality defects (Centers for Disease Control and Prevention, 1986; Montgomery, 1991). Briefly, the control chart is simply a plot of the number of incidents reported over time, with the average value and expected upper and lower ranges drawn in for comparison. So long as the time-series stays within the upper and lower bounds, the problem is in "statistical control"—not getting better, but not getting worse, either (Deming, 1986). If the time-series drops below the lower bound after an intervention has been implemented,then the problem-solvers can be reasonably certain that they have ameliorated the problem.

In Newport News, VA, a patrol sergeant used a control chart to monitor burglaries from automobiles parked near a large factory (Eck and Spelman, 1987). As shown in Figure 5, the number of burglaries had been close to statistical control for the three years preceding the intervention (labeled *A* on the figure). In March 1985, two gangs of youthful burglars were broken up, and in the next month the number of burglaries went below the lower bound for the first time on record. It stayed there for several months, establishing a new level with a lower average (B). About one year after the initial intervention, the number of burglaries rose above the new upper bound, signaling that the system had changed again (Q. Investigation showed that a pair of high-rate juvenile offenders had entered the area, and the burglary rate went down after they were arrested and incapacitated (D).

As this example demonstrates, the control chart can be an invaluable assessment tool. The officers assigned to this problem could show within two months that their efforts were having an effect (in part because the effect was so large). The chart also showed when conditions changed, making an early response possible. Finally, the chart fulfills some administrative objectives. Because they are easy to use, control charts can be maintained by operational personnel. This sustains the focus on results and improves the link between the line officers' actions and the consequences of their actions for the public.

The classic control chart only aims to separate random variations from long-run averages, and upper and lower bounds are set under the assumption that there are no seasonal or trend effects or short-run changes in risks. As shown above, these are unrealistic assumptions for most public places (although they were fairly realistic for the auto burglaries problem). They are particularly unrealistic for the housing project series, for which random processes explain only 4% of the variation but seasons, trends and other short-run changes explain 13%. When these assumptions were

Figure 5: Control Chart
Parking Lot Thefts, 1982-1986



applied to the 105 housing project series (crime, disturbance, and service call time-series for each of 35 projects), the results were predictably disastrous. For most series, some months lay well above the bounds, even though the long-term risks remained constant. The underlying distribution was more skewed than the Poisson for all series.

Use of a control chart with inappropriate bounds would be misleading, but the skewed nature of the distribution suggests an effective adjustment. There are reasons to believe that temporary changes in risks should be multiplicative in nature (Spelman, 1992), thus the distribution of underlying risks for each series should be roughly logNormal. Although random deviations about these (changing) underlying risks add to the variability, Table 2 shows that temporary changes are roughly three times as important as random deviations in explaining the variability within these series. So the number of calls reported over time for any given location should be roughly logNormal-distributed. Thus, appropriate bounds can be set by using the following procedure:

1. Estimate the mean of the empirical series through the usual methods.

2. Take the logarithm of all cases, and estimate the mean and standard deviation of the log distribution.

3. Multiply the log standard deviation by 3.0 (as usual), and add it to and subtract it from the mean of the log distribution. These are the appropriate control limits for a logNormal distribution.

4. Take the exponent of the upper and lower bounds to form bounds for the control chart.

This structures the control chart to screen out random deviations plus short-term changes in underlying risks that are like those found in the recent past. If a new program or policy is effective enough to reduce risks by more than these amounts, the control chart should correctly reflect it.

When upper and lower bounds are set in this way, 95% of time-series in the Boston calls-for-service data set are consistently within bounds:

- 76% fit the control chart perfectly (that is, they were in control throughout the four-year period);
- 19% had more points near the lower bounds than expected, but did not go below the bounds; and
- only 5% included months that went outside the upper or lower bounds.

Since this would happen in about 5% of the series just by chance,[4] the control charts appeared to be working perfectly. Thus the Boston data confirm that control charts can work on calls-for-service data in a wide variety of locations. By adopting this simple and effective way to monitor problems and assess the impact of responses, line officers can make sophisticated statistical judgments.

CONCLUSION

The promise of sitting ducks, ravenous wolves and dens of iniquity is that we can accomplish a lot by focusing our efforts on a few. In theory, we can rehabilitate, deter or at least incapacitate a few very dangerous offenders; we can help especially vulnerable victims to avoid crime and defend themselves; and we can solve problems at especially dangerous locations, reducing them without displacing them. As with dangerous offenders and vulnerable victims, some of this promise is lost upon close examination. Much of the concentration of crime among locations is due to random and temporary fluctuations that are beyond the power of the police and the public to control reliably.

On the other hand, there is much left to be gained through community problem-solving. Among public places, at least, the worst 10% of locations reliably account for some 30% of all calls. The results described above suggest that crime, disturbances and other calls for service can be reduced by something like 50% in the most dangerous locations, simply by focusing on the unique characteristics of those locations that create opportunities for crime and disorder. It remains to be seen whether we can develop the tools to analyze and respond to these problems adequately. But the potential benefits remain vast and worthy of further study.

✦

NOTES

1. Accuracy is defined here as the squared correlation between the number of calls received during the short-run period and the long-run expectations for each location, averaged over all locations.

2. In general, this is referred to as the problem of "left-out variable error." For more information, see Judge et al., 1985.

3. For example, the most widely used statistical package, SAS, has only recently made an interrupted time-series analysis program available as a standard option (SAS Institute, 1991). No such program is as yet available on SPSS.

4. Bounds were set so that 0.1% of observations would go beyond the bounds due to random variation, so the expected number of beyond-bound observations is 52 x .001 = .052 per series. Thus, 5.2% of the series should include one or more beyond-bounds observations.

REFERENCES

Berk, R.A., D.M. Hoffman, J.E. Maki, D. Rauma and H. Wong (1979). "Estimation Procedures for Pooled Cross-sectional and Times-series Data." *Evaluation Quarterly* 3:385-410.

Blumstein, A., J. Cohen, J.A. Roth and C.A. Visher (eds.) (1986). *Criminal Careers and "Career Criminals."* Washington, DC: National Academy Press.

Box, G.E.P. and G.M. Jenkins (1976). *Time-series Analysis: Forecasting and Control.* San Francisco, CA: Holden-Day.

Breusch, T.S. and A.P. Pagan (1980). The LaGrange Multiplier Test and Its Applications to Model Specification in Economics." *Review of Economic Studies* 47:239-254.

Cavanagh, D.P. (1991). *A Cost-Benefit Analysis of Prison Cell Construction and Alternative Sanctions.* Cambridge, MA: BOTEC Analysis Corp.

Centers for Disease Control and Prevention (1986). *Principles of Epidemiology: Disease Surveillance.* Manual 5, Self-Study Course 3030-G. Atlanta, GA: author.

Chaiken, J.M. and M. Chaiken (1982). *Varieties of Criminal Behavior.* Santa Monica, CA: Rand.

Clarke, R.V. and D.L. Weisburd (1992). "On the Distribution of Deviance." In: D.M. Gottfredson and R.V. Clarke (eds.), *Policy and Theory in Criminal Justice.* Aldershot, UK: Avebury.

Cohen, M. (1988). "Pain, Suffering, and Jury Awards: A Study of the Cost of Crime to Victims." *Law and Society Review* 22:537-555.

Cohen, L.E. and M. Felson (1979). "Social Change and Crime Rate Trends: A Routine Activity Approach." *American Sociological Review* 44:588-608.

Deming, W.E. (1986). *Out of the Crisis.* Cambridge, MA: MIT Press.

Eck.J.E. and W. Spelman (1987). *Problem Solving: Problem-Oriented Policing in Newport News.* Washington, DC: Police Executive Research Forum.

Fienberg, S.E. (1980). "Statistical Modeling in the Analysis of Repeat Victimization." In: S.E. Fienberg and A.J. Reiss, Jr. (eds.), *Indicators of Crime and Criminal Justice: Quantitative Studies.* Washington, DC: U.S. Government Printing Office.

Goldstein, H. (1990). *Problem-Oriented Policing.* New York, NY: McGraw-Hill.

Greene, J.R. and S.D. Mastrofski (eds.) (1988). *Community Policing: Rhetoric or Reality?* New York, NY: Praeger.
Horney, J. and I.H. Marshall (1991). "Measuring Lambda Through Self-Reports." *Criminology* 29:401-425.
Judge, G.G., W.E. Griffiths, R.C. Hill, H. Luetkepohl and T.C. Lee (1985). *The Theory and Practice of Econometrics.* New York, NY: John Wiley & Sons.
McCleary, R. and R.A. Hay, Jr. (1980). *Applied Time Series Analysis for the Social Sciences.* Beverly Hills, CA: Sage.
Mande. M. J. and K. English (1988). *Individual Crime Rates of Colorado Prisoners.* Denver, CO: Division of Criminal Justice, Colorado Department of Public Safety.
Montgomery, D.C. (1991). *Introduction to Statistical Quality Control.* New York, NY: John Wiley & Sons.
Nelson, J. F. (1984). "Multiple Victimization in American Cities: A Statistical Analysis of Rare Events." *American Journal of Sociology* 85:870-891.
Pierce, G.L., S. Spaar and L.R. Briggs (1986). *The Character of Police Work: Strategic and Tactical Implications.* Boston, MA: Center for Applied Social Research, Northeastern University.
Reiss, A.J., Jr. (1980). "Victim Proneness in Repeat Victimization by Type of Crime." In: S.E. Fienberg and A.J. Reiss, Jr. (eds.), *Indicators of Crime and Criminal Justice: Quantitative Studies.* Washington, DC: U.S. Government Printing Office.
SAS Institute (1991). *SAS/ETS User's Guide* (version 6). Cary, NC: author.
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.
Spelman, W. (1992). *Criminal Careers of Public Places.* Final Report to the U.S. National Institute of Justice. Austin, TX: LBJ School of Public Affairs, University of Texas at Austin.
Srivistava, V.K. and D.E.A. Giles (1987). *Seemingly Unrelated Regression Equations Models: Estimation and Inference.* New York, NY: Marcel Dekker.
Stimson, J.A. (1985). "Regression in Space and Time: A Statistical Essay." *American Journal of Political Science* 29:914-947.
Zedlewski, E.F. (1987). *Making Confinement Decisions.* Research in Brief. Washington, DC: National Institute of Justice.
Zellner, A. (1962). "An Efficient Method of Estimating Seemingly Unrelated Regression Equations and Tests for Aggregation Bias." *Journal of the American Statistical Association* 57:348-368.