INTERNATIONAL JOURNAL OF COMPARATIVE AND APPLIED CRIMINAL JUSTICE
WINTER 1988, VOL. 12, NO. 2

Alternatives To Population-Based Crime Rates

BRIAN STIPAK
Portland State University

This paper examines several alternatives to standard, population-based crime rates. These alternatives include rates standardized according to a variable other than population, and rates calculated by statistically standardizing for a number of variables simultaneously. Augmenting traditional crime rates with alternative types of rates can provide more information about the prevalence of crime, can help avoid the limitations of traditional rates, and can enhance analysts' flexibility in choosing crime rates appropriate to their analytical purpose.

CRIME RATES are almost always expressed as population-based rates—for example, the number of crimes per thousand population. Despite the widespread use of population-based rates, there are alternative types of crime rates. Some criminologists have even argued the superiority of such alternatives. The purpose of this paper is to investigate the use of alternatives to population-based crime rates.

Past Research

The existing research literature contains a few applications of non-population-based crime rates. Half a century ago Lottier (1938) calculated state automobile theft rates based on automobile registration. A hiatus of almost thirty years followed Lottier's innovation until Boggs (1965) related occurrence of crime to environmental opportunity. Boggs (1965: 900) argued that a crime rate "should form a probability statement, and therefore should be based on the risk or target group appropriate for each specific crime category." Boggs used a variety of crime-specific denominators for standardization, including number of resident females as a denominator for forcible rape, and number of occupied housing units as a denominator for residential burglary. Similarly, Frisbie *et al.* (1977) computed residential burglary rates based on the number of residential units, commercial burglary rates based on the number of commercial units, automobile theft rates based on registered vehicles, and vandalism rates based on the number of buildings. Cohen *et al.* (1985) computed burglary rates based on the number of households, and automobile theft rates based on the number of registered automobiles.

The research literature has also pointed out some of the problems that result from standardizing by population. Harries (1974: 4) cited a simple example from before 1958, when yearly rates were computed using decennial census population figures. Rates consequently tended to rise between decennial censuses, and fall immediately following a census. The official robbery rate in California, for example, fell from 136.1 in 1949 to 85.6 in 1951. The use of decennial population totals also introduced relative distortions across jurisdictions experiencing unequal growth. This simple problem has already, of course, been recognized and rectified by using yearly population estimates.

As an example of a less obvious distortion produced by population-based rates, Boggs (1965) pointed out that central business districts often have spuriously high crime rates. Although such areas offer many opportunities for crime during the day, they contain a small number of residents. Boggs (1965: 900) argued that:

> Although many crimes do take place in such areas, valid occurrence rates would be low relative to the number of potential targets or environmental opportunities for crime.

In a similar vein, Harries (1981: 148) commented that "the uncritical application of population as a denominator for all crime categories may yield patterns that are at best misleading and at worst bizarre."

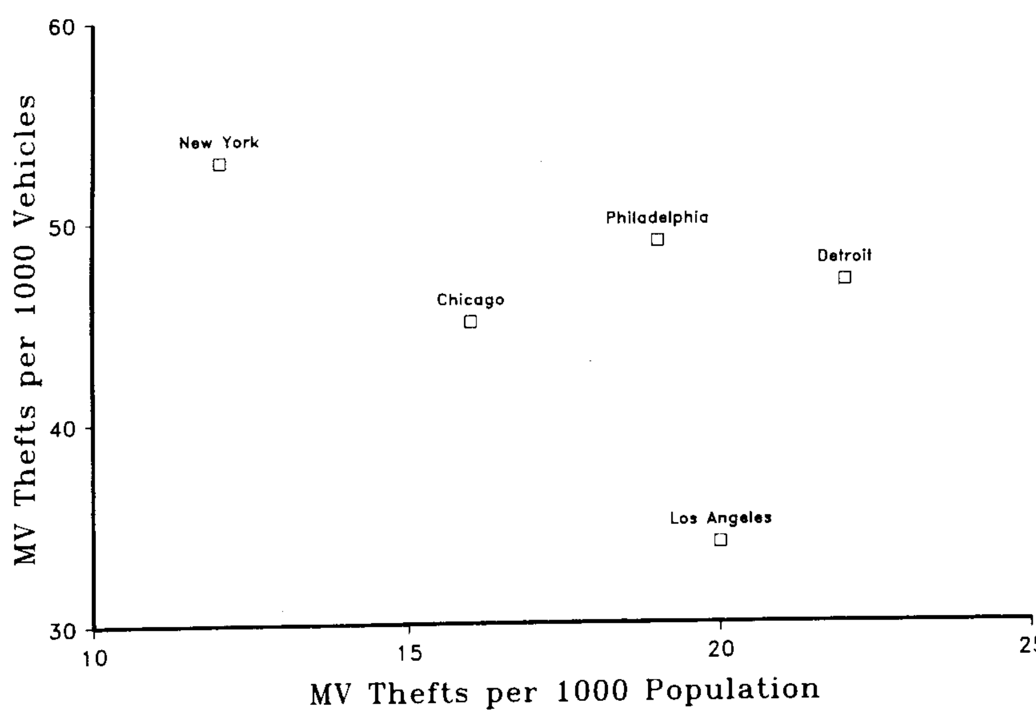
Other researchers have also criticized the utility of population-based rates. Skogan (1978: 4) argued that per capita rates typically do not provide useful measures of the risk of victimization. Boydell (1969: 33) commented that cities have different sleeping, working, and evening populations, and that census population figures are appropriate only for the sleeping population. Similarly, Gibbs and Erickson (1976: 606) pointed out the inappropriateness of census population counts as indicators of the number of potential victims or offenders, and Harries (1974: 5) lamented the inadequacy of standard crime rates as measures of victimization risk or of criminal propensities. Most generally, Sparks (1981: 54-55) argued that the purposes for which crime rates are used require that the rates reflect criminal tendencies in society. Since variations in the opportunities for criminal behavior can affect rates even when the underlying criminal tendencies remain constant, Sparks further argued the need to incorporate that variation in rate calculations.

These criticisms implicitly raise a question that usually remains ignored—what type of variable is appropriate to use as a standardizing variable? Unfortunately, as Brantingham and Brantingham (1984: 55) have commented, there exists no widespread agreement on the appropriate alternative bases for computing rates. One general approach is to standardize on the basis of a measure of the prevalence of potential criminal opportunities. Skogan (1978: 4) argued that this approach yields the most meaningful statistics. However, for some types of crimes a number of different variables may indicate differences in criminal opportunity, raising the problem of which variable to use. For other types of crimes, as Frisbie *et al.* (1977) pointed out for assault, no good opportunity indicators may exist. In contrast to standardizing by opportunity measures, Harries (1981: 154-155) suggested standardizing by risk measures. For example, an assault rate could be computed using as the standardizing denominator the number of males in the age group that commit the majority of assaults.

Some research exploring non-population-based crime rates has found that the choice of the rate denominator can make a significant difference. Skogan (1978), for example, ranked the five largest cities according to motor vehicle thefts per 1000 population, and compared this ranking to the ranking according to motor vehicle thefts per 1000 vehicles. Figure 1 displays both theft rates for these cities on a scattergram. As Figure 1 shows, Skogan found that New York City has the lowest population-based rate but the highest vehicle-based rate, and that the ranking for Los Angeles shifted dramatically in the opposite direction. Boggs (1965) computed correlations across census tracts in St. Louis between population-based rates and opportunity-based rates. She found relatively low correlations for business-related offenses, and higher correlations for burglary, rape, and assault. Following up on Boggs, Phillips (1973) analyzed data from Minneapolis and found not only low, but also negative, correlations between population-based rates and alternative rates. In fact, the product-moment correlation for the data in Figure 1 is also negative (-.52).

Figure 1

Example of Divergent Alternative Crime Measures



The finding that alternative rates may diverge greatly from standard, population-based rates increases the need to understand alternative rates and how they may differ from population-based rates. The purpose of this paper is to make a modest step towards such understanding by examining several alternative crime measures using data from Oregon cities. This paper will compare the alternative measures to standard population-based rates, and will assess the implications of the choice of measures.

Data for this Study

This study uses 1980 crime data and census data for cities in the state of Oregon. The crime data are standard crime report statistics published by the State of Oregon.[1] The census data are standard demographic and housing data obtained through the Center for Population Research and Census, Portland State University. The total number of privately-owned motor vehicles in each city was estimated from the sample census items asking respondents to report the number of automobiles and vans/trucks at that residence. Merging the crime data and census data created a dataset of 133 cases (cities) for analysis.

For the sake of manageability, this study uses only two types of crimes for analysis—motor vehicle theft and burglary. One reason for choosing these crimes is that they have obvious alternatives to population for standardization. A second reason is that they are probably less prone than other felonies to under-reporting biases. Gove *et al.* (1985) concluded, based on a review of seven previous studies, that the reported UCR crime rate correlated very highly with the true rate for both motor vehicle theft (r almost 1) and burglary (r greater than .9). Similarly, O'Brien (1985: 102-103) included motor vehicle theft and burglary among the few crimes for which cross-jurisdiction comparison of UCR data is relatively justified.

Comparison of Alternatively Standardized Rates

Figure 2 replicates the Figure 1 scattergram using the Oregon cities data.[2] In contrast to Figure 1, Figure 2 reveals a very strong positive relationship between motor vehicle thefts per 1000 population and motor vehicle thefts per 1000 vehicles. The correlation is .97, indicating that these alternative measures share 95% of their variance. One city is a notable outlier, a city with an especially high average number of cars per person.[3] Removing this city from the analysis produces a correlation of .99. In short, these two alternative versions of the motor vehicle theft rate are very highly related for these cities.

Figure 3 contrasts the standard burglary rate measure, burglaries per 1000 population, with an alternative measure, burglaries per 1000 housing units.[4] The scattergram shows a strong relationship, although not quite as strong as Figure 2. The correlation between the alternative measures is .94, indicating they share 88% of their variance. Removing the single worst out-

Figure 2

Comparison of Alternative MV Theft Rates

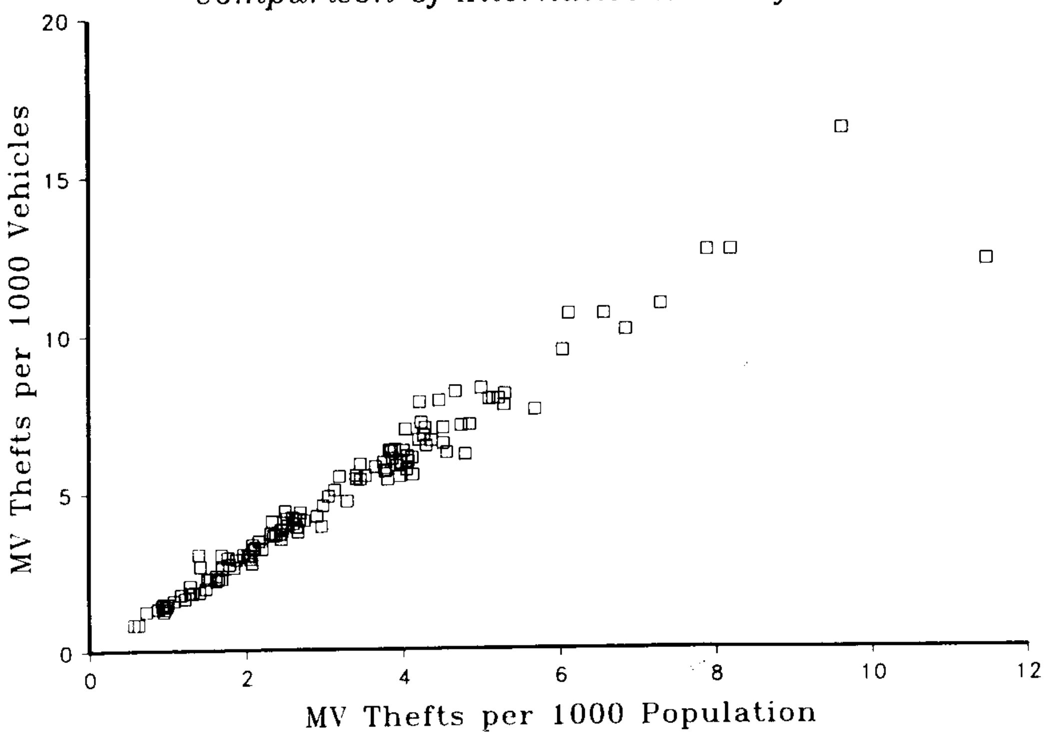
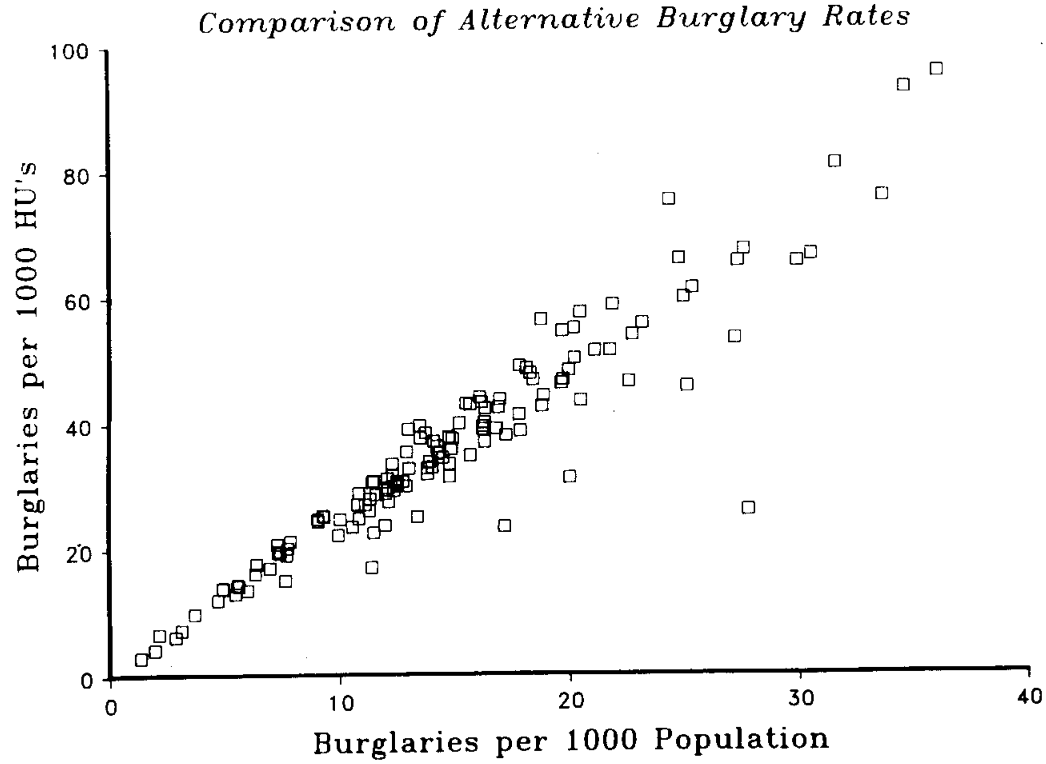


Figure 3

Comparison of Alternative Burglary Rates



lier yields a correlation and shared variance of .96 and 93%, respectively.[5] Besides the single most obvious outlier, Figure 3 shows about ten cities having low burglaries per 1000 housing unit rates, compared to the standard population-based rate. Most of these cities are seaside resort communities with an older than average age structure, a higher than average proportion of second homes, and consequently a low number of persons per housing unit.

Both of the alternative measures in Figures 2 and 3 are standardized according to the number of targets of opportunity, the type of standardization that proponents of non-population-based measures have advocated. How much of a difference does it make to use these alternative measures, compared to the standard population-based measures?

For most of these cities the choice between the standard and the alternative measures makes little difference in how that city compares to other cities, since the measures correlate so strongly. For cities that have unusual ratios of targets of opportunity (e.g. vehicles, houses) to population, as noted above, large differences can result. However, even for the most extreme outliers the alternative measures do not always provide an extremely different perspective on the cities' crime rates. For example, the worst outlier in Figure 2 ranks number one out of 125 cities on motor vehicle thefts per 1000 population, compared to a rank of four on thefts per 1000 vehicles. Either way, this city is in the top 5%. In contrast, the worst outlier in Figure 3 ranks seven out of 132 using the conventional burglary rate, compared to a rank of 95 on the alternative measure. For this city, the conventional rate greatly exaggerates the seriousness of the crime problem when viewed from a target of opportunity perspective. In short, most of the time these alternative rates are essentially equivalent to the standard rates, but for a few unusual cities the alternative measure may provide a markedly different perspective on the cities' crime problem.

Other Statistical Approaches to Rate Construction

Besides using alternatively standardized rates based on denominators other than population, a modified approach to crime rate construction is possible using standard multivariate statistical methods. This section will examine the logic and implementation of such an approach.

To understand the logic of this approach, first consider the underlying objective of standardization. Dividing crime counts by population takes into account or "controls" for population. We do this because we expect a larger number of crimes in larger jurisdictions, and computing the rate per 1000 population accounts for the difference in population when comparing large and small jurisdictions, or when examining time series data for a jurisdiction that has changed in population. Similarly, by computing an opportunity-based rate, such as motor vehicle thefts per 1000 vehicles, we in effect take into account the number of theft opportunities since we expect more crimes to result simply from greater prevalence of opportunities. As Sparks (1981:

58) argues, by standardizing we remove effects we consider trivial so as to observe effects we consider more interesting.

Despite its advantages, standardizing by a single variable limits us to removing the effects of the one variable used in the denominator. What if we want to take into account both changes in population **and** changes in the number of vehicles when examining motor vehicle theft rates over time? This would require a more flexible approach than simple rate construction.

The standard statistical technique for taking into account the effects of many variables on one interval-level variable is multiple regression analysis. Regression analysis offers a straight-forward statistical solution to the need for a more flexible method of rate construction. The first step in constructing a crime rate would be to estimate a crime-specific regression model using cross-sectional data. The dependent variable would be either the number of crimes or the standard crime rate, depending on which of two slightly different alternative statistical approaches the analyst decided to implement.[6] The independent variables would include the "standardizing" variables—such as size, opportunity, and risk factors:

$$c = a + B'X$$

where: c = number of crimes, or crimes/1000 population
X = vector of independent variables, including size, opportunity, and risk factors
a,B = parameters to be estimated (B is a coefficient vector corresponding to the X variables.)

Once the model is estimated, the residual for each jurisdiction (the difference between that jurisdiction's actual and predicted crime rate or number) becomes a measure of that jurisdiction's crime problem, analogous to a simple crime rate. Like a rate, the residual reflects the seriousness of the jurisdiction's crime problem, but unlike a simple rate, it takes into account not just one standardizing variable, but as many variables as are included in the regression equation.

Although computationally more complex and certainly less intuitive, this residualized approach to measuring jurisdictions' crime levels is logically a generalization of the standard simple rate approach. To help in presenting residualized crime rates to others, analysts could explain that these rates show how much lower, or higher, a jurisdiction's crime rate is compared to what we would expect based on the jurisdiction's characteristics. The results could be expressed as a deviation from the expected number of crimes, as a deviation from the expected crimes per 1000 population, or as a ranking or percentile ranking of that jurisdiction within the population used for analysis (such as the cities in a state).

To illustrate this technique, this study estimated regression models for motor vehicle theft rates and for burglary rates using the Oregon cities data.

The first specification decision concerns the form of the dependent variables, and for this analysis the dependent variables were the vehicle theft rate per 1000 population and the burglary rate per 1000 population.[7] The most important specification decisions concern the choice of the independent variables. These should ideally include all the important opportunity or risk factors the researchers want to take into account. Past research (e.g. Danziger, 1976; Harries, 1980; System Development Corporation, 1974; Davidson, 1981; Boydell, 1969; Blau and Blau, 1982) provides some guidance by identifying crime predictors, such as income level, education level, racial composition, age composition, unemployment, region, growth rate, and density. The independent variables in the regression equations for this analysis include opportunity, risk, and demographic factors for which data were available. Population was included to allow for the possibility that crime rates vary by city size. Specific opportunity factors included as predictors in the burglary regression were percent vacant housing units, percent urban housing units, percent rental housing units, and average persons/housing unit. For the motor vehicle theft regression, the average number of motor vehicles per person was included as an opportunity factor.[8]

Table 1 shows the results for both regression equations. The burglary regression has about twice the predictive power—34% variance explained compared to only 17% for the motor vehicle theft regression. In both equations the population variable reveals a strong size effect; larger cities appear to have higher crime rates, **ceteris paribus**. Higher percentages of rental housing strongly predict to higher burglary rates. Higher percentages below the poverty line also predict to higher burglary rates but to lower motor vehicle theft rates (although the effect is weaker). The age variables, included as risk indicators, do not show the expected results.[9] To guard against the possibility that the wide size range of cities in the analysis greatly distorted these results, precautionary checks were made and no large distortions were found.[10]

A comparison of the residualized motor vehicle theft rate to the rate per 1000 population and per 1000 vehicles reveals correlations of .91 and .90, respectively. These correlations are notably lower than the correlation (r=.97) between the population-based and vehicle-based rates. Similarly, the residualized burglary rate correlates .81 and .82 with the rate per 1000 population and per 1000 housing units, respectively. These correlations are notably lower than the correlation (r=.94) between the population-based and housing-unit based rates. In short, although the residualized rates correlate quite highly with the other rates, they differ more from the population-based and target-based rates than the population-based and target-based rates differ from each other.

Figure 4 shows a scatterplot of the residualized burglary rate compared to the standard population-based rate. The largest negative outlier (point below the scatter) is for the city of Portland, which has a high per capita rate

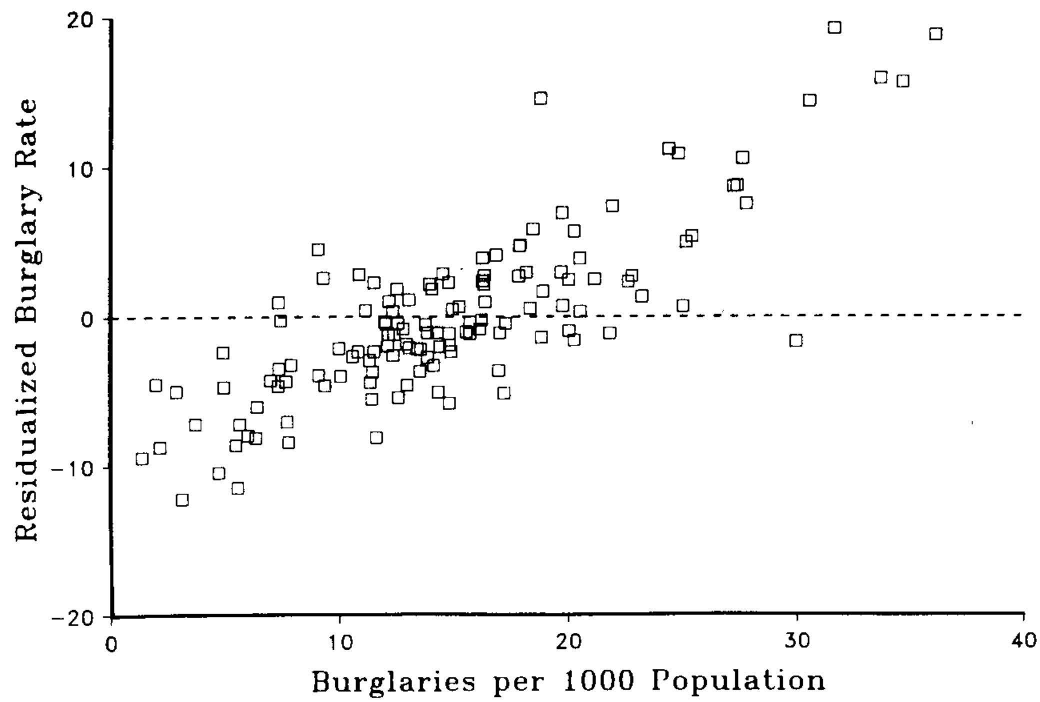
Table 1

MOTOR VEHICLE THEFT RATE AND BURGLARY RATE REGRESSION RESULTS

	M.V. Theft Regression			Burglary Regression		
Independent Variables	b	t	st. b	b	t	st. b
Economic						
Median Income ($1,000's)	-.15	-1.54	-.24	.20	.52	.08
Percent Below Poverty Line	-.11	-1.67	-.23	.52	2.36	.29
Percent Men 16+ Unemployed	.03	.67	.07	.15	.95	.08
Race						
Percent Blacks	-.56	-1.24	-.23	-1.21	-.77	-.13
Percent Hispanic	.05	1.30	.14	-.19	-1.49	-.14
Education						
Percent 18+ Not H.S. Graduate	-.03	-.63	-.12	-.11	-.7	-.13
Percent College Graduate	-.04	-.68	-.11	-.09	-.51	-.07
Age						
Percent 10-13 Males	.24	.85	.09	.97	1.15	.11
Percent 14-17 Males	-.63	-2.70	-.27	.43	.51	.05
Percent 18-21 Males	-.02	-.12	-.01	-1.52	-2.47	-.28
Size						
Population (1000's)	.02	2.2	.41	.06	1.83	.31
Housing						
Percent Vacant Housing Units	---	---	---	.18	1.33	.17
Percent Urban Housing Units	---	---	---	.02	.76	.08
Percent Rental Housing Units	---	---	---	.38	4.10	.43
Persons/Housing Unit	---	---	---	2.26	.61	.10
Motor Vehicles						
Motor Vehicles/Person	2.51	.74	.08	---	---	---
Multiple Correlation (R)	.41			.58		
% Variance Explained (R^2)	17%			34%		
Number of Cases (N)	125			132		

Note: The "b" columns show the unstandardized partial regression coefficients for that independent variable. The "t" columns give the t-statistics for the coefficients. The "st. b" columns give the standardized partial regression coefficients.

Figure 4
Residualized vs. Standard Burglary Rate



but a lower than average residualized rate. Portland's low residualized rate results from its high percentage of rental housing, its somewhat higher than average percent below the poverty line, and its very large population. Taking into account these characteristics and their expected effects, the residualized rate tells us that Portland's rate is fairly low, despite its high per capita rate. In contrast to Portland, the largest positive outlier is Dundee City, which has only a slightly above average per capita rate but a very high residualized rate. Dundee City's high residualized rate results from its very low levels of rental housing, poverty, and population. The residualized rate tells us that taking into account these characteristics, Dundee City has a high level of burglaries.

Assessment of Differences Among Alternative Rate Measures

Now that we have examined several alternatives to constructing crime rates, the obvious question arises—how much difference does it make which alternative we use? As noted earlier, the intercorrelations among the three alternative motor vehicle theft rates used in this study range from .90 to .97, and the intercorrelations among the three alternative burglary rates range from .81 to .94. These strong positive correlations mean that cities having high (or low) crime rates according to one measure will usually have high (or low) rates on an alternative measure. Nonetheless, these measures differ sufficiently so that the choice of measures can affect quite radically, as we saw earlier, the apparent seriousness of the crime problem for some cities.

In order to explore further how the choice of measures can affect the apparent seriousness of the crime problem for selected cities, Table 2 presents

Table 2

RANKINGS OF TOP TEN BURGLARY AND MV THEFT CITIES ON ALTERNATIVE RATE MEASURES

Alternative Burglary Rates			Alternative MV Theft Rates		
POP	HU	RES	POP	VEH	RES
1	1	2	1	4	1
2	2	4	2	1	2
3	4	3	3	2	5
4	3	1	4	3	3
5	7	6	5	5	4
6	10	80	6	8	8
7	95	12	7	6	6
8	6	9	8	7	67
9	9	10	9	9	9
10	20	11	10	19	28

Note: The POP column gives the ranking according to crimes per 1000 population, the HU column per 1000 housing units, and the VEH column per 1000 vehicles. The RES column gives the ranking on the residualized rate measure.

the rankings of the top ten per-capita burglary rate cities on the alternative burglary rate measures, and also presents the rankings of the top ten per-capita motor vehicle theft rate cities on the alternative motor vehicle theft rate measures.[11] The rankings show that for most of these cities the choice of rate does not make a remarkable difference, since the choice shifts the ranking only slightly. However, for several of these cities the choice of the rate does make a remarkable difference. For example, Portland ranks number six on burglaries/1000 population and number ten on burglaries/1000 housing units, but ranks a lowly 80 on the residualized measure, for the reasons discussed earlier. Similarly, Canon Beach ranks number seven on burglaries/1000 population and number 12 on the residualized rate, but ranks a lowly 95 on the burglaries/1000 housing units.

The shifting burglary rankings for Canon Beach illustrate well some of the factors underlying these rate computations. Since Canon Beach is a seaside resort city with a very low average persons per housing unit, standardizing the rate by housing units greatly lowers the city's ranking, compared to standardizing by population. This simply tells us that although

Canon Beach has a high rate compared to its population, it has a low rate compared to its housing supply. The residualized rate, however, not only takes into account the population size and the persons per housing unit, but also takes into account the percent rental housing units and percent vacant housing units. As a resort city, Canon Beach not surprisingly has very high percentages of rental and vacant housing units which increase the regression-computed expected rate, leading to a higher ranking on the residualized burglary rate.[12] In short, the high residualized rate for Canon Beach, compared to its low housing unit rate, results because the residualizing methodology takes into account the unusual housing characteristics of the city, not just one characteristic.[13]

Conclusion

Exclusive reliance on population-based crime rates stems more from blind tradition than from logic or merit. This study has demonstrated the feasibility of using several alternative types of rates, and has shown the additional information these rates can convey. By augmenting traditional population-based rates, alternative types of rates can increase the flexibility criminologists, researchers, and other users of crime data have in choosing an appropriate rate, and can provide more information on the prevalence of crime.

Population-based crime rates alone provide a distorted picture of the crime problem for some jurisdictions. Granted, alternative crime measures intercorrelate strongly enough that they will lead to similar conclusions most of the time. For unusual jurisdictions, however, alternative crime measures can provide very different perspectives, and using population-based rates alone can mislead more than illuminate. Restricting ourselves to one type of measure in such instances creates a self-imposed myopia.

Moreover, for some purposes alternative rates may prove superior to population-based rates. Opportunity-standardized rates lend themselves better than population-standardized rates to probability interpretations. Residualized rates give analysts greater flexibility in deciding which factors to control when comparing rates over time or across jurisdictions. Regardless of exactly which type of rate proves most useful for what purposes, we need to unshackle ourselves from the parochial allegiance to one type of rate alone.

Acknowledgments

The author thanks Edward Schafer, Director, Center for Population Research and Census, Portland State University, for providing the census data used in this study, and for his helpful comments.