*Journal of Quantitative Criminology, Vol. 18, No. 3, September 2002 (© 2002)*

# A Note on the Use of County-Level UCR Data

**Michael D. Maltz[1,2] and Joseph Targonski[1]**

County-level crime data have major gaps, and the imputation schemes for filling in the gaps are inadequate and inconsistent. Such data were used in a recent study of guns and crime without considering the errors resulting from imputation. This note describes the errors and how they may have affected this study. Until improved methods of imputing county-level crime data are developed, tested, and implemented, they should not be used, especially in policy studies.

**KEY WORDS:** UCR, crime statistics, imputation, concealed weapons, guns, county-level crime data.

## 1. INTRODUCTION

County-level data are often intrinsically more appealing than state-level or national data. They allow for analysis of smaller units and can control for variation within states. Aggregating statistics over an entire state can mislead the true nature of a state's population characteristics; for example, although Nevada is considered a rural state, over 85% of its population lives in the two counties containing Reno and Las Vegas (Census, 2000).[3] It is for this reason that county-level analyses are considered so useful, and are used in displaying nationwide variations in crime (Zawitz, 1983, 1988), death due to disease (Pickle *et al.*, 1996), and voting behavior (New York Times, 2000).

However, users should be aware of potential pitfalls in using county-level crime data. A telling example of the problems one can encounter in using such data occurred in *More Guns, Less Crime*, Lott (1998, 2000; see also Lott and Mustard, 1997). These studies (referred to henceforth as MGLC) used county-level crime data to conclude that the imposition of

[1]University of Illinois at Chicago, Department of Criminal Justice (M/C 141), 1007 West Harrison Street, Chicago, IL 60607-7140.

[2]To whom correspondence should be addressed: Email: mikem@uic.edu

[3]Moreover, 99.7% of the population of Clark County lives in the Las Vegas Metropolitan Statistical Area (Goodall, Kafadar, and Tukey, 1998).



0748-4518/02/0900-0297/0 © 2002 Plenum Publishing Corporation

"right-to-carry" gun laws reduced the incidence of homicide and other violent crimes.

Because the findings run counter to the intuition of many, they have been subject to unusually close scrutiny (e.g., Ayres and Donohue, 1999; Black and Nagin, 1998; Zimring and Hawkins, 1998). The criticisms, for the most part, focus on the methods used and on the inferences drawn from the findings; they take the data as given. But the crime data are not "given"; in fact, due to problems in the reporting of crime data there are many gaps that need to be filled. The organization that prepared the data filled these gaps using two different estimation (i.e., imputation) procedures; moreover, the change in these procedures was not recognized by and incorporated into the MGLC analysis. Thus, there are so many problems with the county-level crime data sets used in MGLC that its analyses are called into question. We note, however, that the second edition of *More Guns, Less Crime* (Lott, 2000) includes state- and city-level analyses, which are not subject to this particular problem.

The county-level crime data used by the authors were the data sets aggregated to the county level and archived at the National Archive of Criminal Justice Data (NACJD), part of the Inter-University Consortium on Political and Social Research of the University of Michigan. They were based on the "Crime by County" data file provided by the FBI's Uniform Crime Reporting (UCR) Program. The NACJD data sets have a number of features that militate against their use in analyses of the type conducted in MGLC, which we discuss in this note.

We should note that the authors of these studies were not aware of the extent of the problems with the data they used. In fact, they went to great lengths to secure the best possible data and conferred with many people (including an author of this paper) prior to performing their analysis—but the problems were not known at that time. And not just by them: it is safe to say that few if any researchers were aware of the magnitude of the problems with county-level UCR data, including ourselves, prior to embarking on this analysis.[4] It is only recently that sufficient attention has been paid to the detailed characteristics of the UCR to bring its limitations to the fore. Moreover, they graciously permitted any and all critics access to their data.

In this note we first describe the relevant aspects of the FBI's Uniform Crime Reporting Program. We then provide the characteristics of the data used in MGLC and in the present study. The imputation procedures used by the FBI and by NACJD are then described, followed by how the imputed data were used in the MGLC analysis. Based on the problems introduced

[4]In fact, a paper coauthored by one of the authors of this paper (Olson and Maltz, 2001) assumes that the data are correct.

by the imputed data, we conclude that county-level crime data, as they are currently constituted, should not be used, especially in policy studies.

## 2. THE FBI'S UNIFORM CRIME REPORTING PROGRAM

The Uniform Crime Reporting Program was begun in 1930 (Maltz, 1977), in an effort to forestall newspapers from manufacturing "crime waves" in order to sell more papers. The UCR is essentially a voluntary program, in which reports are submitted by police agencies to the FBI (or, as has become more common, to state agencies that review and compile data from agencies in their states and forward them to the FBI). Because of this, the FBI has no control over the reliability, accuracy, consistency, timeliness, or completeness of the data they receive. The FBI is, of course, aware of the UCR's shortcomings and has made extensive efforts to assess and improve its quality.

Not all police agencies provide 12 months of crime data to the FBI: natural disasters, budgetary restrictions, personnel changes, inadequate training, and conversion to new computer or crime reporting systems all have affected the ability of police departments to report consistently, on time, completely, or at all. And some agencies may not fill out crime reports simply because they rarely have any crime to report.

To maintain consistency in their crime estimates in the face of these omissions, since the 1960s the FBI has been imputing crime data so that the year-to-year trends can be ascertained. Although it imputes crime data at the agency level, *the FBI does not publish these imputed figures for agencies or for counties.* They are used only to aggregate to the state, regional, and national level: they are not aggregated to the county level because they would not be accurate: a single non-reporting agency may account for a substantial fraction of a county's crime (see example below), but rarely do the non-reporting agencies in a state account for a major fraction of that state's crime.[5]

As mentioned above, the FBI also produces a "Crime by County" file, from which the NACJD county-level data file is generated. It includes, for each jurisdiction that has reported *any* crime data that year, the crime count for that jurisdiction, the months reported, and the jurisdiction's population. The crime count of agencies found in more than one county (e.g., Atlanta is in three counties) is distributed among the various counties according to the counties' share of the jurisdiction's population. It also includes the totals

[5]If there are missing or anomalous crime data from an agency with "a population over 100,000, personnel from the [FBI's] UCR Section call the agency directly to verify the data" Maltz, 1999: 16).

of crimes reported and populations for the agencies that submitted reports; however, it does not include the populations of those jurisdictions that submitted no reports, so the *listed* county total population may not be the *actual* county population (Maltz, 1999, p. 20). This is also true for the NACJD county-level data file.

The FBI publishes warnings about comparing cities and counties; they can be found in every issue of their annual report, *Crime in the United States*. The warning reads, "These rankings [of cities and counties based on their Crime Index figures] lead to simplistic and/or incomplete analyses which often create misleading perceptions" (FBI, 1998: iv). However, their rationale for not comparing cities and counties covers only "[g]eographic and demographic factors, ... [such as t]he transience of the population, its composition by age and gender, education levels, and prevalent family structures."

Such factors can often be accounted for by analyzing the data using sophisticated computer models. But these models cannot compensate for missing data of the type and magnitude encountered in the UCR. MGLC used county-level crime data from the FBI (as aggregated by NACJD) under the impression that the data were sufficiently accurate and complete for their analyses. They are not.[6]

## 3. DATA SOURCES AND CHARACTERISTICS

The county-level UCR data files from NACJD are available at http://www.icpsr.umich.edu/nacjd. The crime rates in MGLC combined the NACJD crime data (the numerators) with county-level population data (the denominators) from the Census Bureau.

In our analyses we used the original FBI-collected data, as organized and maintained by the National Consortium on Violence Research (NCOVR), located at Carnegie-Mellon University, rather than the NACJD database. The agency-level population data used by the FBI in that database were obtained from the Census Bureau.

The NCOVR data reside in an Oracle® database, accessible through the Internet (http://www.ncovr.heinz.cmu.edu), that permits off-site querying. Unfortunately, however, the NCOVR database begins in 1980, so our comparison with the MGLC data omits the first three years. It is more detailed than the NACJD county-level database; specifically, it contains data fields

[6]One reviewer of an earlier version of this article noted that "[t]he weakness of the ICPSR county-level data file is [obvious] to anyone who carefully reads the ICPSR codebook." Yet none of the critics of MGLC noted or mentioned it; apparently many users of the data do not read the ICPSR codebook carefully.

detailing the month-by-month reporting history of every policing jurisdiction. This permitted us to determine, for each reporting police agency, how many (and which) months in each year its crime counts were reported to the FBI. Thus we are able to estimate the extent to which data are fully reported to the FBI. Using the NCOVR data, we have estimated the extent of crime reporting gaps in every county for every year between 1980 and 1992.

### 3.1. Crime Data Characteristics

With this data source we were able to obtain about the same crime counts (numerators) as those included in the MGLC data set. The differences between our figures and MGLC are quite small; they were not exactly the same for two reasons. First, NACJD allocates crime levels from statewide agencies to each county in proportion to the county population; the consequence of this procedure will be discussed later.

Second, the FBI data set can change over time (the version used in publishing the report, *Crime in the United States*, for example, has an earlier deadline than the "Crime by County" data file), but not by very great amounts.

### 3.2. Population Data Characteristics

Similarly, the population data (denominators) contain some discrepancies, one affecting only a few counties, the other two affecting most counties. The first discrepancy is due to double-counting of agencies' populations; the second is due to varying estimates of agency populations; and the third is due to agencies with overlapping jurisdictions.

#### *3.2.1. Double-Counting*[7]

Care should be taken in using the original FBI data files, which are archived at NACJD. We found that doing so led us to double-count some agencies' populations, a problem that NACJD did not experience, since its source, the FBI's "Crime by County" data file, accounted for this factor.

Not all police agencies report their own data directly to the FBI; some report through other, usually larger, agencies located nearby. An example is found in Mineral County, Colorado, which contains two police agencies, the Mineral County Sheriff's Department and the Creede Police Department. In 1985, the Creede Police Department became "covered by" the Mineral County Sheriff's Department; that is, the crimes occurring in

[7]We thank Florenz Plassuman of SUNY-Binghamton for bringing this additional problem with FBI population data to our attention.

Creede were reported through and consolidated with those reported by the Mineral County Sheriff.

However, this change was not fully noted in the data set containing the raw UCR data. The *crime data* were reported accurately but the *population data* were not. As seen in Fig. 1, from 1985 onward, the Creede population was double-counted. In its compilation of county-level crime statistics, NACJD properly accounted for these "covered-bys," so this error is not found in the NACJD population counts. Nor would it be found in the MGLC data, since Census data were used in those studies.

For the most part this problem occurs in sparsely populated counties; however, it is also found in the UCR population data for Nassau and Suffolk Counties, New York, for the entire 1980–92 period. That is, if we add up the agency populations for these counties as included in the FBI data file, they exceed the Census (i.e., MGLC) populations by considerable amounts (Fig. 2; 1985 data are given in Table I). The excess populations are approximately equal to the combined populations of the agencies that are "covered by" the county-level agencies. Our initial investigation of this problem has turned up very few cases, with these cases the only ones in a large county.

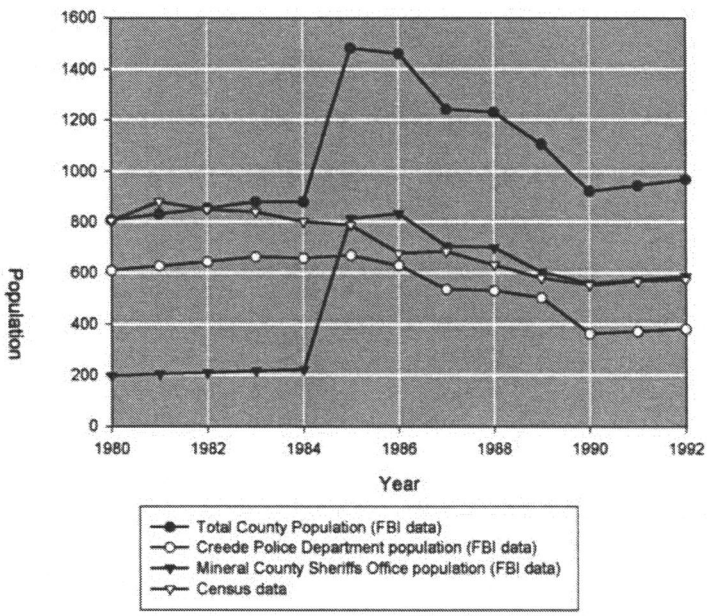


**Fig. 1.** FBI-listed population for Mineral County, CO, ORIs.

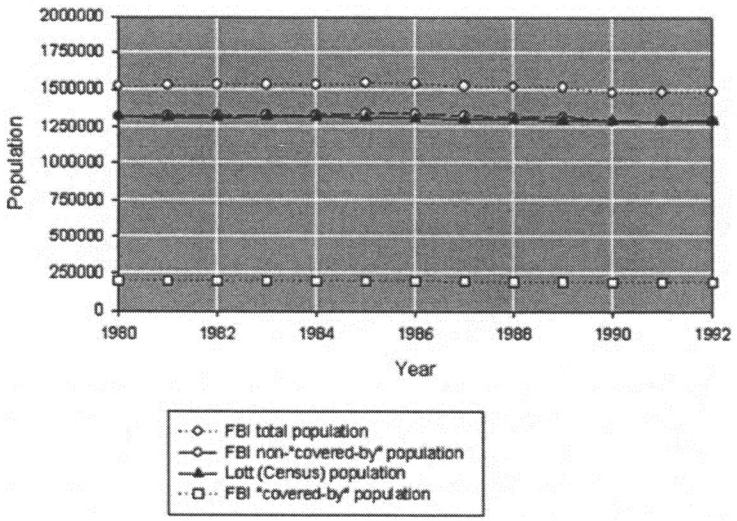


**Fig. 2a.** Nassau County, NY, populations.

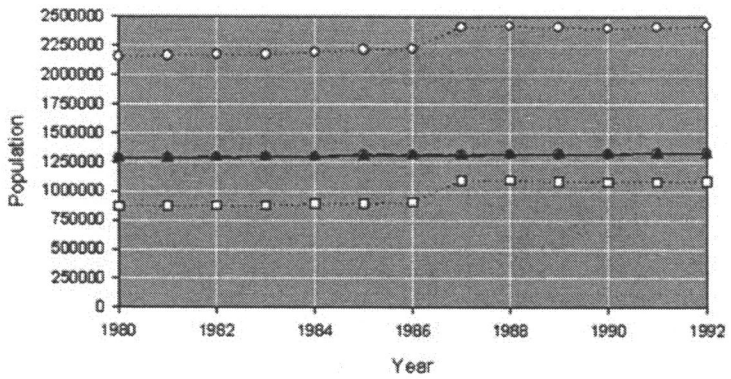


**Fig. 2b.** Suffolk County, NY, populations.

**Table I.** 1985 Population Data for Nassau and Suffolk Counties

| | FBI (total of all agencies in county) | FBI (total of "covered-by" agencies) | FBI (difference) | Census Data |
|---|---|---|---|---|
| Nassau County | 1,543,992 | 202,208 | 1,341,784 | 1,314,652 |
| Suffolk County | 2,216,536 | 897,814 | 1,318,722 | 1,306,894 |

### *3.2.2. Varying Estimates*

Summing the Census population estimates for each agency in a county (used by NACJD) will not necessarily equal the Census county population, for a variety of reasons. Some agencies may be in more than one county;[8] the estimates may have been made at different times; and the procedure used by the Census Bureau for estimating intercensal populations may be different for counties than for municipalities.

### *3.2.3. "Zero-Population" Agencies*

Not all police agencies can be dealt with in a straightforward manner. Some jurisdictions in a county provide crime reports without being recorded in the UCR as having a population. This is the case when policing jurisdictions overlap. For example, there are actually four policing jurisdictions in Delaware County, Indiana: the Muncie Police Department, the Delaware County Sheriff's Office, the Indiana State Police barracks in Delaware County, and the Ball State University campus police. The latter two jurisdictions are called "zero-population" jurisdictions because the FBI assigns no population to them.

That is, the Indiana State Police barracks in Delaware County is a "zero-population" agency because its coverage is the entire county, but the total county population is already included in Muncie and the area policed by the Delaware County Sheriff. In addition, those policed by the Ball State University campus police are already counted in the Muncie population. So, to avoid double-counting, their crime reports are not associated with a population. Crime reports provided by these agencies are added to the county's numerator, but no population is added to the denominator.[9]

During the entire period in question the Indiana State Police barracks in Delaware County provided complete crime data. The Ball State University police only recently began reporting crime data to the FBI, in 1990, perhaps in response to federal legislation regarding the reporting of crime by colleges and universities—see the first section of Table II.

[8]From the Data Collection Description of the 1994 County-Level Detailed Arrest and Offense Data (ICPSR dataset 6669): "In the county-level crimes reported files, the population and crime data for jurisdictions located in multiple counties are provided by the UCR proportioned to each county (maximum of three) in which the jurisdiction is located."

[9]If a "zero-population" police agency does not report fully, no imputation is made to compensate for the missing data, either by the FBI or by NACJD. With the assistance of the American Statistical Association, the FBI is currently considering new procedures for handling this problem (ASA, 2000).

**Table II.** UCR Data for Delaware County, Indiana, 1980–1992

| Year | Delaware Cty. Sheriff's PD | Muncie PD | Ball State Univ. PD | Indiana St. Pol. (Del. Cty.) | NACJD Weighted Crime |
|---|---|---|---|---|---|
| *Months Reported by Police Agencies, Delaware County, IN* | | | | | |
| 1980 | 0 | 12 | | 12 | |
| 1981 | 0 | 12 | | 12 | |
| 1982 | 12 | 12 | | 12 | |
| 1983 | 10 | 12 | | 12 | |
| 1984 | 5 | 12 | | 12 | |
| 1985 | 0 | 12 | | 12 | |
| 1986 | 0 | 12 | | 12 | |
| 1987 | 6 | 12 | | 12 | |
| 1988 | 6 | 12 | | 12 | |
| 1989 | 0 | 12 | | 12 | |
| 1990 | 0 | 5 | 12 | 12 | |
| 1991 | 0 | 0 | 12 | 12 | |
| 1992 | 0 | 0 | 12 | 12 | |
| *Violent Crime* | | | | | |
| 1980 | 0 | 903 | | 18 | 921 |
| 1981 | 0 | 775 | | 9 | 784 |
| 1982 | 20 | 441 | | 6 | 467 |
| 1983 | 29 | 314 | | 15 | 364 |
| 1984 | 10 | 217 | | 14 | 231 |
| 1985 | 0 | 219 | | 16 | 235 |
| 1986 | 0 | 185 | | 16 | 201 |
| 1987 | 48 | 183 | | 12 | 291 |
| 1988 | 44 | 208 | | 13 | 309 |
| 1989 | 0 | 203 | | 9 | 212 |
| 1990 | 0 | 112 | 51 | 11 | 61 |
| 1991 | 0 | 0 | 33 | 8 | 42 |
| 1992 | 0 | 0 | 28 | 6 | 34 |
| *Property Crime* | | | | | |
| 1980 | 0 | 5018 | | 114 | 5132 |
| 1981 | 0 | 4602 | | 106 | 4708 |
| 1982 | 661 | 4589 | | 83 | 5333 |
| 1983 | 569 | 4682 | | 96 | 5461 |
| 1984 | 213 | 4315 | | 73 | 4388 |
| 1985 | 0 | 3962 | | 83 | 4045 |
| 1986 | 0 | 3665 | | 98 | 3763 |
| 1987 | 238 | 3473 | | 70 | 4019 |
| 1988 | 243 | 3749 | | 65 | 4300 |
| 1989 | 0 | 3889 | | 59 | 3948 |
| 1990 | 0 | 1544 | 952 | 43 | 995 |
| 1991 | 0 | 0 | 994 | 29 | 1023 |
| 1992 | 0 | 0 | 941 | 45 | 986 |

### 3.3. Crime Rates

In its county-level data files, NACJD provides estimates of county crime counts and the population to which these crimes are attributed. Because of the way NACJD imputed crime data between 1977 and 1993, sometimes the crime count is nonzero but the population is zero. That is, before 1994 *NACJD used an imputation procedure that omitted substantial numbers of crimes and populations from a county's estimated crime count.* Moreover, *MGLC omitted the same crimes as NACJD, but not the populations.* The following section details the imputation procedures used by the FBI, the procedures used by NACJD, and how MGLC used the NACJD-imputed data.

## 4. IMPUTING UCR DATA

### 4.1. The FBI's Imputation Procedure

When gaps appear in the crime data, the FBI fills in the gaps with estimates; this permits them to provide comparisons from year to year. The FBI imputes crime data for agencies in the following way. If an agency reports three or more months of crime data, the total number of crimes for the year is estimated to be equal to $C \times 12/N$, where C is the total crime reported and N is the number of months reported. Thus, an agency that only reports four months of data, and has experienced 10 robberies during that period, will be estimated to have experienced $10 \times 12/4$, or 30 robberies over the entire year.

If an agency reports 0, 1, or 2 months of crime data, its crime count is replaced by a number equivalent to the rate experienced by the aggregate crime rate of "similar" agencies (Maltz, 1999: 23). A "similar" agency is one that is in the same state and same population group as defined by the FBI (Maltz, 1999, p. 21). The first line of Table III summarizes this procedure.

Thus, below a certain threshold an agency's data is replaced by equivalent data from "similar" agencies; and above that threshold, the agency's number of crimes (inflated to reflect a full-year estimate) is used for the crime count. This is obviously a crude estimate, since it does not take into account seasonality or other specifics of the situation, but the FBI feels that this estimation procedure is adequate for the limited purpose of making state (and regional and national) estimates of crime. Moreover, *every jurisdiction (and population) is included in the aggregate statistics.*

However, no imputation is done for "zero-population" or statewide police agencies; if they are missing some (or all) monthly reports, no provision is made for inflating them to a full-year equivalent.

**Table III.** Description of the FBI and NACJD Imputation Procedures

| Imputation procedure | Use in MGLC | NACJD's calculation of an agency's imputed crime: Number of months reported, N: 0 to 2 | 3 to 5 | 6 to 12 |
|---|---|---|---|---|
| FBI (used by NACJD after 1993) | 1994–1996 | $C_S * P_A/P_S$ | $C_A * 12/N$ | $C_A * 12/N$ |
| Used by NACJD from 1977–1993 | 1977–1993 | 0 | 0 | $C_A * 12/N$ |

$C_A$, $P_A$: The agency's crime count and population, respectively, for the year in question.
$C_S$, $P_S$: The crime and population counts, respectively, of similar agencies in that state, for the year in question. Similar agencies, as currently used by the FBI, are those in the same state and FBI-defined population group. Population groups are cities according to population (Group I—250,000 and over; II—100,000–249,999; III—50,000–99,999; IV—25,000–49,999; V—10,000–24,999; VI—less than 10,000) and counties (VIII—rural county; IX—suburban county). Group VII consists of cities with populations under 2,500 and colleges and universities, to which no population is attributed; they are now included in Group VI.

## 4.2. The NACJD Imputation Procedure

The NACJD imputation procedure is very different from the FBI's, because it initially had a very different purpose. In support of a BJS desire to provide the country with a cross-sectional "snapshot" of crime in the U.S., NACJD aggregated the 1980 jurisdiction-level crime data to the county level. The report (Zawitz, 1983) contained a choropleth map of the U.S.; i.e., each county was shaded according to its crime rate. Jurisdictions that reported fewer than six months were excluded from consideration, and counties that had less than half the population represented in the data were considered to have missing data; they were represented in white in the map. The crime rates in the other counties were represented by one of five different shadings, representing five different levels of crime. That is, there was no attempt (or need) to represent the crime rate with any degree of precision. The second line of Table III summarizes this procedure.

The report was very well received, and as a consequence NACJD was asked by BJS to repeat this effort with 1984 data (Zawitz, 1988). Moreover, BJS commissioned NACJD to generate county-level data sets routinely, starting with the 1977 UCR data, and make them available for others to use. "The deficiencies or consequences of using the *ad hoc* imputation procedure were not considered, because up to that time the county-level data had only been used for cross-sectional comparisons and not for more rigorous analytic purposes" (Maltz, 1999, p 11).

A description of the imputation procedure was provided by NACJD so others would be mindful of its limitations. In the NACJD County Crime Codebook covering the years 1977–83, the following description of the imputation procedure is given (see http://www.icpsr.umich.edu/cgi/ab.prl?file = 8703:

> Data have been aggregated to the county level. Within each county, data for agencies reporting 6 to 11 months of information were weighted to yield 12-month equivalents. Agencies reporting fewer than 6 months of data were excluded from the aggregation. Data from agencies reporting only statewide figures were allocated to counties in proportion to each county's share of the state population (which totaled the populations of those county agencies reporting six months or more of data).

The same paragraph is given in the abstract for datasets through 1993. That procedure was employed in the 1977–93 county-level UCR data sets by NACJD in estimating county crime: the data for agencies reporting six or more months are inflated as in the FBI's procedure, by $C \times 12/N$. However, *both the crime and population counts of agencies reporting five or fewer months are dropped completely* in calculating the county's total crime and population counts; see Table III. Thus, a crime rate calculation (crimes per person) using the NACJD data makes the implicit assumption that the excluded agencies have the same crime rate as the rest of the county.

After 1993 NACJD used the same imputation procedure as the FBI in its county-level computations.

### *4.2.1. "Zero-Population" and Statewide Agencies*

Similar to the FBI, NACJD does not impute data for "zero-population" agencies. In addition, however, the way NACJD handles agencies that provide statewide crime data can have major effects on county crime rates in some states. Ordinarily these agencies (like the state police) do not report a great deal of crime. NACJD distributes the crime counts of statewide agencies to counties in proportion to their population.[10] While this would not make a great difference in urban counties that have a great deal of crime, it would tend to greatly undercount the crime count in rural counties with little crime: not only does this procedure spread the crime throughout the state, but it is often the case that *the state police perform crime-related functions only in rural counties.*

For example, the Connecticut State Police have major criminal justice responsibilities in the four Connecticut counties east of the Connecticut River—New London, Middlesex, Tolland, and Windham—and Litchfield County in the northwest corner of the state. They have virtually no criminal justice responsibilities in the three more highly urbanized counties—Fairfield, Hartford, and New Haven—which contain about three-quarters of

[10]Not all statewide agencies provide state-level data. For example, the Indiana State Police provides county-level crime data. When the data are available by county, they are allocated properly.

the state's residents.[11] The NACJD allocation procedure therefore allocates three-quarters of the Connecticut State Police-reported crime to the counties where they do little or no non-traffic policing, rather than to the counties that generated the crime.

### 4.3. Applying the NACJD Data in MGLC

The first edition of MGLC analyzed data from 1977–92. It relied on NACJD crime data as imputed, but not their population figures. Instead it used the Census Bureau's county-level population figures. This procedure, in effect, assumes that the missing agencies have no crime. In other words, *when a major police agency in a county did not report completely, the crime rate calculated in MGLC is substantially lower than the actual crime rate.* This is equivalent to basing the Cook County, Illinois, crime rate on the aggregated crime data from all Cook County jurisdictions except Chicago, divided by the entire county population. An example of its effect is given in Section 5, below.

The second edition of MGLC (Lott, 2000) extended the county-level analysis to include 1993–96. This was probably done without checking the codebooks for the later data, under the assumption that the methods used by NACJD to aggregate the data did not change. This assumption was incorrect, since (see Table III) NACJD has been using a different imputation procedure since 1994. The 1994 NACJD codebook (ICPSR dataset 6669) explicitly notes this in a major heading, "Break in Series," and describes the new imputation procedure it will use from 1994 on. It goes on to state,

> These changes will result in a break in series from previous UCR county-level files. Consequently *data from earlier year files should not be compared to data from* 1994 *and subsequent years* because changes in procedures used to adjust for incomplete reporting at the ORI[12] or jurisdiction level may be expected to have an impact on aggregates for counties in which some ORIs have not reported for all 12 months. [Emphasis added.]

We did not investigate the implications of the second error on the MGLC analysis, and confine our remarks to the analysis from the first edition.

[11]Personal communication, William G. Lopez, Connecticut Uniform Crime Reporting Program, September 5, 2000.

[12]An ORI is an *O*Riginating agency *I*dentifier, the identification number used by the FBI to identify police agencies.

## 5. AN EXAMPLE

Figure 3 shows the number of months reported between 1980 and 1992, for the two major police agencies in Delaware County, Indiana: the Muncie Police Department and the Delaware County Police Department. Note that in some years one of these agencies did not provide at least six months of reports, in other years the other agency did not, and in 1991 and 1992 neither agency reported at all.

The Delaware County population figures from Census and NCOVR are shown in Fig. 4a; they are contrasted with the NACJD figures, based on the months of UCR data they reported, as per the NACJD imputation procedure: if fewer than six months of crime data are reported NACJD drops that agency's population from the denominator.

Figure 4b shows the violent and property crime rates for this county in the Lott data set; as can be seen, the apparent drop in crime in 1990–92 is not a real drop in crime (nor is the earlier dip) but reflects the fact that neither of the major police agencies in Delaware County had adequate crime reporting in those years. [The reason that the crime rates are not zero in 1990–92 is that the two "zero-population" police agencies (the Indiana State Police barracks in Delaware County and the Ball State University Police Department) reported crime in those years.] The crime rate calculations for Delaware County are given in Table IV.

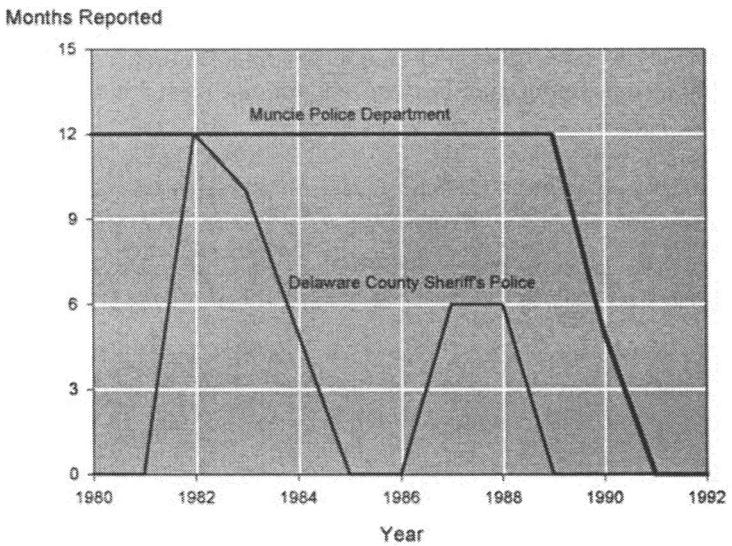


**Fig. 3.** Months of crime data reported by police agencies in Delaware County, IN, 1980–1992.

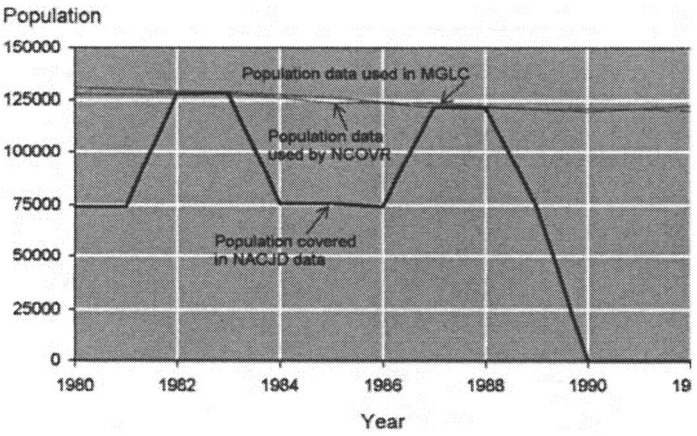


**Fig. 4a.** Population covered in NACJD data Delaware County, IN, 1980–1992.

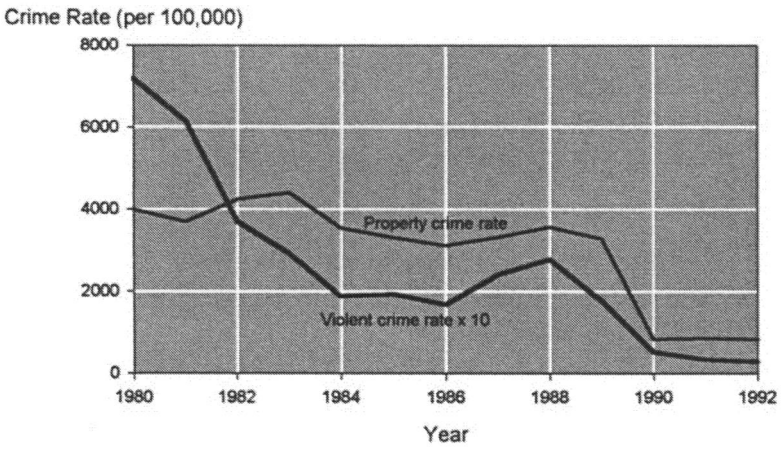


**Fig. 4b.** Crime rates used in MGLC.

This example is not atypical; Fig. 5 portrays the gaps in crime reporting coverage of Georgia's counties.[13] The helter-skelter nature of the figure shows clearly that there is no underlying pattern to the non-reporting behavior. Although many of Georgia's counties have very little unreported crime, this is not generally the case.

[13]Some 20 states have similar reporting profiles.

**Table IV.** UCR Population and Crime Rates for Delaware County, Indiana, 1980–1992

*Population Covered*

| Year | Delaware Cty. Sheriff's PD | Muncie PD | Ball State Univ. PD | Indiana St. Pol. (Del. County) |
|---|---|---|---|---|
| 1980 | 53,359 | 74,096 | | 0 |
| 1981 | 53,407 | 74,162 | | 0 |
| 1982 | 53,456 | 74,230 | | 0 |
| 1983 | 53,534 | 74,338 | | 0 |
| 1984 | 50,343 | 75,655 | | 0 |
| 1985 | 47,948 | 75,661 | | 0 |
| 1986 | 49,441 | 74,272 | | 0 |
| 1987 | 48,541 | 72,961 | | 0 |
| 1988 | 48,519 | 72,929 | | 0 |
| 1989 | 47,083 | 73,795 | | 0 |
| 1990 | 48,624 | 71,035 | 0 | 0 |
| 1991 | 49,201 | 71,878 | 0 | 0 |
| 1992 | 49,657 | 72,544 | 0 | 0 |

*Population of Delaware County, IN*

| Year | Sum of agency populations | MGLC (Census) population | NACJD (weighted) population |
|---|---|---|---|
| 1980 | 127,455 | 128,448 | 74,096 |
| 1981 | 127,569 | 127,328 | 74,162 |
| 1982 | 127,686 | 125,888 | 127,686 |
| 1983 | 127,872 | 124,064 | 127,872 |
| 1984 | 125,998 | 123,568 | 75,655 |
| 1985 | 123,609 | 122,032 | 75,661 |
| 1986 | 123,713 | 120,768 | 74,272 |
| 1987 | 121,502 | 121,072 | 121,502 |
| 1988 | 121,448 | 120,656 | 121,448 |
| 1989 | 120,878 | 119,952 | 73,795 |
| 1990 | 119,659 | 119,872 | 0 |
| 1991 | 121,079 | 119,808 | 0 |
| 1992 | 122,201 | 119,776 | 0 |

*Crime Rate Calculations*

| Year | Violent crime rate* using NACJD crime and population estimates | Property crime rate* using NACJD crime and population estimates | MGLC violent crime rate* | MGLC property crime rate* |
|---|---|---|---|---|
| 1980 | 1243.0 | 6926.1 | 717.0 | 3995.4 |
| 1981 | 1057.1 | 6348.3 | 615.7 | 3697.5 |
| 1982 | 365.7 | 4176.7 | 371.0 | 4236.3 |
| 1983 | 284.5 | 4270.5 | 293.2 | 4401.6 |
| 1984 | 305.3 | 5800.0 | 186.9 | 3551.1 |
| 1985 | 310.6 | 5346.2 | 192.6 | 3314.7 |
| 1986 | 270.6 | 5066.5 | 166.4 | 3115.9 |
| 1987 | 239.5 | 3307.8 | 240.4 | 3319.5 |
| 1988 | 254.4 | 3540.6 | 256.1 | 3563.9 |
| 1989 | 287.3 | 5350.0 | 176.7 | 3291.3 |
| 1990 | Indeterminate | Indeterminate | 51.7 | 830.1 |
| 1991 | Indeterminate | Indeterminate | 34.2 | 853.9 |
| 1992 | Indeterminate | Indeterminate | 28.4 | 823.2 |

*Crime rates are calculated per 100,000 people.

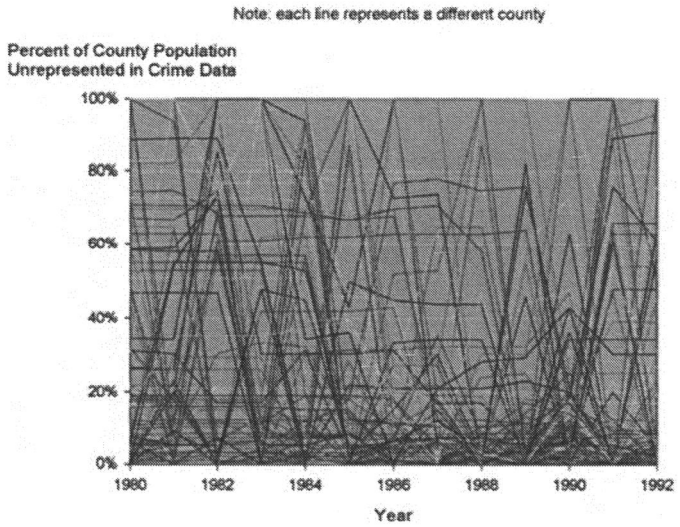


**Fig. 5.** Percent of county population unrepresented in crime data, 159 Georgia counties, 1980–1992.

It should be noted that the smaller counties are more likely to have extensive reporting deficiencies than larger counties. As in MGLC, we deal with all counties. In fact, Lott (2000, pp. 143, 155) criticizes Black and Nagin (1998) for focusing on counties with populations of 100,000 or more, in their critique of his analysis, noting that this procedure eliminates 85% of all counties from the analysis. Including the smaller counties, however, increases the number of error-laden data points: whereas 5% of data points from counties over 100,000 in population are in error by 30% or more, this is true of 14% of the data points from counties under 100,000.

## 6. COUNTY-LEVEL CRIME REPORTING

A broader view of the extent of gaps in county-level crime reporting coverage can be seen in Fig. 6. Note that over 50% of the county-level data points in three states (CT, IN, and MS) are missing crime data from more than 30% of their populations, and in another 13 states more than 20% of the data points contain coverage gaps of this magnitude.[14] Moreover, those states with the greatest gaps in crime reporting coverage tend to be states

[14]We chose 30% as a convenient benchmark; setting it at other levels gives similar results.

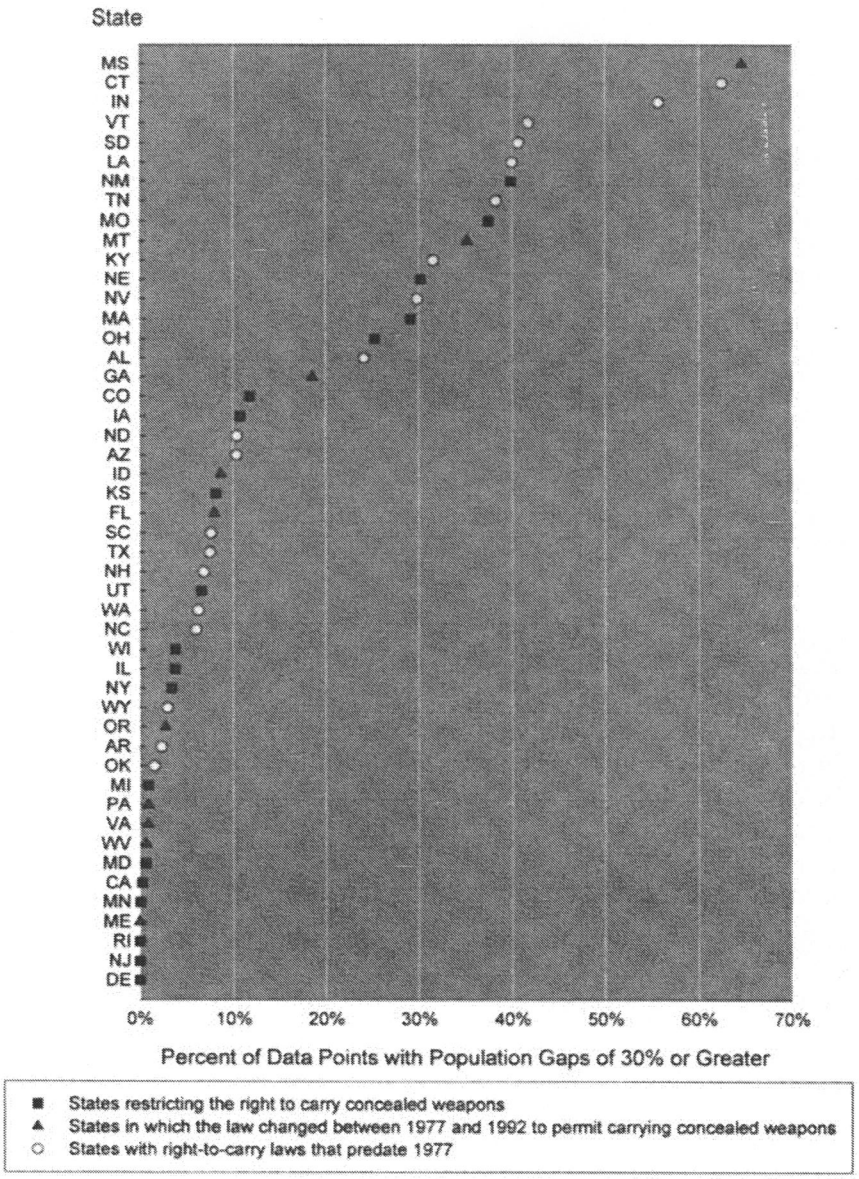


**Fig. 6.** Coverage gaps in county-level state data.

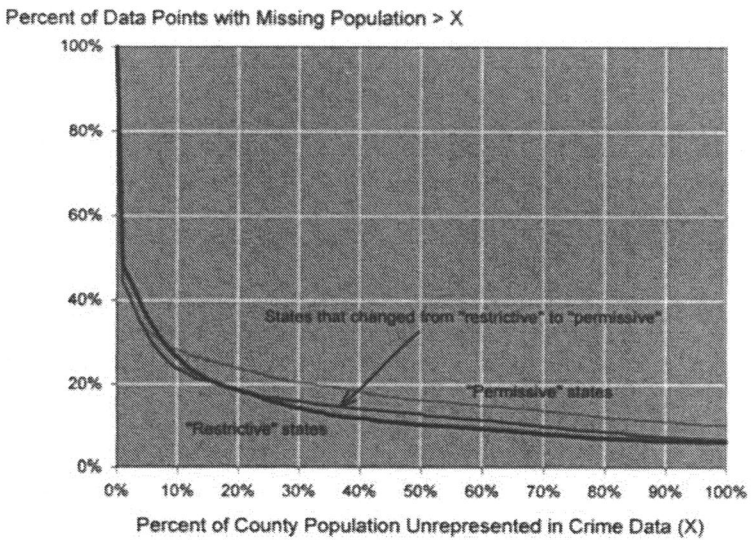


**Fig. 7.** Percent of county population unrepresented in crime data, by type of state "shall-issue" law.

that have (or have changed to) laws permitting the carrying of concealed weapons.[15]

The effect of these gaps on the MGLC data set can be ascertained from Fig. 7. It shows the extent of coverage gaps for counties in three types of states: those that have permitted carrying concealed weapons since before 1977, those that relaxed their restrictions between 1977 and 1992, and those that have restricted the carrying of concealed weapons since before 1977. As can be seen, the "permissive" states have more gaps in coverage; note that these states have a greater fraction of data points with 30% or more coverage gaps than the other states—16% vs. 10%.

The extent of crime reporting gaps varies considerably among the individual states. Figure 8 shows the population coverage gaps of California, Georgia, and Indiana. It shows that in Georgia, 9% of its data points represent less than half of the county population—for Indiana, this is true of 45% of the data points. What this analysis clearly indicates is that, at the county level, for some states the underrepresentation of population in the crime statistics is very extensive.

[15]In this and subsequent figures, darker symbols or lines denote more restrictive conditions. We also note that our classification of states is open to challenge; however, since we are comparing our analysis to that of MGLC, we used the same classification as in MGLC.

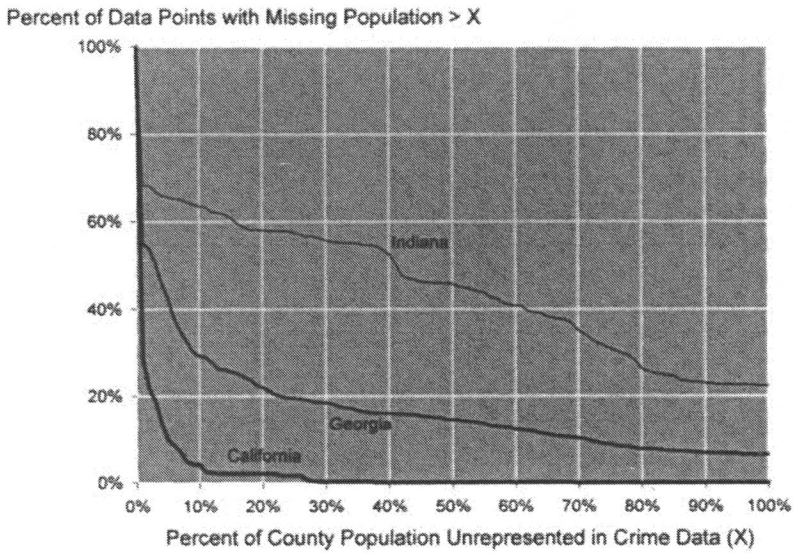


**Fig. 8.** Percent of county population unrepresented in crime data, in California, Georgia, and Indiana, 1980–1992.

## 7. CONCLUSION

The inadequacy of some jurisdictions' reporting practices would not be a problem if the jurisdictions were "consistently inconsistent." That is, if a particular municipality or sheriff's office in a county did not report every year, then the *crime rate* might be incorrect but the year-to-year *variation in crime rate* would arguably not be affected, a sort of "law of continuity of error." This would also hold if a jurisdiction in a county reported crime data for the first six months of every year but not the last six months, year after year. This, however, is not the case: many states display the same variation in county-level reporting as seen in Fig. 5, with different counties taking turns reporting erratically.

At this point, county-level crime data cannot be used with any degree of confidence. The only possible remedy, we feel, is to develop and test new algorithms for estimating crime counts, and to provide estimates of the extent of possible error in each county's data. Under a grant from the National Institute of Justice, we are currently proceeding on this course of action.

The crime rates of a great many counties have been underestimated, due to the exclusion of large fractions of their populations from contributing

to the crime counts. Moreover, counties in those states with the most coverage gaps have laws permitting the carrying of concealed weapons. How these shortcomings can be compensated for is still an open question, one that we are attempting to answer in our ongoing study of different methods of imputation. It is clear, however, that in their current condition, county-level UCR crime statistics cannot be used for evaluating the effects of changes in policy.

In addition, *all* studies that use aggregated UCR data—especially at the county level—should be looked at carefully to determine the extent to which coverage gaps and imputation affect their findings. The FBI is currently developing new methods to improve the old imputation procedures; given sufficient testing, they should enhance our ability to estimate crime and the effect of different policies on its prevalence.

## ACKNOWLEDGMENTS

This work was supported by a Visiting Fellowship to the first author from the Bureau of Justice Statistics, U.S. Department of Justice and by Grant No. 2001-IJ-CX-0006 from the National Institute of Justice. The authors thank Kenneth Candell, John Lott, Marianne Zawitz, and the reviewers for comments on earlier drafts, and Pat Edgar and the NCOVR staff for their assistance with data issues. The opinions expressed herein, however, are our own and do not reflect the opinions or policies of these individuals or of the U.S. Department of Justice.

## REFERENCES

American Statistical Association (2000). Meeting of the UCR Subcommittee of the ASA Committee on Law and Justice, held at the offices of the ASA, August 9, 2000.

Ayres, I., and Donohue, J. J., III (1999). Nondiscretionary concealed weapons laws: a case study of statistics, standards of proof, and public policy. *Am. Law and Econ. Rev.* 6: 436–470.

Black, D., and Nagin, D. (1998). Do "Right-to-Carry" laws deter violent crime? *J. Legal Studies* 27: 209–219.

Census Bureau (2000). County data are from http://www.census.gov/population/cen2000/phc-t4/tab01.pdf; the total Nevada population is found at http://www.census.gov/population/cen2000/phc-t2/tab01.pdf.

Federal Bureau of Investigation (1998). *Crime in the United States.* FBI, Washington, DC.

Goodall, C. R., Kafadar, K., and Tukey, J. W. (1998). Computing and using rural versus urban measures in statistical applications. *Am. Statistician*, 52: 101–111.

Lott, J. R., Jr. (1998). *More Guns, Less Crime*, University of Chicago Press, Chicago.

Lott, J. R., Jr. (2000). *More Guns, Less Crime*, second edition. University of Chicago Press, Chicago.

Lott, J. R., Jr., and Mustard, D. B. (1997). Crime, deterrence, and right-to-carry concealed handguns. *J. Legal Studies* 26: 1–68.

Maltz, M. D. (1977). Crime statistics: a historical perspective. *Crime and Delinq.* 23: 32–40.

Maltz, M. D. (1999). *Bridging Gaps in Police Crime Data*, a discussion paper from the BJS Fellows Program. Report No. NCJ-1176365, Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice, Washington, DC. Available at http://www.ojp.usdoj.gov/bjs/pub/pdf/bgpcd.pdf.

New York Times (2000). "The Presidential Vote." Thursday, November 9, 2000, Page B1.

Olson, D. E., and Maltz, M. D. (2001). Right-to-carry concealed weapon laws and homicide in large U.S. counties: The effect on weapon types, victim characteristics, and victim–offender relationships. *J. Law and Econ.*, in press.

Pickle L. W., Mungiole M., Jones G. K., and White, A. A. (1996). *Atlas of United States Mortality*, National Center for Health Statistics, Hyattsville, Maryland.

Zawitz, M. W., Ed. (1983). *Report to the Nation on Crime and Justice: The Data*, BJS Report NCJ 87068, U.S. Department of Justice, Washington, DC.

Zawitz, M. W., Ed. (1988). *Report to the Nation on Crime and Justice: Second Edition*, BJS Report NCJ 105506, U.S. Department of Justice, Washington, DC.

Zimring, F., and Hawkins, G. (1997). Concealed handgun permits: the case of the counterfeit deterrent. *Responsive Community* 7: 46–60.