The Dynamics of Crime Regimes*

Richard Berk
John MacDonald
Department of Statistics
Department of Criminology
University of Pennsylvania

January 20, 2009

Abstract

Crimes have many features, and the mix of those features can change over time and space. In this paper, we introduce the concept of a crime regime to provide some theoretical leverage on collections of crime features and how the collection of features can change. Key tools include the use of principal components analysis to determine the dimensions of crime regimes, visualization methods to help reveal the role of time, summary statistics to quantify crime regime patterns, and permutation procedures to examine the role of chance. Our approach is used to analyze temporal and spatial crime patterns for the City of Los Angeles over an 8 year period. We focus on the number of violent crimes over time and their potential lethality.

*Richard Berk's work on this paper was funded by a grant from the National Science Foundation: SES-0437169, "Ensemble methods for Data Analysis in the Behavioral, Social and Economic Sciences." A special thanks goes to Andreas Buja for a number of helpful suggestions on how to analyze the data. We are also indebted to the Los Angles Police Department for providing the data used in this paper. Finally, we benefitted enormously from comments made when an earlier version of this paper was given at the Deparment of Criminology of the University of Maryland and at the Department of Criminal Justice of Temple University.

1 Introduction

The distribution of crime in time and space has long been of significant interest to criminology. Nineteenth century European criminologists Guerry, Bonger, and Quetelet were among the early positivists to examine the spatial and temporal distribution of crime and consider its causes (see Beirne, 1987; Vold and Bernard, 1986). Arguably the most influential perspectives on spatial and temporal variation in crime draw heavily on the social disorganization theory of Shaw and McKay (1942). The distribution of crime in time and space can be explained by differences across neighborhoods in poverty, residential stability, and ethnic heterogeneity that, in turn, affect the development of common values and maintenance of informal social controls (Bursik, 1988; Kornhauser, 1978; Byrne and Sampson,1986; Sampson and Groves, 1989). Fredrick Thrasher (1927) suggested similar mechanisms to explain the spatial and temporal distribution of youth gangs in Chicago neighborhoods.

The past 30 years have seen a host of studies building on the social disorganization tradition, capitalizing on new theory, better data, and statistical tools more capable of capturing spatial and temporal variation. By and large, work in this tradition has focused on variation in crime over relatively large geographical areas and units of time: cities from one census to the next, neighborhoods from year to year, police beats from month to month and the like. Sometimes driven by the data available, and sometimes by theory, there is now a rich literature at these scales.

Yet, crime can vary at more micro levels in interesting and important ways. Local opportunity structures are a key instance (Clarke and Felson, 1993; Cohen and Felson, 1979; Wilcox, Land, and Hunt, 2003). Recently, criminologists also have begun to exploit relatively fine-grained temporal patterns, often analyzed as crime trajectories (Nagin and Land, 1993) within neighborhoods (Weisburd et al, 2004; Griffith and Chavez, 2004). A rich set of proposed causal mechanisms has followed.

In this paper, we step back from the important job of explaining the causes of crime in time and space to reconsider crime patterns themselves. We work across spatial scales ranging from small police administrative units to an entire city, and across temporal scales ranging from days to weeks. To do this, we introduce the concept of a *crime regime* and provide data analysis tools to characterize how crime regimes can change over time. Among the benefits are potential links to work on dynamical systems, concepts that may ultimately contribute to theory development, and a practical means to detect changes in crime over time and space that might otherwise go unnoticed or be characterized in an unsatisfactory manner. Although many of the components on which we build have been available to criminologist for some time, we combine them in what we hope is a novel manner and introduce some new ideas along the way. For better or worse, the final product is distinctive.

Section 2 provides a brief look at past work to help set a context for our approach. Section 3, combines important conceptual material with proposed statistical methods. The concepts are central because they determine how one should proceed in practice. Indeed, we believe that insofar as our approach has merit, it lies far more in how to think about crime in time and space than with the tools we suggest for analyzing it. Section 4 provides a lengthy illustration using data from the City of Los Angeles for the years 2000-2007. It is intended to be an illustration of our approach, not a definitive statement about the dynamics of crime. Sections 5 and 6 extract from earlier material a number of implications and conclusions.

2 Setting the Context

With roots in the Chicago school tradition of urban sociology (Park, 1915), there is now a large literature on social and economic processes thought to be responsible for crime patterns over time at city, state, and neighborhood levels (Bursik, 1988; Land et al., 1990; Messner et al., 2001; Osgood and Chambers, 2000; Ousey and Augustine, 2001; Sampson and Groves, 1989; Sampson et al., 1997; 2005; Shihadeh and Steffensmeier, 1994). Family disruption and concentrated poverty, for instance, are routinely linked to high crime rates at varying geographic scales (see Land et et.,1990; Sampson, 1987; Sampson et al., 2005; Shihadeh and Steffensmeier, 1994). But, the mix of environmental influences can vary geographically (Taylor, 2001). Osgood and Chambers (2000) show, for example, that there are differences between urban and rural counties in the variables associated with aggregate youth violence rates.

Various mechanisms, sometimes competing (Wolfgang and Ferracuti, 1967) and sometimes overlapping, have been proposed to explain how such spatial features can affect crime (see Bursik, 1988 for discussion). Changes in opportunity structures and concentrations of unsupervised peer groups are two examples (Elliot et al., 1996). Gangs and drug trade feeding predatory environments, which disable informal social control, are also invoked (Anderson, 1998). Frequent mention is made of a lack of jobs and economic opportunities that can lead to increased idleness, fewer employed men who are desirable marriage partners, and less community supervision of young people (Sampson, 1987; Wilson, 1987). Spatial differences in economic disadvantage are also associated reduced participation in social organizations and the willingness of residents to engage in informal social control (Janowitz, 1975; Sampson and Groves, 1989; Sampson et al., 1997; Sampson and Raudenbush, 1999).

A more recent approach has focused on features of highly localized spatial environments that define crime opportunity structures and make some places and times more attractive to criminals than others (Clarke and Felson, 1993; Cohen and Felson, 1979; Wilcox, Land, and Hunt, 2003). These micro environments can be situated within communities, neighborhoods, and police jurisdictions to allow for theories operating at more than one spatial level (Eck and Weisburd, 1995; Sherman et al., 1989). Recent work by Weisburd and colleagues (2004), for example, indicates that a relatively small share of street segments in Seattle are responsible for most of the city's overall drop in crime during the 1990s.

The rich literature on crime in time and space has, by and large, used crime rates or crime counts of various kinds as the outcomes to explain. These outcome measures have served criminology well. But perhaps there are other productive ways to think about crime in time and space. It is widely appreciated that crimes can have a number of attributes that can vary. To take a simple example, the number of crimes *and* the seriousness of those crimes can vary by location and by month. It may be useful, therefore, to think about a collection of crime attributes and how that collection of attributes changes over time and space. We use the term *crime regime* to represent a collection of crime attributes. A more formal and complete definition will follow shortly.

The central role of crime regimes in this paper will also highlight the need to consider crime dynamics in an explicit and systematic manner. Interest in patterns of crime over time dates back at least to the 19th century when, for example, Bonger summarized work describing how from 1853-1892 the yearly price of "provisions" was related to crimes against persons (Bonger, 1967: 68-71). If one fast forwards to 2008, there are a number of closely related issues addressed in an NRC publication *Understand Crime Trends:*

Workshop Report (Petrie, 2008). In both instances, a fascinating effort is made to describe how long terms trends in one or more explanatory variables are related to long term trends in one or more response variables and to explain why such relationships exist. However, time is primarily a marker by which observations can be arrayed. The passage of time does not play a direct role in formal theory explaining why the crime trends take the form they do.

Treating time primarily as a marker represents standard practice in criminology, with only occasional exceptions. For example, Berk and his colleagues (1983) propose and test a model of prison populations over time in which a key feature is a self-regulating mechanism helping to keep the number of inmates within certain bounds. Other examples can be found in theoretically informed time series modeling (Land et al., 1995) and in systemic approaches to the interlocking functions of criminal justice agencies (Blumstein and Larson, 1969; Berk and Cooley, 1979). Even when change is the explicit subject of inquiry, most criminologists treat time as little more than an ordinal variable with which to determine a sequence of events (Bursik, 1986; Taylor and Covington, 1990).

3 Conceptual Framework and Methods

As a way to provide some initial sense of the conceptual issues, consider the following example. Suppose one were interested in characterizing the state of drug trafficking in a city. Drug trafficking has a number of important attributes. Suppose that one were interested in and could obtain data on three of them: the number of transactions, whether the transactions took place in "open markets" (Reuter and MacCoun, 1992), and the number of associated homicides or attempted homicides. (See Blumstein et al., 2000 for a discussion of drug markets and violence.) For a given time period, one could summarize the state of drug trafficking with a value on each dimension. In a given month, for example, there might be 28 transactions, 75% of which were open air, with 6 associated homicides or attempted homicides. If one were then able to obtain the same information for a series of months, one would be able to report the state of drug trafficking month to month, and it is likely that over time the values on all three dimensions would vary.

Imagine now plotting the state of drug trafficking in that city, month to month, in a three-dimensional scatter plot. Each dimension in the plot would correspond to one of the three attributes of drug trafficking. Each point in that three-dimensional space would characterize the state of drug trafficking in a given month. And one could follow the path of such points over time to see how the state of drug trafficking was changing. The time path through the scatter plot might be of great interest and in principle could be statistically analyzed. For example, are there smooth trends or do the points tend to cluster in some regions of the three dimensional space? And if the latter, what might account for the clustering and the rapid transitions from one cluster to another.

What we consider below has some of this look and feel. We too examine the location and movement of crime attributes over time in two or more dimensions. But the crime attributes of interest are not prejudged; they are determined by the data available. We do not, for example, assume there are crime regime clusters and do not set out to construct such clusters. This heavy reliance on data means that we must begin by discussing what such data are like and how such data can be organized. We then turn to how crime regimes over time can be effectively displayed. Finally, we will suggest how summary statistics and ways to address uncertainty can be usefully applied.

3.1 Data Structure

Crime regimes can be operationalized in a many ways using many different data types. For example, a variety of information may be extracted from 911 calls, police dispatches, arrests, parole violations, coroner's reports, survey data, and many other sources. In this paper, we will use "offense reports" from the Los Angeles Police Department (LAPD).

When Los Angeles police officers are dispatched to a potential crime scene or come upon a possible crime during a routine patrol, they are usually required to fill out an offense report. Time when the incident occurred may be reported to the minute, and location may be reported by street address or some other small spatial unit such as a police beat. A variety of other crime attributes also are usually recorded: the kind of crime, victim and perpetrator characteristics, circumstances surrounding the crime, so on.

In this paper, we will construct crime regimes from four pieces of information, (1) when a crime occurred, (2) where a crime occurred, (3) the kind crime (within commonly used, broad categories of offenses), and (4) some salient features of those crimes. These crime attributes are almost universally recorded, have relatively simple interpretations, and can lead to interesting results that may be of interest to law enforcement agencies and to criminologists seeking to develop causal explanations for crime regimes patterns.

In principle, offense data can be arrayed with locations as columns and time as rows. Each cell would contain a count of the number of crimes by locale and time. Each column could then provide a time profile of the frequencies of crime events in given locale. Somewhat more complicated data structure also could be used. One could allow for columns representing different kinds of crime in each area. For example, for each police precinct in a city, there could be a column for the number of homicides and a column for the number of assaults. In this paper, each geographical locale will have two features arrayed over time: the number of violent crimes and the proportion of times in those crimes a handgun of assault weapon was used. More details follow shortly, but at this point it is important to stress that the application is meant primarily to be illustrative.

For each column, we use the term "temporal profile" rather than "time series" to underscore that we will be looking for patterns over time. If there are no temporal patterns, then crime counts over time are just "noise." That might be an interesting finding, but also a dead end. Temporal fluctuations in crime would have no systematic content. And were that the case, there would be no point in reading the rest of this paper.

Figure 1 is a mock-up of a data matrix of the sort we will consider. Time is indexed by $t = 1, 2, \ldots, T$, with T here equal to 20. Location is indexed by $s = 1, 2, \ldots, S$, with S here also equal to 20. Each cell contains a count for the number of crimes. In practice, T will need be be much larger than 20 if temporal profiles are to be usefully characterized, and because of the statistical procedures we will apply shortly, S cannot be larger than T.[1] The former is likely to be far more critical than the later. Interesting temporal patterns are unlikely to be well described with less than 50 observations and more commonly, at least a few hundred time points are needed. When T is of the requisite size, the constraint on the size of S will usually not cause serious problems. In our later illustration, there will be two columns under each of the S locations, one for the violent crime counts and one for the proportion of times a handgun or assault weapon was involved. Still, T will

[1]This is much the same as the usual restriction on conventional regression analysis. One cannot estimate more regression parameters than there are observations. Alternatively put, the number of predictors in the equation cannot exceed one minus the number of observations.

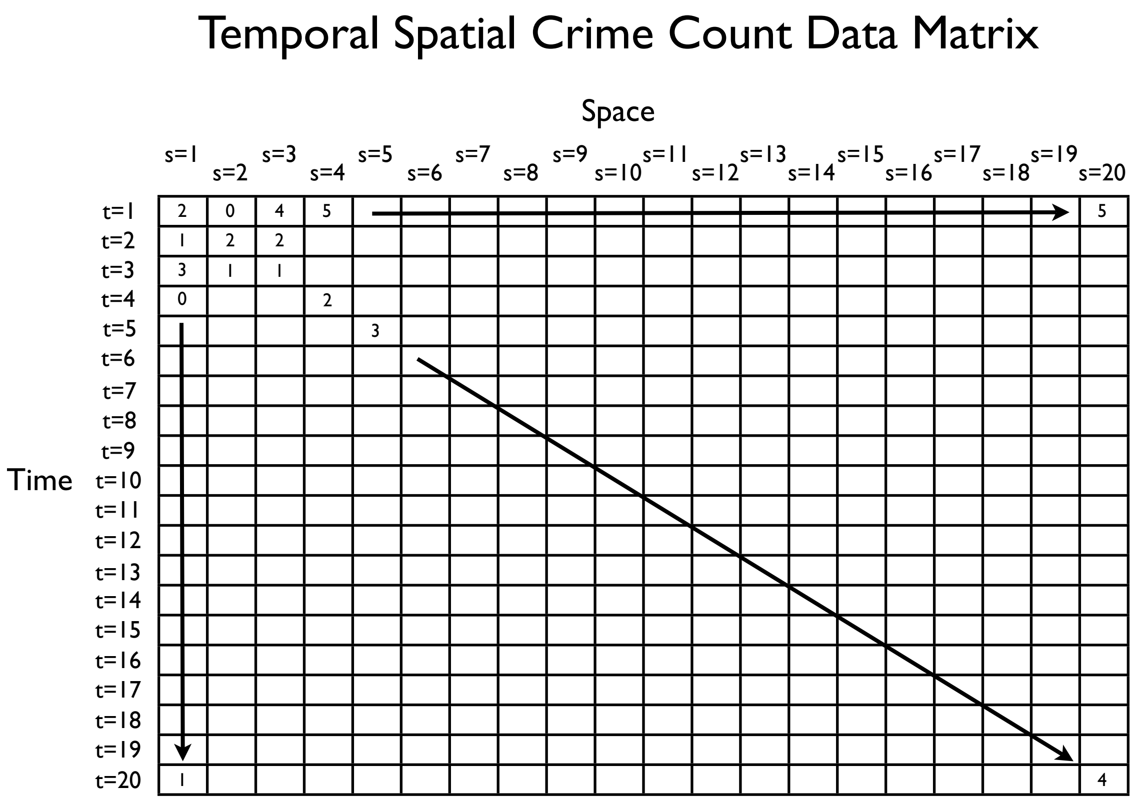

Figure 1: Crime Regimes Data Matrix

be greater than S.

Such data can be very useful for law enforcement and other practical purposes. Data use protocols such as COMPSTAT exploit just this kind of information (Henry, 2003). However, without some means to characterize significant patterns in such data, the data are of limited use for comprehending systematic variation of crime in time and space. And without this comprehension, there can be no deeper insight.

Conventional maps of crime in time and space are a step in the right direction, but map displays are meant to locate crimes in real geographical space. Our intent is to represent abstract properties of data that can serve to complement and extend conventional maps. As Cook and Swayne observe (2007: 2), "When we visualize data, we are interested in portraying abstract relationships among ... variables." In the spirit of scatter plots, therefore, the space represented in our analyses will be a space defined by variables. This will lead to crime regimes and how they can change over time in variable, not geographic, space.

3.2 Visualization Methods and Some Definitions

It would not be surprising if the temporal profiles across the spatial units had some important commonalities: monotonic trends, seasonal patterns, shifts in level or slope, spikes up or down, outliers and others. Should such patterns across spatial units exist, one strategy would be to extract the shared temporal patterns and use those in place of the spatial units. Figure 2 provides an idealized version of some time profiles. Crime counts is on the vertical axis and time (e.g., weeks) is on the horizontal axis. Real time profiles would be perturbed by noise, and could well be composed of several temporal patterns such as those in Figure 2.

One straightforward way to proceed is with principal components analysis (PCA). Consider again the data matrix shown in Figure 1. PCA does nothing more than construct S linear combinations of the columns so that the linear combinations are orthogonal.[2] The linear combinations are built so that the first accounts for the largest share of the variance, the second accounts for

[2]Each principal component is composed of the weighted sum of the columns of the data matrix, with weights for each principal component determined by the procedure so that principal components are uncorrelated with one another. That is, the data matrix is reconfigured so that it contains the same amount of information, but with new columns (variables) that are uncorrelated.

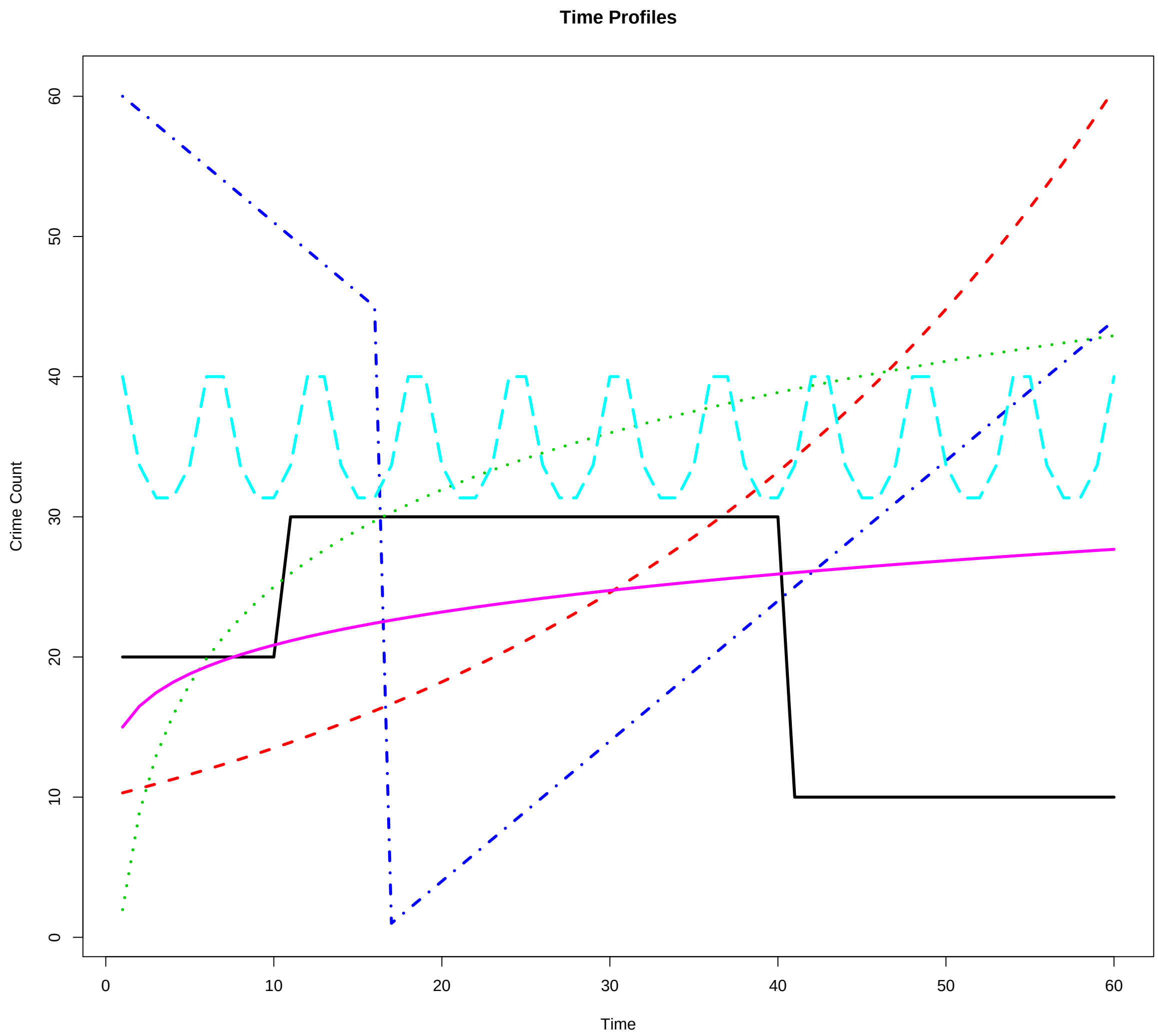

Figure 2: Examples of Time Profiles

the second largest share of the variance, and so on up to the Sth principal component that accounts for the least amount of the variance. The variance contributions over the S principal components sum to the 100%. Usually, only the first two or three principal components are needed because the remaining variance accounted for is very small. For the kinds of data likely to be found in Figure 1, these principal components would capture the most salient temporal patterns in crime shared by the spatial units. For example, there may be a pendulum-like alternating process over time as a balance between processes that generate crime and processes that restrain it. One nice feature of PCA is that in contrast to such procedures as factor analysis, PCA is model free (Morrison, 1976: Chapter 8). PCA is data reduction tool, not a method to represent how the data were generated.[3]

It is often useful to plot the principal components against one another. Figure 3 shows an idealized version of such a scatter plot. For ease of exposition, the units of both principal components are crime counts centered on zero. Because all of the time profiles have the same units (i.e., crime counts), there is no need to work in standard deviation scores; we will move to standardized units when we later turn to real data. The principal components can be seen as summaries of the salient temporal patterns in the amount of crime over all of the spatial units. In other words, even though the units are crime counts, these are *not* conventional crime indices. And, a very important feature is that each plotted observation represents a moment in time.

To help make the discussion concrete, suppose the data come from a large American city. The spatial units are police precincts, and the temporal units are days. Further suppose that the first principal component (PC1) summarizes the *temporal profile* of gang violence, and the second principal component (PC2) summarizes the *temporal profile* of commercial burglaries. The temporal profile for PC1 might be dominated by the ebbs and flows of gang rivalries. The temporal profile for PC2 might be dominated by smooth, repeating ups and downs, perhaps distinguishing between weekdays and weekends. Note that the principal components are not responding to shared variation in a set of S different variables from T cross-sectional units, which is the usual application of PCA. There are not, for example, 20 variables from the census for large U.S. cities. The "variables" here are crime

[3]A modern discussion of PCA as a special case of a much large set of data reduction procedures can be found in Gifi (1996).

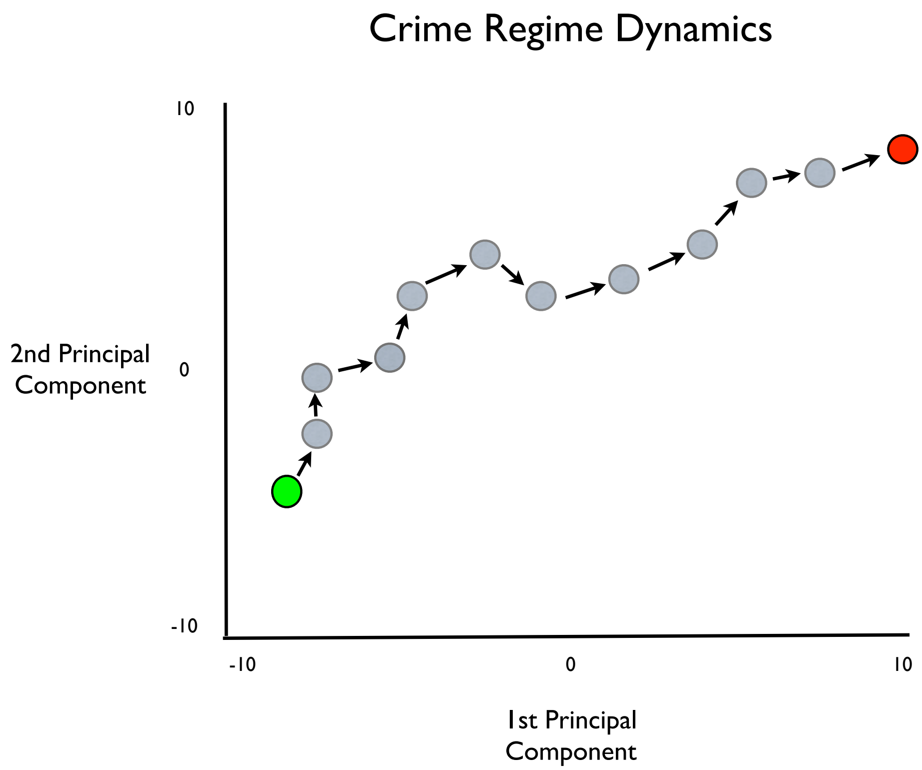

Figure 3: Smooth Movement of Crime Regimes Over Time

counts in different locations for different points in time.

Insofar as other kinds of crimes beyond gang violence and commercial burglary are not represented, they do not possess time profiles that are similar to those of these offenses. If the gang violence and commercial burglary principal components capture most of the variance, it can mean that in this city at the time the data were generated, other crimes did not have temporal profiles that were similar to one another. If there were such profiles, there would be additional principal components to plot and examine.

To get an initial sense of what Figure 3 conveys, consider the green circle toward the left hand side of Figure 3. It is the first day in the data set and is characterized by being relatively low on PC1 and PC2. There are roughly 10 fewer crimes of gang violence than the average and roughly 5 fewer commercial burglaries than the averate. Consider the red circle toward the right hand side of the figure. It is the last day in the data set and is characterized by being relatively high on PC1 and PC2. For both, there are about 10 more crimes than the average. The gray points that fall in between will be discussed shortly.

3.2.1 Crime Regimes Defined

We are now ready to define a crime regime for purposes of this paper. *Each possible location in a space defined by its principal components is a crime regime.* Thus, a space defined by two principal components contains all possible crime regimes along those two dimensions. Conceptually, the plane is continuous so there are in theory a limitless number of crime regimes in this two-dimensional plane. In practice, there will a large, but finite, number because there are limits to the precision with which the principal components are measured; the surface is not continuous in practice. Consistent with Figure 3, the actual number of crime regimes that materialize in a given data set, what one might call the *realized* crime regimes, will usually be much smaller still. In order to keep the prose from becoming unnecessarily cluttered, we will only use the qualifier "realized" when we need to make clear that the crime regimes being considered come from real data.

3.2.2 Crime Regime Dynamics

The arrows overlaid on Figure 3 connecting the points are meant to convey the passage of time from the earliest observation to the latest. In this idealized example, there are 12 days, although in practice that would likely be far too few to capture interesting over time patterns in crime. The arrows allow one to follow crime regimes over time and discern such things as whether crime regimes that are close in time tend to be alike, and whether some regions in the principal component space act like attracters around which realized crime regimes gather.

Figure 3 shows a relatively smooth temporal pattern. Both kinds of crime are gradually increasing over time. This is the sort of pattern often found in studies of crime trends and can lead to explanations in which the relevant predictors also change gradually. Neighborhood demographic changes are one example. It would also be possible, of course, to have violent crime increasing and commercial burglaries decreasing if, for instance, commercial enterprises were leaving a neighborhood.

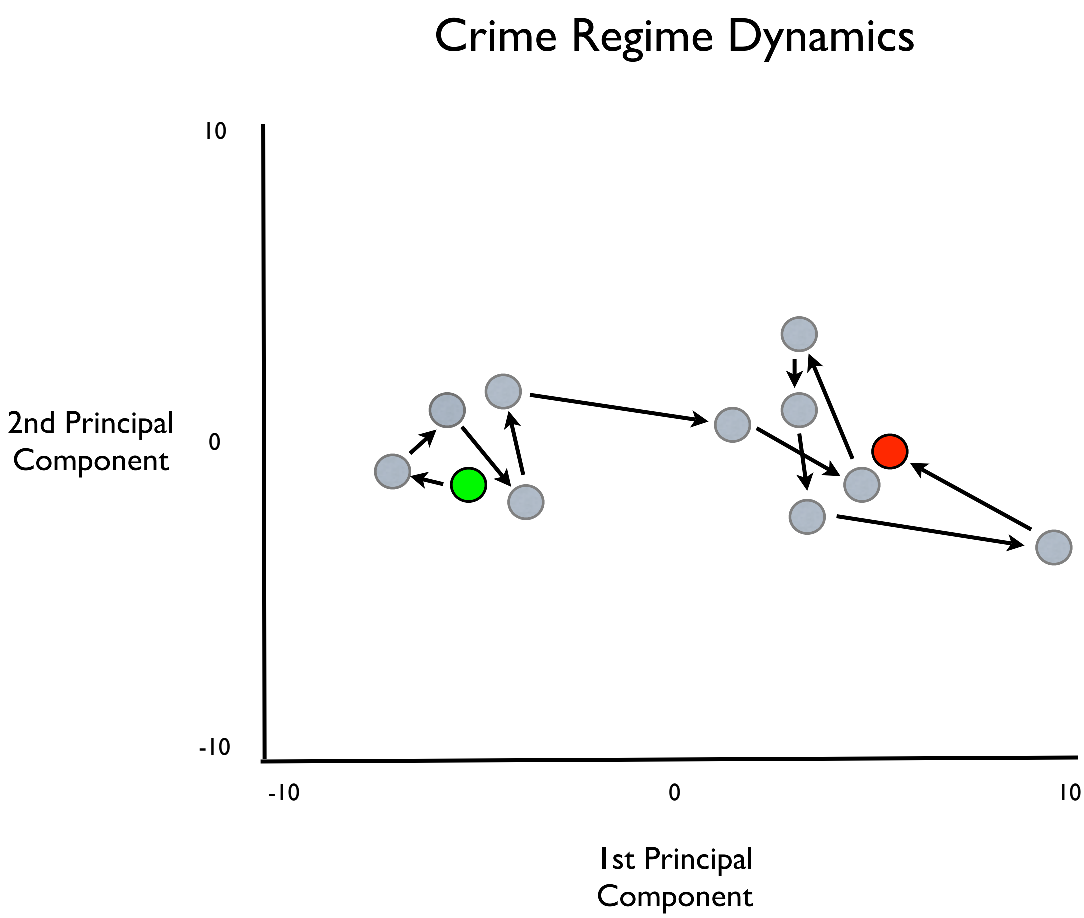

Figure 4: Clustered Movement of Crime Regimes Over Time

Alternatively, very different and very challenging patterns might materialize. Consider Figure 4. For the first five days, the realized crime regimes cluster in a region where the number of violent gang-related crimes is below average, and the number of commercial burglaries is about average. Between the fifth day and the sixth day, there seems to be a transition. On days six through twelve, the realized crime regimes cluster in a region where the number of commercial burglaries is still about average, and where there is a larger number of violent gang-related crimes than average. In addition, the eleventh day appears to be an isolate at the far right hand side of the plot representing a spike in crime along the first principal component. This could represent an especially ugly dispute between rival gangs, or some problem in the data itself.

It is important to appreciate that for Figure 4 there can be clustering in principal component space and in time but that the two are not necessarily linked. Realized crime regimes will tend to gather in certain regions of the principal component space insofar as the marginal distributions of each principal components are not rectangular. Where there are greater densities of observations on both dimensions, there will be clustering in the plot. For example, if there are a large number of observations that have much higher than average counts on the first principal component, and if there are a large number of observations that have much higher than average counts on the second principal component, there will be a large number of crime regimes clustering in the upper right hand side of the Figure 4. But such clustering does not represent an association between the two principal components because by construction, they are orthogonal. Also, such a cluster by itself is silent on the role of time. The cluster may be created by crime regimes from days that are close to one another or not. For example, a cluster may be composed of crime regimes falling on 20 consecutive days or of crime regimes that are each several months apart. In Figure 4, there is clustering in time as well as space. The early, realized crime regimes are located to the left, and the late realized crime regimes to the right. This can have important implications, as we will soon see.

Beneath the patterns in Figure 4, there could be interesting information to be found plotting each principal component over time. One might see, for instance, sharp and erratic shifts in PC1, with gang rivalries and police crackdowns a possible cause. Or one might see cyclical increases in PC2 associated with weekends or national holidays, perhaps because there can be long intervals during which commercial buildings are unoccupied. Such patterns can help to provide interpretations of the principal components and explanations for the dynamics shown in Figure 4. For example, the transition between the fifth and six day might be explained by a long memorial day weekend. Unexplained however, would be why the cluster persists for more than three days.

Given the content of our expository illustration, one might rightly wonder why one needs to plot the principal components against one another. Why not just plot a count of the number of violent gang crimes against a count of the number of commercial burglaries for the city as a whole? If one knew in advance that there were similar temporal profiles for gang violence across the police precincts, a simple sum for each day of the number of such crimes for the city as a whole might well suffice. The same would hold for commercial burglaries. But, how would that be known in advance? And how would one know not to pool commercial burglaries with residential burglaries, for example?

Moreover, there is no guarantee that the principal components would represent such intuitively clear distinctions. Somewhat more abstract time profile features might be captured. For example, the first principal component might represent changes in the density of the interactions between potential perpetrators and potential victims, and the second might represent the aggressiveness of police patrols.[4] Then, the movement into clusters may represent self-organization, and the movement out of clusters may represent the process of dissolution. This in turn could raise such questions as what kinds of temporal changes in crime regimes are reversible.

One might also wonder about how to proceed if more than two important principal components surface. For three principal components, there is the option of a three-dimensional plot. Such plots are available in many popular software packages.[5]

[4]Still higher levels abstractions might emerge, perhaps capturing some general properties of dynamical systems (Bergé et al., 1984). For example, one principal component could be a process of self-organization and another principal component could be competing pressures toward entropy (Prigogine, 1980). The application of such ideas to social phenomenon is controversial and risks what can be called reductionism. In this paper, we are agnostic on these issues.

[5]The main problem would be to keep the plot sufficiently uncluttered so that patterns in time and space could be easily seen. One nice device is to make the three-dimensional plot dynamic so that with the help of a computer, the space can be viewed from many different directions. Such tools exist in many software package with *GGobi* (www.ggobi.org/) one of the best.

If there are more than three important principal components or a three-dimensional plot is too cluttered to be effectively read, a good option is to examine all two-way plots. These are essentially projections of a high dimensional space onto a set of two-dimensional spaces, which can be displayed as a scatter plot matrix. In some implementations, one can mark particular observations so that they are highlighted automatically in each of the scatter plots. For example, one might mark the dates during which a curfew on teenagers was in place. Then one could more easily determine if there are any crime regime shifts during that period along any of the principal components. Yet another approach is to animate the plot. Time, for instance, could be captured by show the movement of crime regimes week to week.

3.3 Summary Statistics

One might hope that in many applications, the dynamics of crime regimes could be apprehend with little more than the tools already discussed. But there may also be instances in which summary statistics would be helpful. Basically, any summary statistics that one might apply to a scatter plot could be useful for describing crime regimes. Beyond statistical features of the principal component marginal distributions (e.g., mean and variance), the usual measures of association are prospects along with the full array of bivariate fitting procedures such as smoothers.

However, what may be most interesting about the scatter plots could be missed by conventional scatter plot statistics. Clusters, for example, can imply the need to apply spatial statistical tools. To take a simple example, one might want to report the ratio of the average Euclidian distance between clusters compared to the average Euclidian distance within clusters as a way to describe the degree of clustering.[6]

3.4 Characterizing Uncertainty

Insofar as one can make the case that one's data were generated by probability sampling or a stochastic process that is well understood, a wide variety of statistical procedures exist with which one might construct conventional

[6] These and many other ways to summarize spatial patterns are discussed in any number of fine textbooks and references (Cressie, 1993; Haining, 2003) and can be found in a variety of software such as in R (www.r-project.org/). Similar opportunities exist for characterizing the time dimension and again, there are ample resources readily available.

hypothesis tests and confidence intervals. In practice, however, the data for the study of crime regimes will rarely fit within such a framework. They are not likely to be a probability sample or the product of a well understand stochastic process.

However, one can examine certain *theoretical* properties that one thinks are responsible, at least in part, for the realized crime regimes. For example, one might wonder what the pattern of realized crime regimes would look like if time played no systematic role. A natural approach would be to randomly scramble the temporal order of the each principal component and plot the data again. A visual inspection of the two plots may then be all that is needed. It may then be readily apparent whether or not scrambling time leads to a similar looking scatter plot. If the plots look very different, one may conclude that time is systematically relevant.

A step toward more conventional statistical inference might then be made within the framework known as permutation tests (Good, 2004). One would begin with one or more summary statistics of interest, such as a measure of the degree of clustering. One would then simulate the distribution of each summary statistic under the assumption that time did not matter. If a summary statistic obtained from the original scatter plot was beyond either tail of the simulated distribution, one might conclude that time was actually an important feature of the how realized crime regimes are organized in principal component space. In effect, one simulates a null distribution, which in turn, provides a benchmark for the summary statistic obtained from the data on hand.[7]

3.5 Revisiting Geographical Variation

We have been focusing on temporal variation in crime regimes across geographical units. One might wonder about the reverse: focusing on geographical variation across temporal units. In theory, it would be a simple matter to transpose a data matrix like the one shown in Figure 1 and apply principal

[7] The basic idea is an old one, dating back to R.A. Fisher (Pitman, 1937; Wald and Wolfowitz, 1944). The null hypothesis is that the observations are independently and identically distributed with an arbitrary marginal distribution. The order statistic from the data is the sufficient statistic. Conditioning on the sufficient statistic, all permutations of the data are equally probable. The distribution of a test statistic under the null hypothesis is called the permutation distribution for that test statistic and can be used in a manner that has many similarities to conventional statistical tests.

components analysis, assuming that for the transpose there were not more columns than rows. If this is not the case, the temporal units and/or spatial units would need to be redefined.

However, it not clear what would follow. Geographical location would not automatically lead to a single, equal interval metric like time with which to order the observations; what would be the units of the principal components? A distance metric, for example, would need an anchor from which the distances for all geographical units could be computed. Like time, one would need a single quantitative variable to use as a metric. But without such an anchor, one would be left with a symmetric distance *matrix* containing the distances between each geographical unit. And it is not apparent how an instructive anchor point could be determined or even that a conventional distance metric would be appropriate to begin with.[8]

An alternative and promising strategy is to undertake separate crime regime analyses for different geographical areas. One might, for example, consider whether the temporal patterns of crime regimes are more alike for neighborhoods that have similar features than for neighborhoods that are different in important ways. Likewise, one might compare adjacent neighborhoods to neighborhoods that are a substantial distance apart. All of the concepts and procedures considered above could be adapted to allow for such options, but a serious discussion is well beyond the scope of this paper.

3.6 A Brief Interim Summary

The moment one acknowledges that crimes have many features, and that it can be instructive to look at the most important of those features simultaneously, existing approaches to crime in time and space become less satisfactory. In effect, there are now two or more "dependent variables" whose movements needs to be to be examined at once.

Although the raw data can be arrayed in conventional time and conventional (geographic) space, we exploit space that is defined by variables not locations. This allows us to summarize shared temporal patterns over a large number of spatial units. These summaries are the few, leading principal components that in turn provide a formal way to define crime regimes and a

[8]Recall, the crime patterns in cities are often shaped by the manner in which potential perpetrators and potential victims move about. Ease of travel and travel time (e.g., public transportation routes) can be far more important than literal distance.

space in which crime regimes move. The rest — temporally shuffled scatter plots, summary statistics, and statistical tests — is just details.

The movement of crime regimes in principal component space is the key source of insights about the dynamics of crime. Where in the principal component space are there crime regimes? What are the patterns of movement? Are there clusters in time? Answers to these and other related questions are meant to stimulate new thinking about how crime is produced and how crime changes. Whether these claims have merit, however, depends on how they hold up in practice. It is time, therefore, to consider an application.

4 Dynamic Crime Regimes for the City of Los Angeles

We will illustrate and extend the concepts and tools just discussed with data from the City of Los Angeles. The raw data are based on offense reports which specify the kind of crime, the time of day, the date, and the location, as well as many other features of the crime, the victim(s) and the perpetrator(s). There are eight years of data: from January 1st, 2000 though 2007.

4.1 Determining the Temporal and Spatial Units

An initial issue is what the most useful levels of temporal and spatial aggregation should be. On the one hand, if the data are made too coarse, some interesting patterns might be obscured. On the other hand, data that are not coarse enough might lead to crime counts that are very small, too often zero, and significantly perturbed by noise. For example, some preliminary work at the spatial scale of census tracts indicated that even when aggregated up to weeks, counts for common crime categories were typically indistinguishable from a simple Poisson process, even though at larger spatial scales, systematic temporal patterns were evident. Aggregating in time and/or space can help to cancel out noise.

It is also important to consider theoretical justifications for different temporal and spatial scales, as well as what scales make sense from a policy perspective. For example, census tracts do not necessarily correspond to neighborhood boundaries that Los Angeles residents recognize. Likewise, counts for six hour time intervals do not map well into the natural cycles of daylight and darkness or the average work day.

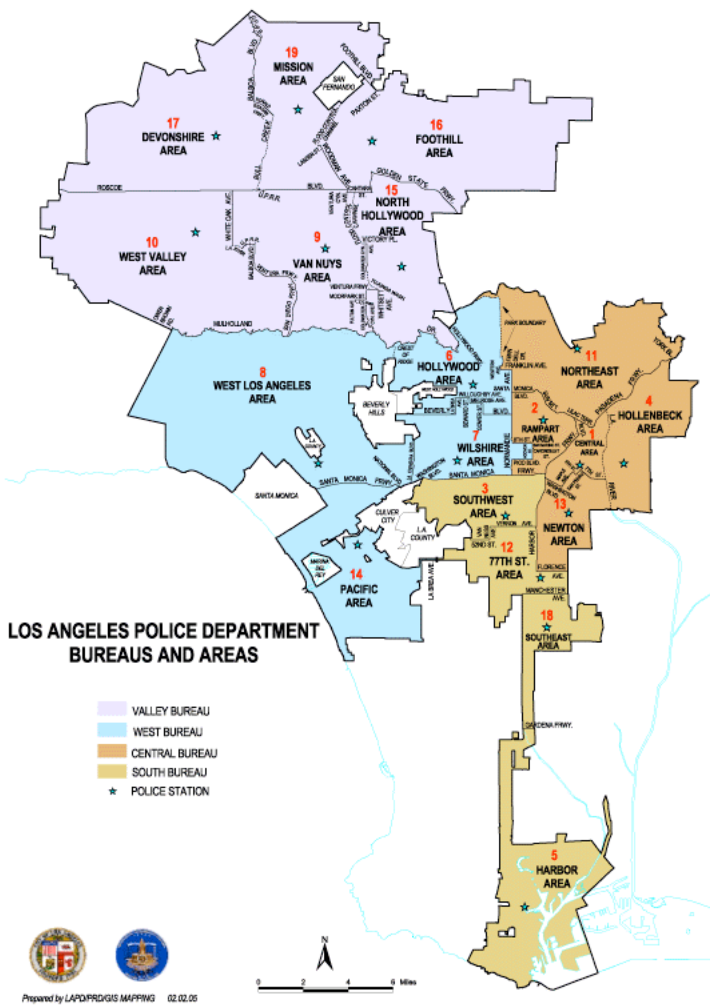

Figure 5: Police Divisions for Los Angeles

Trying to balance all of these issues led first to an initial decision to aggregate the temporal dimension to weeks, and the spatial dimension to police divisions. These units are consistent with practices of the Los Angeles Police Department (LAPD), which often allocates its personnel and other assets on a weekly basis at the level of the 19 police divisions. If one of the factors affecting crime regimes is police practices, using other temporal or spatial units could have made it more difficult to find evidence of police impact. A map of the police divisions is shown as Figure 5.

Even with nearly 4 million crimes in the data set, the data were very sparse when spread over 19 police divisions and 424 weeks. As a way to reduce the sparsity yet construct a data set of some interest, two simple indices were constructed: (1) the sum of homicide, rape, robbery aggravated assault and (2) the proportion of those crimes in which a handgun or assault weapon was used. The first can be seen as a measure of the amount of violent crime, and the second can seen as a measure of its potential lethality. Our hunch was that the two indices would have somewhat different temporal patterns. In future work, crime attributes other than the kinds of weapons used will be considered.[9]

The 19 police divisions were reduced to 18 because one division (i.e., Mission) did not exist before 2005. We were also constrained by the requirement of PCA to not have more columns in the data set than rows. With weeks as the temporal unit, there were 424 rows. With 18 police divisions and 2 kinds of crimes, there were 36 columns. Had we used the LAPD reporting district as the spatial units and the two crime categories, there would have been over 1500 columns. In short, our temporal and spatial units reflect a number of compromises that we will revisit in the future as more data for Los Angeles are made available.

One might wonder if other temporal and spatial units could lead to different results. The answer should be yes, and that is a good thing. It seems unnecessarily risky to assume that crime dynamics and their causes are the

[9]One might consider applying a data reduction tools like principal components analysis to the crimes within each of the indices rather than computing their sums. We considered this and rejected it for three reasons. First, the data would have become very sparse and more perturbed by noise. Second, another layer of data reduction would have placed more distance between the raw data and summaries of important patterns. Finally, the results we obtained seemed sensible and easily interpreted; there seemed to be no problem that needed fixing. Of course, this decision can be revisited in future work, and researchers with different data might quite properly proceed in other ways.

same regardless of scale at which they are studied. And insofar as different scales show different patterns, the challenge to criminological theory is that much richer.

4.2 Results

A principal component analysis produced one dominant dimension and one other of considerable statistical importance. Together they accounted for well over two-thirds of the variance. No other principal component accounted for more than 4% of the variance and most accounted for far less.[10]

The two leading principal components represent a summary of temporal profiles for the number of violent crimes and the proportions in which a handgun or assault weapon were used over police divisions for the City of Los Angeles as a whole. In the process, 36 columns were reduced to 2. Had other measures of crime been used, the results could have been rather different.[11]

Figure 6 is a biplot for the first two principal components. Because counts and proportions are very different units, we have followed the conventional practice of transforming all of the variables into standard scores. The axes on the bottom and to the left are for the weights used to construct the principal components. The axes at the top and to the right are for the principal components themselves. Both are in standard deviation units. The plot symbols are numbers denoting the week in which the crime incident occurred: 1 through 424. Each vector's direction and length represent the weight used to construct the principal components. The direction captures the associations with the two principal components. For example, if a vector is parallel to the principal component on the horizontal axis, that column in the data matrix is associated with that principal component and not with the principal component on the vertical axis. A vector that is at 45 degrees to the principal component on the horizontal axis has an association of the same magnitude with both principal components. And, the longer the vector, the larger the weight.

[10]The principal components procedure was implemented with singular value decomposition in the statistical programming language R (www.r-project.org/). Singular value decomposition is generally the preferred method for numerical accuracy.

[11]In earlier work, for example, each police division was represented by the number of violent crimes per week and the number of serious property crimes per week. Then, three important principal components surfaced, not two.

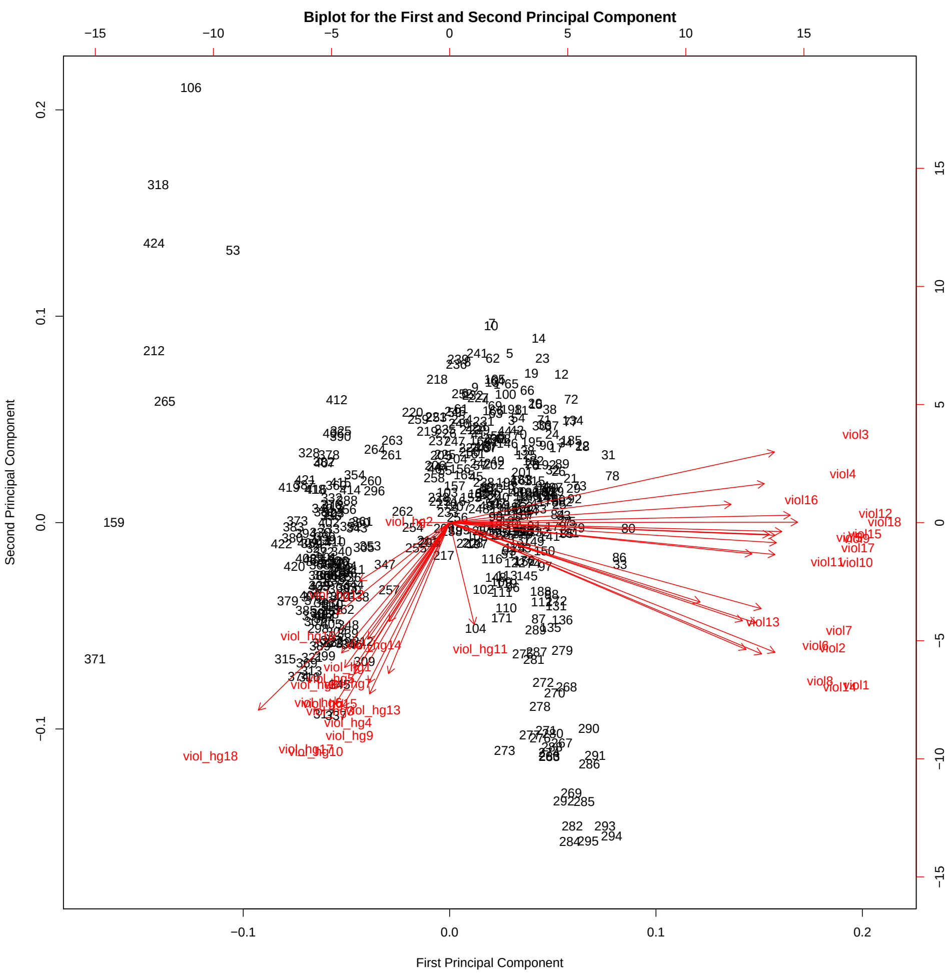

Figure 6: A Biplot of the Data and the Weights for the First and Second Principal Component

The first principal component is most strongly associated with the violent crime index. The vectors viol1, viol2, ..., viol18 are roughly parallel to the axis of the first principal component. The second principal component is most strongly associated with potential lethality. The vectors viol_gh1, viol_hg2, ..., viol_hg18 are roughly parallel to the axis of the second principal component. The vectors for the number of violent crimes and the vectors for the proportions in which a handgun or assault rifle were used are approximately orthogonal. They represent rather different features of crime regimes. That is, temporal variation in the numbers of violent crimes is nearly unassociated with temporal variation in potential lethality; weeks when there is a greater number of violent crimes are neither more nor less likely to be weeks in which potential lethality is worse. Even by itself, this may be a finding of some interest and begins to illustrate the possible benefits of a crime regimes perspective.

Insofar as the two sets of vectors are not exactly orthogonal, there is a slight tendency for the potential lethality vectors to be negatively related to the vectors for the number of violent crimes.[12] This may represent how the mix of violence crimes is associated with the choice of weapon. Aggravated assault is the most common of the violent crimes included in our violent crime counts. Among those crimes, it is also the one in which the use of a handgun is least likely (Rand, 1990). Therefore, when variation in aggravated assault is an important driver of our overall counts, increases in the counts could be plausibly associated with less use of handguns and assault weapons.[13]

4.2.1 Displaying the Results

Figure 7 is a scatter plot of the first two principal components. The first principal component is on the horizontal axis. The second principal component is on the vertical axis. The units for both are in standard deviations because of standardization. All observations are numbered chronologically from week

[12] A normalized eigenvector is only determined up to a factor -1, that is, a simultaneous change of all signs. For singular value decomposition in particular, pairs of (normalized) singular vectors u and v are only determined up to a simultaneous change of all signs. That is, the pairs (u, v) and $(-u, -v)$ are equivalent with no material change of interpretation.

[13] The issues this raises generalize. If one were to consider even a single kind of crime defined by statute, there would still be heterogeneity, perhaps very important heterogeneity, within that collection of crimes. Not all homicides, for instance, are alike. Consequently, the mix of crime types as well as their numbers can change over time. And they can change in different way in response to different factors.

1 to week 424. Observations that are adjacent in time are connected with a colored line. Three different regions in the plot are shown by color and are also labeled.

Ignoring the role of time for the moment, it is readily apparent that some regions in Figure 7 are unpopulated. For example, there are no realized crime regimes that are very high on the first principal component and very high on the second principal. Likewise, there are no realized crime regimes that are very low on potential lethality and other than very high on the amount of violent crime.

One reason is that the marginal distribution of the principal component for the proportion of violent crimes in which a handgun or assault weapon was used is effectively unimodal with the greatest mass of the data located around the mean and a long tail towards smaller proportions. At the same time, the marginal distribution of the principal component for the numbers of violent crimes is bimodal. Where the modes of the two distributions tend to cross, there will necessarily be a higher density of observations in their joint distribution. But as a substantive matter, why would such marginal distributions materialize? For example, what is it about the social processes that generate violent crime that would lead to especially high concentrations of observations on either side and some distance from the mean? We return to this issue shortly when we consider the role of time.

One can see in Figure 7 several isolated observations toward the left side of the plot. These are artifacts of how the data were assembled. When the data were aggregated by week, one would often find that the last week of the year did not have seven days. So the violent crime counts for these weeks would tend to be biased downward. There are a variety of adjustments that could have been made, such as scaling up the counts to adjust for the length of the week. However, we decided to leave the counts untouched to show that our methods can reveal certain kinds of artifacts and that a few extreme outliers do not affect one's ability to see where and how the data cluster. Dropping the outliers does not materially affect the substantively important patterns in Figure 7 nor the story that follows.

Clearly, there are three temporal clusters in the scatter plot. For first 265 weeks, the realized crime regimes hover in a region that is a little higher than average for the amount of violent crime and around average for potential lethality. We call this region 1, and the connecting lines are red.

Beginning in week 264 (early in 2004), there is a rapid and dramatic shift in the region occupied by the realized crime regimes. The green cluster

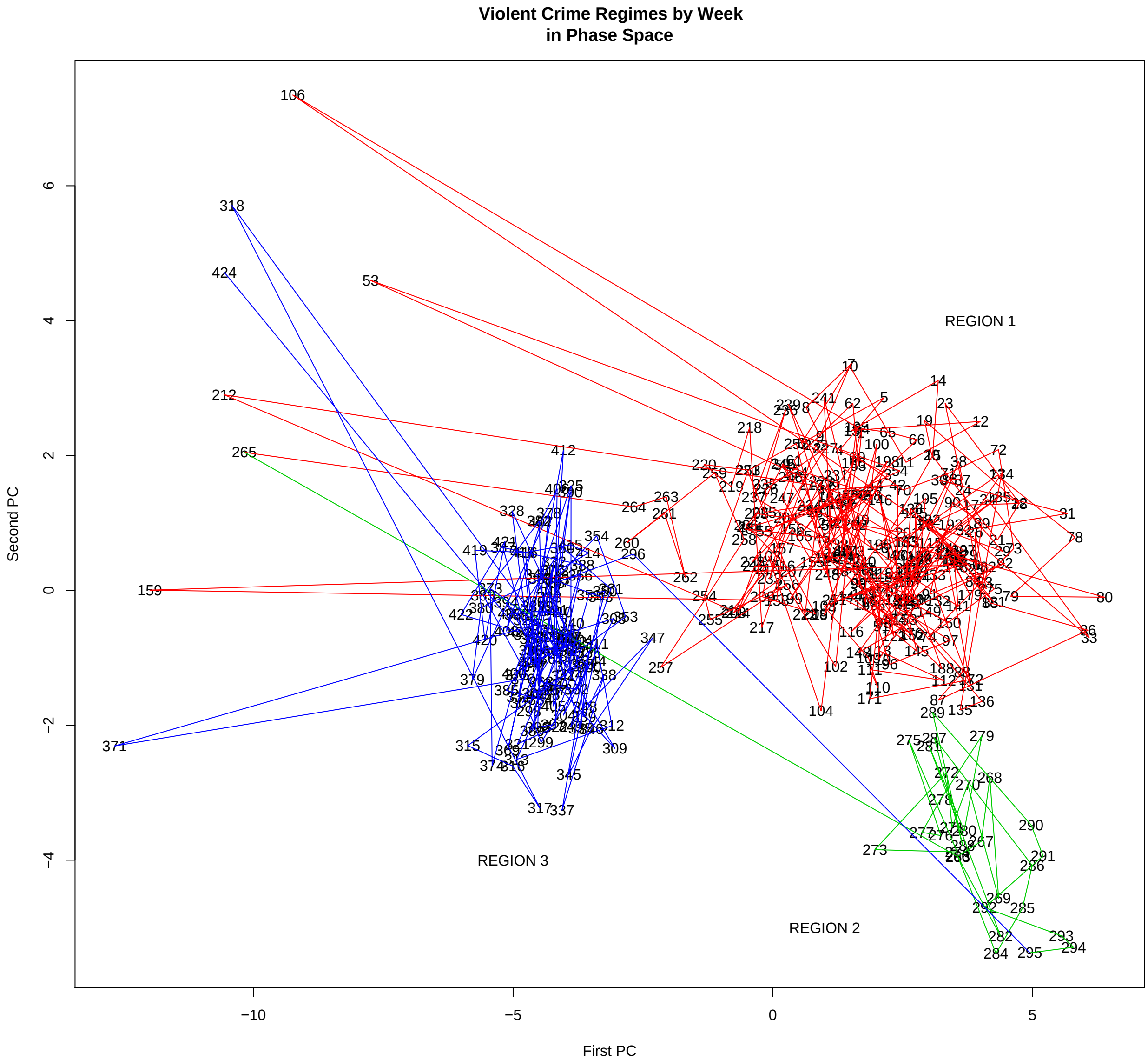

Figure 7: Crime Regimes in Two-Dimenstional Space: First and Second Principal Components

(Region 2) toward the lower right of the plot represents an area where there is a substantial increase of about two standard deviations on the average in the principal component associated with the amount of violent crime. At the same time, there is a substantial decrease of about four standard deviations on the average in the principal component associated with potential lethality. The pattern is maintained for about 30 weeks.

With the start of week 295 (around the end of 2004), there is another rapid, dramatic shift to to the blue area and Region 3. There is a very dramatic drop of about ten standard deviations in the first principal component and a substantial increase of about three standard deviations in potential lethality. There are far fewer violent crimes, but the use of handguns and assault weapons becomes more common.

Two questions naturally follow. First, why do crime regimes cluster in a given region for many weeks. There is variation within a cluster to be sure, but no crime regimes stray very far and behave as if pulled back toward the mass of the data. The beginnings of possible explanations are provided later.

The second question is what accounts for the rapid transitions. A grounded explanation might be found in factors that also can change quickly such as gang tensions and police practices. A strictly enforced curfew or an aggressive stop-and-frisk policy, for instance, could lead to abrupt changes in the numbers and kinds of violent crimes. A more abstract and perhaps more intriguing explanation is that the crime regimes are moved about as part of a highly nonlinear system. Small changes in inputs can lead to large changes in outputs that materialize as abrupt regime shifts (Strogatz, 1994: Part III). Such patterns would be somewhat akin to neighborhood tipping effects, but on much smaller time scales of a week or two. We return to this question as well later.

4.2.2 Showing the Importance of Time

Visual inspection of Figure 7 should make a convincing case. There are three distinct clusters of realized crime regimes defined by their location in the plot combined with a chronological ordering. It is important to appreciate that both features of these clusters are significant because there could be distinct groupings in which the observations were not in time order. The strays toward the left side of the plot can be thought of in this manner. Some are located in about the same region but are widely separated in time.

It is possible to establish a temporal baseline from which to compare

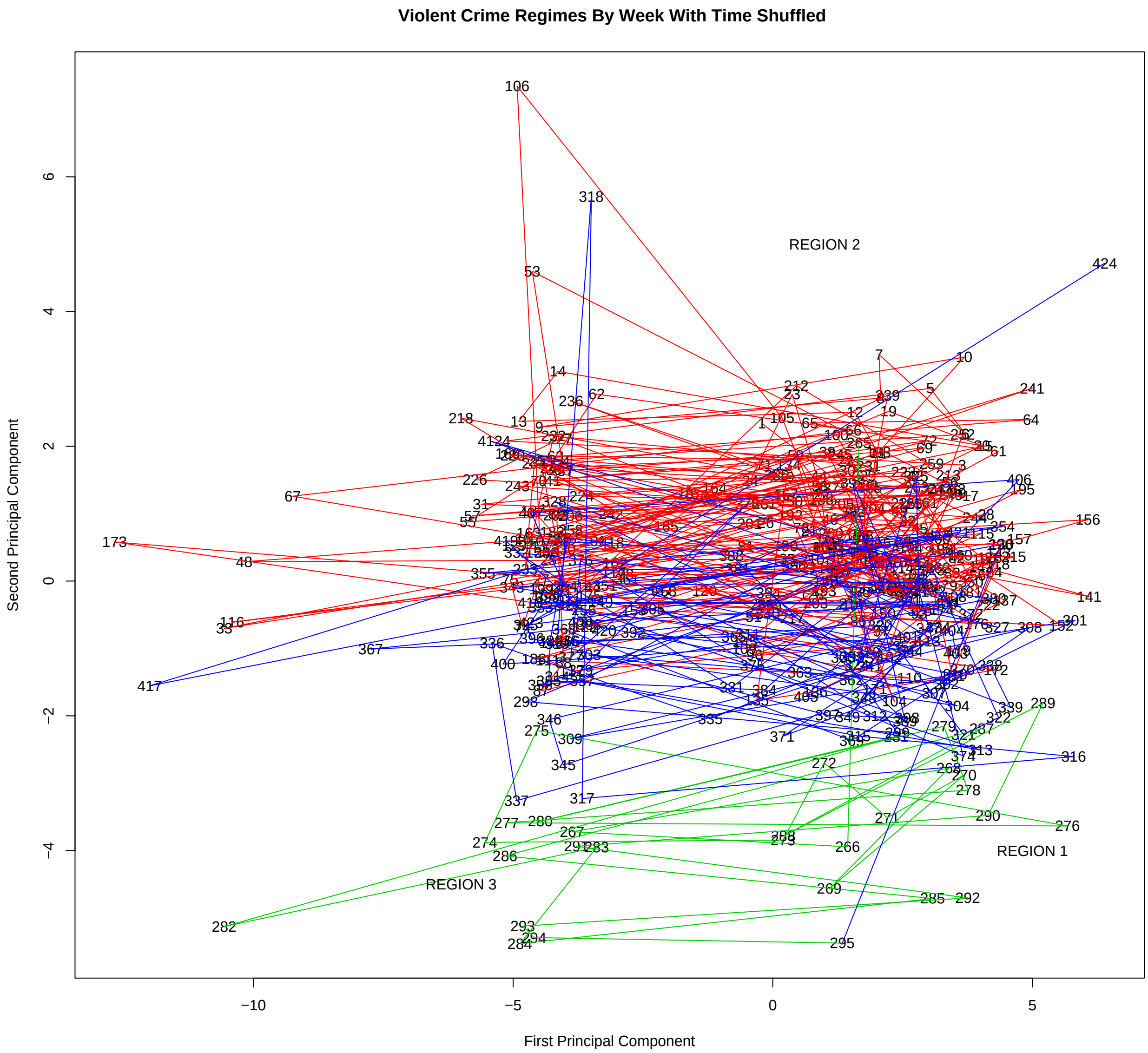

Figure 8: Crime Regimes with Time Shuffled: First and Second Principal Components

Figure 7. Suppose that time did not matter whatsoever. What would such a plot look like? It is easy to construct an instructive plot by shuffling time randomly. Figure 8 is the result. The pattern of lines in Figure 8 looks very different from the pattern of lines in Figure 7. No one would confuse the two. Figure 7 shows what clustering in time can look like. Figure 8 shows what happens to the realized crime regimes when there is no clustering in time.

4.2.3 Attaching Summary Statistics

It can be useful to complement plots like Figure 7 and Figure 8 with descriptive statistics. A good place to begin is with a quantitative definition of a cluster. There is no "right" definition, but some definitions are more instructive than others. Building on outlier detection traditions (Cook and Weisberg, 1999: sections 1.5.2 and 15.4), our definition has three components.

1. The metric will be Euclidian distance between realized crime regimes.

2. The earliest point in a realized crime regime cluster must be at a distance of at least three standard deviations from the immediately preceding realized crime regime.

3. There are at least three immediately subsequent realized crime regimes with no more distance between each pair of temporally adjacent crime regimes than one standard deviation.

There are no doubt many sensible variations on this basic definition. For example, one might require more than four crime regimes in a cluster. We picked four because that is about a month, and some important law enforcement decisions are made on a monthly basis. Or, if the distribution of the Euclidian distances has a long right tail, one might use the interquartile range, rather than the standard deviation, as a measure of dispersion. Indeed, Figure 9, a histogram of the Euclidian distances between temporally adjacent crime regimes, is very long tailed. In our case, however, using the standard deviation is conservative. It is more difficult to find definitive patterns.

The mean of the distances between temporally adjacent crime regimes is about 1.7 standard deviations. Using that value and our three-part definition, all of the clusters pass muster. Even with the short week counts included, the clusters in Figure 7 are easily identified by our formal definition. If

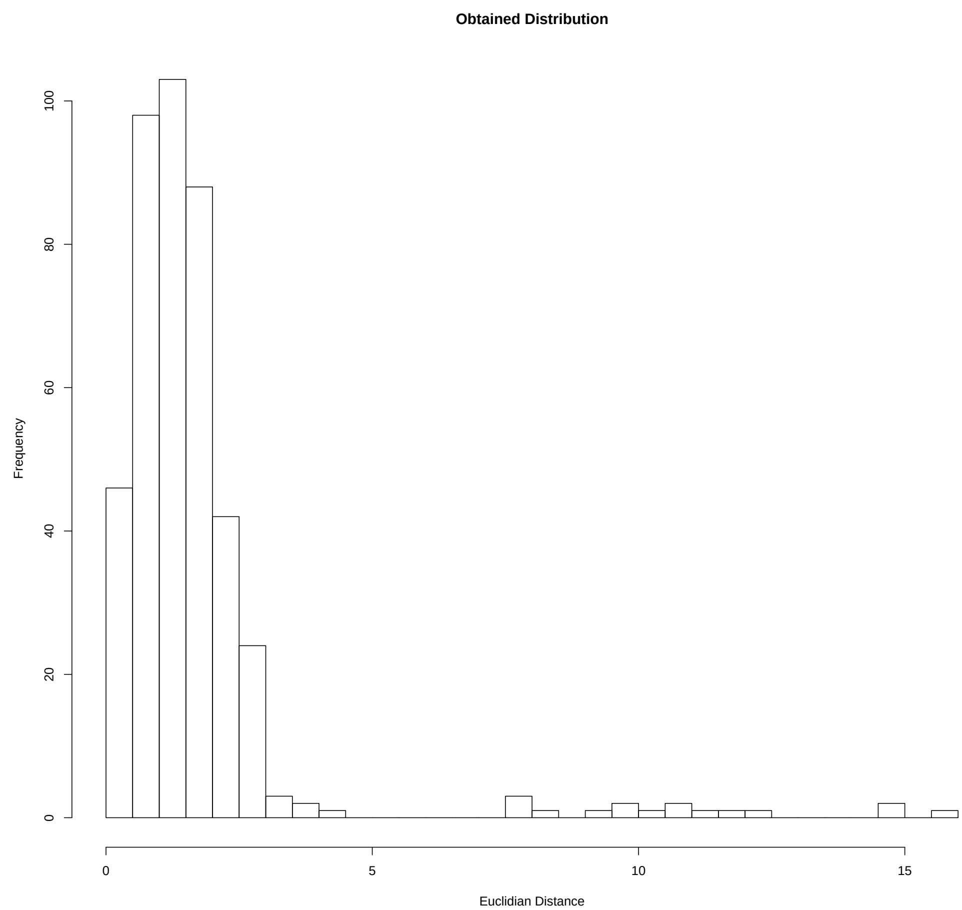

Figure 9: Distribution of Distances between Adjacent Points in Standard Deviation Units

the outliers are dropped, or the mean is replaced by the median, the average distance become approximately 1.3. The identification process becomes more powerful. In contrast, there are by our definition no clusters in Figure 8.

4.2.4 Time and Uncertainty

The contrast between Figure 7 and Figure 8 is so stark that one would probably not think that Figure 7 could be easily produced by chance through the shuffling used to construct Figure 8. But for the skeptic, is it relatively easy to address that question within a framework that is broadly consistent with conventional statistical inference.

We begin with the need to define a useful test statistic. Again, there is no right answer, but some test statistics will be more useful than others. From Figure 7 and Figure 8, it is apparent that the average the Euclidian distances between time-adjacent observations is greater for the latter. Because of the time shuffling, time-adjacent crime regimes will tend to be farther apart in the two-dimensional principal component space.

Figure 10 is a histogram of the Euclidian distances after shuffling between temporally adjacent observations. Compared to Figure 9, the distribution is far less skewed and shifted to the right. A reasonable inference is that the typical distances distances after shuffling are greater than the typical distances before shuffling. And in fact, the mean after shuffling is about 4.6, which nearly three times as large as the mean without shuffling.

But how likely is it that one could get an observed mean of 1.7 or smaller from a permutation distribution centered in 4.6? If one is prepared to think about this distribution as a product of random shufflings of the times linked to each realized crime regime — which is reasonable if the comparison between Figure 7 and Figure 8 is instructive — such as distribution is easily simulated.

Figure 11 shows the distribution in standard deviation units of the Euclidian distances based on 10,000 reshufflings of time. This is just a conventional permutation approach that can serve as the basis for statistical tests (Good, 2004). In this case, the observed mean of the Euclidian distances is 1.7, and the smallest mean of the Euclidian distances from the simulated permutation distribution is about 4.2. One can conclude that the chances of getting an observed mean as small or smaller than 1.7 from a null distribution of means centered on 4.6 is less than 1 in 10,000 (or a p-value of .0001 level), and in fact a lot less. The pattern of distances seen in Figure 7 is extremely unlikely to be the result of happenstance.

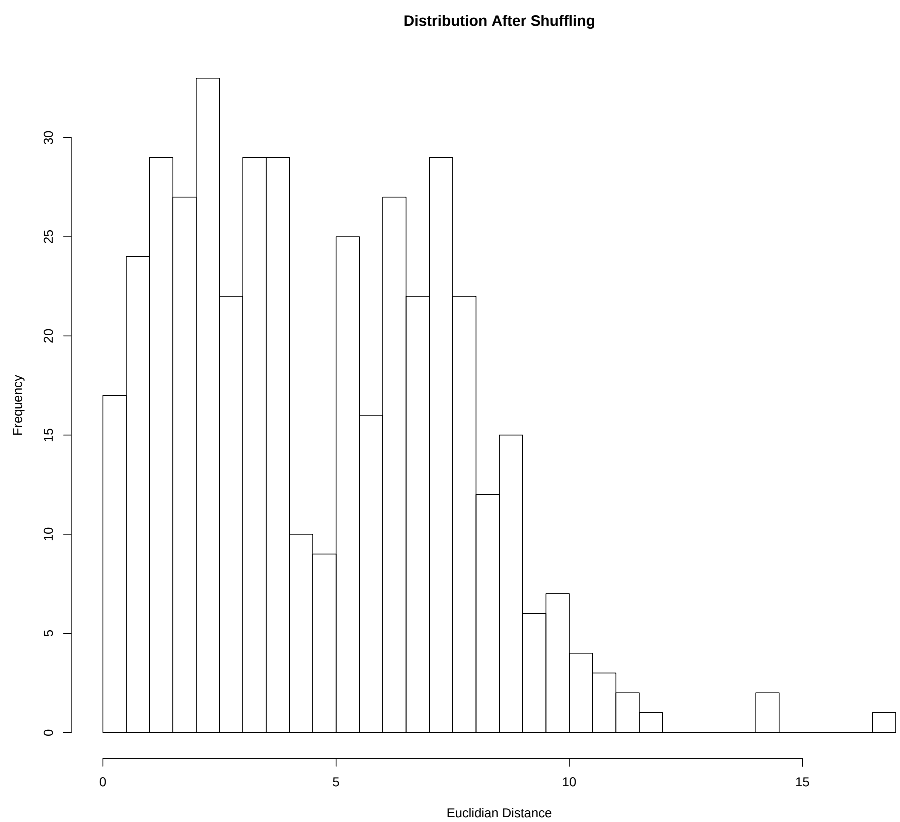

Figure 10: Euclidian Distances between Adjacent Points in Standard Deviation Units with Time Shuffled

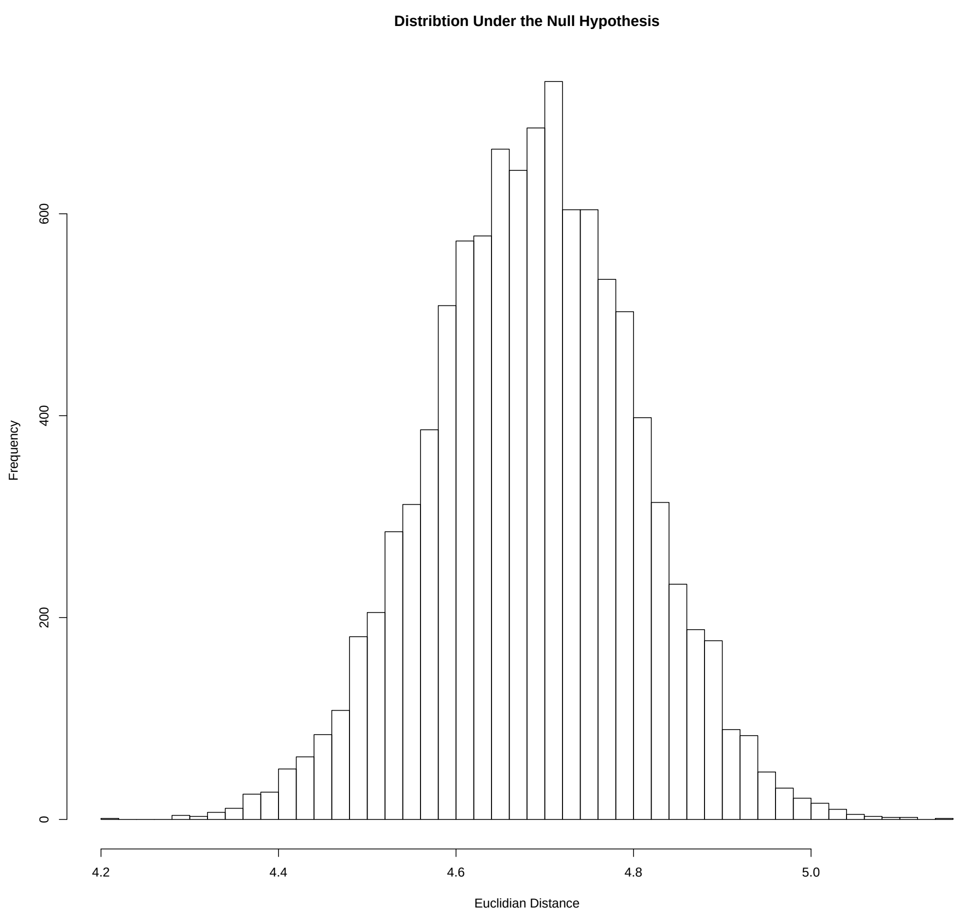

Figure 11: Distribution of the Mean Under the Null Hypothesis

The same approach can be taken for parts of Figure 7. For example, the third region (in blue) may correspond to a period in which the police changed their practices, including such things a putting many more police officers in high crime areas. There looks to be a sharp decline in violent crime. Could this be happenstance?

The strategy just illustrated can be undertaken for just those realized crime regimes constituting this cluster. The mean of the Euclidian distances within the cluster is about 1.6 standard deviations. The mean of the Euclidian distances in that same cluster after shuffling is around 5.4 standard deviations. Obtaining a mean of 1.6 from a null distribution with a mean of 5.4 is highly unlikely.

5 Moving Toward Explanations

5.1 Why Clustering?

Consider again the the clustering. Why does a sequence of realized regimes reside in one small portion of the crime regime space for many months? One possible approach is that variation within a cluster results from nothing more than independent random perturbations in the two principal components around the two within cluster principal component means. That is,

$$pc_{1,t} = \overline{pc}_1 + \varepsilon_{1,t-1}, \tag{1}$$

and

$$pc_{2,t} = \overline{pc}_2 + \varepsilon_{2,t-1}, \tag{2}$$

where pc stands for principal component, $\overline{pc}$ represents the within cluster mean of the principal component, the first subscript denotes the pc, the second subscript denotes time, and ε is a random perturbation, generated independently within and across the two equations and with an expectation of 0.

Should equations 1 and 2 apply, a number of questions naturally arise. Why, for example, would each principal component's perturbations be independent or one another. What it is about the generation of violent crime and its potential lethality that make the perturbations independent? One might well have expected at least some temporal dependence; one week's random perturbation might be related to next week's random perturbation.

A wide variety of other formulations might apply. For example, one might have anticipated

$$pc_{1,t} = pc_{1,t-1} + \varepsilon_{1,t-1}, \tag{3}$$

and

$$pc_{2,t} = pc_{2,t-1} + \varepsilon_{2,t-1}, \tag{4}$$

which would be two-dimensional random walk (Berk, 2003), but that is not what Figure 7 seems to show. The clustering indicates that crime regimes do not wander off.

Another formulation, perhaps consistent with Figure 7, is that there is on the average a negative dependence between time-adjacent locations within a cluster. Random perturbations pushing a crime regime away from the statistical center of the cluster tend to be followed by perturbations drawing the next crime regime toward the statistical center of the cluster. In equations 1 and 2, the random perturbations within each equation could then be negatively related for adjacent weeks. In fact, such negative dependence is not consistent with the data, but were the crime generation process to be homeostatic within a cluster, such a model might well apply. And if so, by what crime-related mechanisms would this occur?

In short, in this instance equations 1 and 2 seem to be at least descriptively useful. But the key point is this: the existence of the clusters needs to be explained not just in statistical terms, but subject-matter terms. We are pursuing this in ongoing work.[14]

5.2 Why Large Discontinuities?

Perhaps the dominant temporal story from these data is the rapid and dramatic shifts in realized crime regimes from one cluster to another. Explanations may take two generic forms. One kind of explanation would build on rapid shifts in potential predictors. If predictors with this property could be found that showed changes at appropriate times, the transitions might be easily understood. For example, a rapid escalation in gang violence could affect both the number of violent crimes and their mix. The latter could

[14]The same sorts of issues would arise were there only one cluster. We might have seen no sharp discontinuities. All of the crime regimes would then be moving about within a single region of the principal components space. The natural question is the same: why? What mechanisms could be drive such a pattern?

lead to changes in proportion of violent crimes in which a handgun or assault weapon is used.

This implies a data analysis strategy building on an interrupted time design but with a two-dimensional response variable. Because the principal components are orthogonal, the two-dimensional response variable could perhaps be analyzed as two, separate one-dimensional response variables. These ideas generalize well to principal component spaces of more than two dimensions.

These kinds of considerations can lead to a second, complementary explanation that is deeper and more challenging. As noted several times, if a dynamical system has highly nonlinear relationships, one can observe very rapid changes in response variables even if the predictors variables are changing slowly and in very small increments. Because such small changes in inputs can lead to large changes in outputs, the system may appear to be be erratic and unpredictable. Moreover, until the small changes in inputs compound to produce large changes in output, one might well observe realized crime regimes hovering in a particular region of the crime regime space.

Finally, even if the first kind of explanation seems appropriate, the second may also need to be invoked. One can ask the causal question one step back in the dynamic causal system: why do the *predictors* change in such an abrupt manner? In addition, it may be useful to move to higher levels of abstraction even if the immediate measures have grounded interpretations. In short, we suspect that concepts from dynamical systems may help in the study of crime regimes and crime in time and space more generally.

6 Summary and Conclusions

Crimes come bundled with a number of attributes. The mix of these attributes can change over time and space. We have suggested that the concept of a crime regime can provide some theoretical leverage for these matters, especially when combined with ways to think about how the bundle of attributes can change. Key tools include the use of principal components analysis to determine the dimensions of crime regimes, special scatter plots that help reveal the role of time, summary statistics to quantify crime regime patterns, and permutation procedures to examine the role of chance. We used these tools to analyze crime patterns for the City of Los Angeles. The analysis raised many more questions then it answered, but perhaps that is the point. We were looking at temporal and spatial patterns of crime with a somewhat different lens.

What might be seen through that lens?

1. It can be important to begin with the space of possible crime regimes that is *not* populated by realized crime regimes. These regions may represent crime regimes that are not possible in practice, or at least unlikely. From this would follow the natural question of what it is about these particular crime regime that weighs strongly against some of configurations of attributes. For example, in well-controlled and stable drug markets, turf wars may be seen as bad for business. Consequently, there are few incidents of gang-related violence. Then, the region high in trafficking events and high in violence might have no crime regimes. Or, there may be no clusters in time at the extreme right of Figure 4 because after a serious incident, the police may engage in a "crackdown" and saturate the areas where the violence occurred (Sherman, 1990).

2. Are there clusters of realized crime regimes in particular regions of the principal component space or are clusters spread relatively evenly over that space? Clusters can imply that certain kinds of crime regimes dominate in the spatial-temporal processes responsible for the amount of crime. But what is it about the generation of crime that would lead to clusters of this sort? Such a discussion would begin by considering why the the marginal distributions of the principal components have the form they do. For example, if one of the principal components happens to be associated with commercial burglaries, and has a bimodal distribution, what is it about the nature commercial burglaries that would lead to that pattern?

3. If there are realized crime regime clusters in the principal component space, are the realized crime regimes also proximate in time? If so, it implies that there are dynamic processes at work in which time really matters. In the simplest case, realized crime regimes that are close in time will tend to have similar content. What exactly is producing that kind of result? For example, commercial burglaries may in part reflect patterns of success; one successful burglary may provide a strong incentive to undertake another one very soon. Likewise, one attempted or

successful homicide may lead to a retaliation and a process of violence contagion (Fagan et al., 2007).

4. Is there a tendency for realized crime regimes to move gradually over time in particular directions, perhaps in response to relatively slow-changing features of neighborhoods? What would such forces be? Gentrification could be one example of a slow process of emerging and transitional change (McDonald, 1986; Schuerman and Kobrin, 1986; Taub et al., 1984).

5. Do realized crime regimes tend to cluster in one region, jump to another region, only later to return to the first region. Were this seen several times, it could imply that there is a "home region" in the principal component space to which realized crime regimes can be drawn after they stray. The home region could represent a relatively stable equilibrium where crime regimes would reside were it not for occasional perturbations. Where in the principal component space is home region located and what is it about possible equilibrating processes that lead crime regimes to that location? And what are the forces that pull perturbed crime regimes back? For example, relatively enduring features of neighborhoods may produce stable crime regimes that are temporarily perturbed by outside forces such as a summer school vacation or a temporary rivalry between gangs over drug markets.

Although it is risky to speculate about the the uses to which crime regimes may be put, we suspect that the potential links to research on dynamical systems may be the most important long run contribution. Theory and research in much of criminology has been at least implicitly proceeding within a dynamical framework for much of its history. Crime regimes in time and space may help build a bridge between a very important literature in criminology and a host of useful concepts, theories, and tools increasingly used in the natural and life sciences.

Where then do we go from here? Insofar as concepts and tools introduced in this paper appear to have merit, there needs to be a constructive discussion about how they can be enriched and refined depending on the kinds of theoretical questions to be addressed and the kinds of data to which they will be applied. We have not provided a "methodology," let alone a recipe. We hope we have provided a useful way to think about crime. In addition, it will be important to try out the approach, and creative variations, on a wide variety of data sets. For example, many large cities could provide data of a similar richness and in a similar format to those we assembled for Los Angeles, and a variety of temporal and spatial scales should be considered. Researchers also will need to hand-tailor the details of their analyses in response to the data on hand. The definition of a cluster, for instance, could sensibly vary. Finally, the work should be grounded in existing theory and yet open enough to accept theoretical perspectives from outside of conventional criminology. In short, we hope this paper will help to stimulate innovation. We will have failed if the ideas represented become codified as a pull-down menu in the next release of popular statistical software.

References

Anderson, E. (1998) "The Social Ecology of Youth Violence." In M. Tonry and M. H. Moore (Eds.), *Youth Violence.* (pp. 65-104). Chicago: University of Chicago Press.

Austin, J., Naro, W., and T. Fabelo (2007) *Public Safety, Public Spending: Forecasting America's Prison Population 2007– 2011.* Washington, D.C., The Pew Charitable Trusts.

Beirne, P. (1987) "Adolphe Quetelet and the Origins of Positivist Criminology'". *American Journal of Sociology* 92(5): 1140-1169.

Bergé, P. Pomeau, Y, and C, Vidal (1984) *Order within Chaos: Towards a Deterministic Approach to Turbulance.* New York: John Wiley and Sons.

Berk, R.A. (2003) "Random Walk." In M.L. Beck (ed.) *Encyclopedia of Social Science Research Methods.* Sage Publications, 2003.

Berk, R.A. and T. Cooley (1979) "A Dynamic Decision Theoretic Perspective on Modeling the Performance of the Criminal Justice System." *Social Science Research* 8: 265-286.

Berk, R.A., Messinger, S., Rauma, D. and J. Berecochea (1983) 'Prisons as Self-Regulating Systems: The Case of California from 1851 to 1980." *Law and Society Review* 17(4): 547-586.

Blumstein, A. and R. Larson (1969) "Models of the Total Criminal Justice System." *Operations Research* 17(2) 199-232.

Blumstein, A, Rivara, F.P. and R. Rosenfeld (2000) "The Rise and Decline of Homicide - and Why." *Annual Review of Public Health.* 21:505-541.

Bonger, W.A. (1967) *Criminality and Economic Conditions*, New York: Agathon Press.

Bursik, R.J., Jr. (1986) "Delinquency Rates as Sources of Ecological Change." In J.M. Byrne and R.J. Sampson (eds.), *The Social Ecology of Crime* . New York: Springer Verlag.

Bursik, R.J., Jr. (1988) "Social Disorganization and Theories of Crime and Delinquency: Problems and Prospects." *Criminology* 26: 519-551.

Byne, J.M. and R.J Sampson (eds.) (1986) *The Social Ecology of Crime*, New York: Springer-Verlag.

Clarke, R. V. and M. Felson (Eds.) (1993). *Routine Activity and Rational Choice. Advances in Criminological Theory.* Vol 5. New Brunswick, NJ: Transaction Books.

Cohen, L., and M. Felson. (1979) "Social Change and Crime Rates." *American Sociological Review* 44: 588-08.

Cook, D., and D.F. Swayne (2007) *Interactive and Dynamic Graphics for Data Analysis.* New York: Springer.

Cook, R.D., and S. Weisberg (1999) *Applied Regression Including Computing and Graphics.* New York: John Wiley and Sons.

Cressie, N.A. (1993) *Statistics for Spatial Data*, Second Edition New York: John Wiley and Sons.

Eck, John and D. Weisburd (Eds.) (1995). *Crime and Place.* Monsey, NY: Willow Tree Press.

Elliott, D., Wilson, W.J., Huizinga, Robert J. Sampson, R.J. , Elliott, A. and B. Rankin (1996) "Effects of Neighborhood Disadvantage on Adolescent Development." *Journal of Research in Crime and Delinquency* 33: 389-426.

Fagan, J., Wilkinson, D.L., and G. Davies (2007) "Social Contagion of Violence." In *The Cambridge Handbook of Violent Behavior* (pp. 688-723) Daniel Flannery, A. Vazsonyi, I. Waldman (eds.) New York: Cambridge University Press.

Gifi, A. (1996) *Non-Linear Multivariate Analysis*, New York: John Wiley and Sons.

Good, P.I. (2004) *Permutation, Parametric, and Bootstrap Tests of Hypotheses*, Third Edition, New York: Springer.

Griffiths, E., and J. M. Chavez (2004) "Communities, Street Guns, and Homicide Trajectories in Chicago, 1980-1995 Merging Methods for Examining Homicide Trends across Space and Time." *Criminology* 42 (4):941-978.

Haining, R. (2003) *Spatial Data Analysis.* Cambrige: Cambridge University Press.

Henry, V.E. (2003) *The COMPSTAT Paradigm*, New York: Looseleaf Press.

Janowitz, M. (1975) "Sociological Theory and Social Control." *American Journal of Sociology.* 81: 82-108.

Kornhauser, R. (1978) *Social Sources of Delinquency.* Chicago: University of Chicago Press.

Land, K., McCall, P. L., and Cohen, L. E. (1990) "Structural Covariates of Homicide Rates: Are There Any invariances Across Time and Social Space?" *American Journal of Sociology.* 95, 922-963.

Land, K. Cantor, D. and S. T. Russell, (1995) "Unemployment and Crime Rate Fluctuations in the Post-World War II United States: Statistical Time Series Properties and Alternative Models. (pp. 55-79) In J. Hagan and R. D. Peterson (eds.), *Crime and Inequality.* Stanford: Stanford University Press.

McDonald, S. C. (1986) "Does Gentrification Affect Crime Rates?" In Albert Reiss Jr. and Michael Tonry (eds.) *Communities and Crime*, Crime and Justice: A Review of Research 8: 163-201.

Morrison,, D.F. (1976) *Multivariate Statistical Methods*, Second Edition, New York: McGraw Hill.

Nagin, D. S., Land, K. C. (1993) "Age, Criminal careers, and Population Heterogeneity: Specification and Estimation of a Nonparametric, Mixed Poisson Model." *Criminology* 31: 327-362.

Ousey, G. C., and M.C. Augustine (2001) "Young Guns: Examining Alternative Explanations of Juvenile Firearm Homicide Rates." *Criminology* 39: 933-968.

Park, R. E. (1915) "The City: Suggestions for the Investigation of Human Behavior in the City Environment." *American Journal of Sociology* XX: 577-612.

Petrie, C. (2008) *Understanding Crime Trends: Workshop Report.* Washingon, D.C. The National Academies Press.

Pitman, E.J.G. (1937) "Significance Tests Which May be Applied to Samples from Any Populations." *Supplement to the Journal of the Royal Statistical Society* 4(1): 119-130.

Prigogine, I. (1980) *From Being to Becoming: Time and Complexity in the Physical Sciences.* San Francisco: W.H. Freeman and Company.

Rand, M.R. (1994) "Guns and Crime." Report NCJ 147003, US Department of Justice, Bureau of Justice Statistics, United States.

Reuter, P., and R. MacCoun (1992) "Street Drug Markets in Inner-City Neighborhoods: Matching Policy to Reality" In D. Lyon, J. Steinberg and M. Vaiana (eds.) *Urban America* (pp. 227-251). RAND: Santa Monica, CA.

Rosenfeld, R., Fornango, R., and A. Rengifo (2007) "The Impact of Order-Maintenance Policing on New York City Robbery and Homicide Rates: 1988-2001." *Criminology.* 45: 355-383.

Sampson, R. J. (1987) "Urban Black Violence: The Effect of Male Joblessness and Family Disruption." *American Journal of Sociology.* 93, 348-382.

Sampson, R. J., and W.B. Groves (1989) "Community Structure and Crime: Testing Social Disorganization Theory." *American Journal of Sociology* 94: 744-802.

Sampson, R. J., Raudenbush, S.W., and Earls, F. (1997) "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science.* 277: 918-924.

Sampson, R. J., and Raudensbush, S. W. (1999) "Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods." *American Journal of Sociology.* 105: 603-651.

Sampson, R. J., Morenoff, J. D., and S. Raudenbush (2005) "Social Anatomy of Racial and Ethnic Disparities in Violence." *American Journal of Public Health*. 95: 224-232.

258 Schuerman, L. and S. Kobrin. (1986) "Communities and Careers in Crime." In Albert Reiss Jr. and Michael Tonry *Communities and Crime: Crime and Justice: A Review of Research* 8: pp. 67-100. Chicago: University of Chicago Press.

Shaw, C. R., and H. D. McKay (1942) *Juvenile delinquency and urban areas.* Chicago: University of Chicago Press.

Sherman, L. W., Gartin, P.R., and M. E. Buerger (1989) "Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place." *Criminology* 27: 27-55.

Sherman, L. W. (1990) "Police Crackdowns: Initial and Residual Deterrence" in Michael Tonry and Norval Morris, eds., *Crime and Justice: an Annual Review of Research.* Volume 12, pp. 1-48. Chicago: University of Chicago Press.

Sherman, L.W. (1992) "Attacking Crime: Police and Crime Control." in Norval Morris and Michael Tonry, eds., *Modern Policing: Crime and Justice*. Vol. 15, pp. 159-230. Chicago: University of Chicago Press.

Shihadeh, E. S., and Steffensmeier, D. J. (1994) "Economic Inequality, Family Disruption, and Urban Black Violence: Cities as Units of Stratification and Social Control." *Social Forces*. 73: 729-751.

Strogatz, S.H. (1996) *Nonlinear Dynamics and Chaos.* New York: Westview Press.

Taub, R, Taylor, G., and J. Dunham (1984) *Paths of neighborhood change: Race and crime in urban America.* Chicago: University of Chicago Press.

Taylor, R. (2001) *Breaking Away from Broken Windows: Baltimore Neighborhoods and the Nationwide Fight Against Crime, Grime, Fear, and Decline.* Boulder: Westview.

Taylor, R., and J. Covington (1990) "Ecological Change, Changes in Violence, and Risk Prediction." *Journal of interpersonal Violence* 5(2): 164-175.

Thrasher, F. (1927) *The Gang.* Chicago: University of Chicago Press.

Vold, G. B., and T. Bernard (1986). *Theoretical Criminology.* New York: Oxford University Press.

Wald, A., and J. Wolfowitz (1944) "Statistical Tests Based on Permutation of the Observations." *The Annals of Mathematical Statistics* 15(4): 358-372.

Weisburd, D., Bushway, S., Lum, C., and S-M. Yang (2004). "Crime Trajectories at Places: A Longitudinal Study of Street Segments in the City of Seattle." *Criminology.* 42(2), 283-322.

Weisburd, D., Wyckoff, L., Ready, J., Eck, J.E., Hinkle, J.C., and F. Gajewski (2006) "Does Crime Just Move Around the Corner?: A Controlled Study of Spatial Displacement and Diffusion of Crime Control Benefits." *Criminology.* 44(3): 549-591.

Wilcox, P., K.C., Land, and S.A. Hunt (2003). *Criminal Circumstance: A Dynamic Multicontextual Criminal Opportunity Theory.* New York: Aldine de Gruyter.

Wilson, W. J. (1987). *The Truly Disadvantaged.* Chicago: University of Chicago Press.

Wolfgang, Marvin E., and F, Ferracuti (1967) *The Subculture of Violence: Towards an Integrated Theory of Criminology.* London: Tavistock.

Wolfgang, M. E., Figlio, R. M., Tracy, P. E., and S.I. Singer (1985). *The National Survey of Crime Severity.* U.S. Department of Justice, U.S. Government Printing Office, Washington, D.C.