



Inequality and Political Violence Revisited

T. Y. Wang; William J. Dixon; Edward N. Muller; Mitchell A. Seligson

The American Political Science Review, Vol. 87, No. 4 (Dec., 1993), 979-994.

Stable URL:

<http://links.jstor.org/sici?sici=0003-0554%28199312%2987%3A4%3C979%3AIAPVR%3E2.0.CO%3B2-Y>

The American Political Science Review is currently published by American Political Science Association.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/apsa.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

INEQUALITY AND POLITICAL VIOLENCE REVISITED

In their 1987 article in this Review, Muller and Seligson used logged ordinary least-squares (LOLS) to estimate the effect of income inequality on cross-national levels of deaths by political violence. T. Y. Wang challenges the robustness of the main conclusion and argues for the application of a maximum likelihood approach—the exponential Poisson regression (EPR) model—rather than LOLS. He concludes that the widely used LOLS approach yields misleading conclusions when applied to event count data. Dixon, Muller, and Seligson replicate previous work using both LOLS and EPR approaches and conclude that in most—but not all—respects the two approaches yield similar results, supporting the effect of inequality when the specifications are identical. They also argue (in response to concerns expressed by Brockett 1992) that the inequality results are robust when account is taken systematically of the best information on underreporting of deaths.

COMMENT

Inequality in its various forms has long been regarded as a major determinant of political violence. The theoretical foundation of this argument is the relative deprivation thesis, which holds that discontent generated from the gap between an individual's expected and achieved well-being is a major cause of violent behavior (Gurr 1970).¹ A body of quantitative literature has emerged to examine the causal linkage between inequality and collective violence.² While different findings have been registered, most of the studies share two important methodological similarities: (1) many studies used *event count data*³ as measurements of dependent variables with a measure of deaths from violence as the most widely used indicator (see, e.g., Boswell and Dixon 1990; Hartman and Hsiao 1988; Muller 1985, 1988b; Muller and Seligson 1987; Park 1986; Weede 1981, 1986), and (2) many studies used the logged ordinary least squares (LOLS) method for statistical estimation. As will be demonstrated, LOLS is not appropriate when event count data are used for measurements of dependent variables. The unsuitability comes about because event count variables typically display a Poisson or other count-based distribution; therefore LOLS produces inefficient or biased estimates (King 1988, 1989a–b).

Muller and Seligson (1987) recently examined the relationship between inequality and political violence. Using the LOLS method, they concluded that income inequality (the share of income going to the top 20% of all households) has a strong positive effect on mass violence. I shall replicate this study to demonstrate the unsuitability of LOLS when dependent variables are measured by event count data. I shall then reanalyze their data by using an event count analysis—a statistical technique that has not been used in previous research on political violence but has been successfully employed in studies on presidential appointments of Supreme Court justices (King 1987), on the demand of health insurance and the use of health care services (Cameron and Trivedi

1986), and on patents and research and development activity (Hall, Griliches, and Hausman 1986; Hausman, Hall, and Griliches 1984). The replicated analysis reveals that inequality in general—whether in the form of income or land maldistribution—does not have statistically significant effect on collective violent behavior. Instead, a semirepressive regime and intensity of separatism are the main causes of mass violence.

DIRECT REPLICATION

The frequent adoption of LOLS in the study of violence stems from the fact that the ordinary least squares method assumes a linear and additive function. However, event count data display a Poisson or other count-based distribution, and the relations between variables are not linear in functional forms.⁴ In such cases, ordinary least squares solutions do not provide proper estimates. A frequently adopted remedy in the literature is to make a natural logarithmic transformation of the dependent variable so that a log-linear function is acquired. All of the studies on political violence I have mentioned adopt this approach. In many instances, LOLS is not suitable for event count data when zero is one of the observed values. This is a frequent occurrence in the study of political violence. A constant c is often added to the dependent variable Y_i before obtaining a logarithmic transformation, so that a zero value can be avoided, namely, $\ln(Y_i + c)$.⁵ As will be demonstrated, the value of this added constant has substantial effects on the conclusions reached.

Table 1 reports the LOLS results of a direct replication of Muller and Seligson's study, with a small constant, c , taking on different values added to the dependent variable, the logged death rate of 1973–77.⁶ Since Monte Carlo experiments have shown that there is no optimal choice for c across empirical examples (King 1988), the values of this constant are arbitrarily set at .001, .01, and 1. As columns 1–3 show, when the value of c changes, the magnitudes of all regression coefficients and their corresponding standard errors also change. The size of almost all LOLS coefficients drops as c increases. For instance,

TABLE 1

LOLS Regression of Political Violence on Income Inequality and Other Explanatory Variables

VARIABLES	C = .001	C = .01	C = 1
Upper 20 % income share	.08 (.05)	.07 (.04)	.038* (.017)
In energy consumption per capita	.46 (.34)	.30 (.25)	.11 (.11)
Intensity of separatism	2.71* (1.10)	2.14* (.82)	1.23** (.36)
Semirepressive regime ^a	4.17** (.97)	3.06** (.73)	.98*** (.32)
In negative sanction per 1m			
1973-77	1.11 (.69)	.95 (.52)	.33 (.23)
1968-72	-.22 (.80)	-.06 (.60)	.13 (.26)
Adjusted R ²	.44	.47	.46

Note: Figures in parentheses are standard errors.
^aThe regression coefficient of "Semirepressive regime" is different from that reported in Muller and Seligson 1987, panel 1.3. While all other coefficients are identical to the original ones, it is empirically impossible to have a different coefficient for "Semirepressive regime". I hence conclude that the difference is due to a misprint in Muller and Seligson (1987). N = 61.
 *p < .05.
 **p < .01.

the coefficient of semirepressive regime changes from 4.17 to .98—a decrease of 76%—as *c* increases from .001 to 1. In addition, the level of significance also changes. Income inequality (upper 20% income share), for example, is not significant when *c* is .001. As *c* increases to 1, it is significant at the .05 level. The arbitrary choice of an added value for a logarithmic transformation can thus have enormous effects on one's substantive conclusions. This raises additional questions about the accuracy and reliability of previous findings.

INDIRECT REPLICATION: AN EVENT COUNT ANALYSIS

This analysis demonstrates that the commonly adopted LOLS regression analysis is inappropriate when political violence is measured by event count data. The use of event count analysis represents a more reliable and accurate alternative (Cameron and Trivedi 1986; King 1988, 1989a-b). The advantages of event count analysis are (1) it takes into account the nonnegativity of event count data and leaves no possibility of predicting a negative value (least-squares techniques, by contrast, do not constrain the expected number of events to positive values, which must be the case)⁷ and (2) since zero is a natural outcome of event count models, there is no need for the logarithmic transformation and hence no "zero problem" (Hausman, Hall, and Griliches 1984).

Probably the simplest specification of an event count analysis is the exponential Poisson regression (EPR) model. Specifically, the EPR model takes the form of

$$E(Y_i) = \lambda_i = \exp(\mathbf{X}\boldsymbol{\beta}) = \exp\left(\sum_{k=1}^p X_k\beta_k\right), \quad (1)$$

where $i = 1 \dots N$, and $p < N$; $E(Y_i)$ is the expected value of the dependent variable; \mathbf{X} is a vector of independent variables; and $\boldsymbol{\beta}$ is a vector of regression estimates. The goal is to estimate the vector of regression estimates, $\boldsymbol{\beta}$, in which each element represents the individual effect of X_i , a particular independent variable, on the dependent variable or the observed event count.

The Poisson specification is restrictive in one important way, however: it is based upon assumptions that events occur *independently* over time and that the probability of an event occurring at any point of time is *constant*. These assumptions may not be valid, since either the prior occurrence of an event such as political violence or changes of other factors such as economic conditions could result in a change of the probability of subsequent events. In addition, it is common practice in cross-national studies to aggregate data over a period of time (e.g., 5 or 10 years). The aggregation process may increase cross-national variability since countries tend to have different incident-proneness at different time periods. A technique that explicitly takes into account *overdispersion*⁸ in event counts is *negative binomial regression*. Statistically speaking, this means that the specification of the variance of Y_i is different in the EPR and negative binomial models even though they have an identical specification of expected values (see equation 1). While the EPR model assumes that the expected value and variance of Y_i are equal (i.e., $V(Y_i) = \lambda_i$), negative binomial models relax this assumption.

Cameron and Trivedi (1986) have shown that two negative binomial models—what they called NEGBIN I and NEGBIN II—can be generated by different specifications of variance:

$$V(Y_i) = (1 + \alpha)\lambda_i \text{ (NEGBIN I)} \quad (2)$$

$$V(Y_i) = \lambda_i(1 + \alpha\lambda_i) \text{ (NEGBIN II)} \quad (3)$$

The adequacy of the Poisson model is thus based on tests of the nuisance parameter, $\alpha = 0$. If α is not significantly different from zero, which implies that events are independent, the expected value and variance of Y_i are equal and the negative binomial models will reduce to exactly to the Poisson model. If α is greater than and significantly different from zero, which implies that events are overdispersed and/or have a heterogeneous probability of occurrence, the Poisson model should be rejected. Meanwhile, the choice between NEGBIN I and II is also contingent on tests of α in equation 2 and 3.

With these statistical qualifications in mind, I will replicate Muller and Seligson's study. There are two

modifications in the operationalization. First, Muller and Seligson undertook a logarithmic transformation on three independent variables: energy consumption per capita, negative sanctions per one million of 1973–77, and the same of 1968–72. The following analysis does not have that transformation on these variables, since an event count analysis assumes a nonlinear function, so that there is no reason to make a logarithmic transformation on independent variables to fit the linear assumption. Second, the dependent variable of the original model, political violence, was measured by “the *logged* sum of annual deaths from domestic political conflict during 1973–77 *divided by midinterval population*” (Muller and Seligson 1987, 433–34, emphasis mine). No such transformation was done on the dependent variable in the following analysis. Instead, the natural logarithm of the mid-interval population ($\ln \text{POP75}$) is included on the right-hand side as an additional independent variable. This is done because the original form of the relationship is

$$\frac{E(Y_i)}{\text{POP75}} = \exp(\mathbf{X}\boldsymbol{\beta})$$

or, equivalently,

$$E(Y_i) = \exp\left(\sum_{k=1}^p X_k\beta_k + 1n(\text{POP75})\right).$$

This exercise is due to the fact that a linear transformation on the dependent variable (such as taking an average) is incompatible with event count analysis. This method of measurement is conceptually and algebraically identical to the original operationalization. That is, political violence is measured by the death rate from domestic conflict per one million population. Algebraically, this implies that the coefficient of $\ln(\text{POP75})$ should be constrained to one. It is harmless, however, to free it from this restriction by including it as an independent variable in the equation (King 1988, 1989a).

Furthermore, Muller and Seligson exclude all land inequality variables in the full model, based upon the initial analyses of the relations between all inequality variables and political violence (1987, 435–36). Since their initial analyses are mainly based upon bivariate regression/correlations, the exclusion of land inequality variables from the subsequent analyses is not appropriate. Any relationship or lack of relationship between variables in an incomplete specification such as this may either be spurious or suppressed (Pedhazur 1982). The following analysis will include all inequality variables except “Gini land concentration.”⁹ As Prosterman pointed out, it is implausible to expect that the relatively high Gini coefficients of land concentration in industrialized societies, where only a small agricultural labor force is present, would generate significant discontent (1976, 350).¹⁰

Several countries register extremely large scores on the dependent variable. In order to reduce the prob-

lem of estimation, Muller and Seligson set a ceiling on the values of the dependent variable. The death rates of these countries were arbitrarily set to 50 (1987, 434).¹¹ Their decision to set an upper limit to the death rate is inappropriate, because an arbitrary ceiling could skew the conclusions. I shall include the original values of these cases and examine their robust standardized residuals¹² to determine whether their presence would bias the results.¹³

Before the actual effects of the inequality variables on political violence can be examined, it is necessary to determine whether a Poisson model is adequate. As noted, the adequacy is contingent upon tests of the nuisance parameter, α , being significantly greater than zero. Consider the Poisson residuals, $u_i = y_i - \hat{\lambda}_i$, and let $\hat{V}(y_i) = (y_i - \hat{\lambda}_i)^2$. The following regressions can be obtained (standard errors in parentheses are obtained using White’s heteroscedastic consistent estimator):

$$\hat{V}(y_i) = 1,458.27 \hat{\lambda}_i \quad (771.53) \quad (4)$$

$$\hat{V}(y_i)/\hat{\lambda}_i = 530.54 - .63 \hat{\lambda}_i. \quad (324.13) \quad (.63) \quad (5)$$

From equations 2 and 4, it can be seen that $\hat{\alpha} = 1,458.27 - 1$. The estimated T -statistic is 1.89, which is statistically significant at the .05 level with a one-tailed test.¹⁴ This evidence clearly rejects the adequacy of a Poisson specification in favor of a negative binomial parameterization.

Which of the two negative binomial models should be used? The $\hat{\alpha}$ in equation 5 is not statistically significant at the .05 level. This suggests that the first rather than the second negative binomial model is adequate. Table 2, column 1 presents the results with a NEGBIN I parameterization.¹⁵

Surprisingly enough, none of the regression coefficients in Table 2, column 1 is statistically significant at the .05 level for a two-tailed test. Only intensity of separatism and semirepressive regime are of borderline significance (between the .05 and .10 levels). An analysis of the residuals reveals that Argentina, Peru, and the United Kingdom are outliers.¹⁶ In order not to be biased by a few outliers, I drop these three cases and reestimate the parameters. Parallel regression equations to equations 4 and 5 are first obtained to determine an appropriate specification:

$$\hat{V}(y_i) = 73.01 \hat{\lambda}_i \quad (32.83) \quad (6)$$

$$\hat{V}(y_i)/\hat{\lambda}_i = 99.53 + .02 \hat{\lambda}_i. \quad (39.88) \quad (.05) \quad (7)$$

Equation 6 again suggests that the negative binomial is preferred to the Poisson specification; with $\hat{\alpha}$ in equation 7 being statistically insignificant, it suggests that the NEGBIN I is still a preferred specification.

Table 2, column 2 presents the results after the outliers are removed from the analysis. Again, as in column 1, none of the inequality variables has statis-

TABLE 2
Negative Binomial Regression of Political Violence on Inequality and Other Explanatory Variables

VARIABLES	ALL VALID CASES INCLUDED (N = 45)	OUTLIERS EXCLUDED (N = 42)
Agrarian inequality	-.016 (.023)	-.043 (.031)
Landless	-.004 (.027)	.048 (.032)
Upper 20% income share	.032 (.029)	.032 (.026)
Energy consumption	-.0001 (.012)	.002 (.012)
Intensity of separatism	.733 (.408)	1.180** (.434)
Semirepressive regime	.984 (.508)	1.369** (.550)
Negative sanctions per 1m		
1973-77	-.001 (.032)	.012 (.034)
1968-72	.068 (.116)	-.005 (.108)
ln (Population, 1975)	.297 (.166)	.171 (.150)
Log-likelihood	82,224.5	43,258.7

Note: Figures in parentheses are standard errors.

** $p < .01$.

tically significant effect on collective violence. This finding contradicts Muller and Seligson's strong claim that income inequality is the major determinant of mass violence. It also runs counter to the argument that landlessness, the proportion of the labor force that does not own or possess ownershiplike tenure of land, contributes to violence (Prosterman 1976; Prosterman and Riedinger 1987). Agrarian inequality (the multiplicative interaction between land inequality and the size of the agricultural labor force) also fails to warrant a causal effect on collective violent behavior. The evidence presented here indicates that inequality in general does not have statistically significant effect on mass violence.

The only variables that are associated with political violence are semirepressive regime and intensity of separatism, which serves as a control variable. The positive effect of intensity of separatism can be explained on the grounds that political violence often occurs when groups or regions attempt to gain greater autonomy or independence. The significant semirepressive regime lends some support for the resource mobilization interpretation of political violence. Its advocates argue that the determinant of collective violence is not the discontents generated from inequality per se but rather the organization of discontent.¹⁷ Individuals are more likely to be orga-

nized and engage in violent behavior in a semirepressive regime than in a repressive one. Differences in regime structure are thus relevant because they directly affect individuals' ability and likelihood to get organized.

Moreover, semirepressive regimes also serve as causes of collective violent behavior. The Iranian revolution documented in detail by Green (1982, 1984) and cited by Muller and Seligson (1987, 430) is a case in point. Green argues that the pseudoparticipation tactics undertaken by the shah "served to increase popular hostility among those socially mobilized Iranians eager to have a measure of influence over the manner in which their society was ruled" (1984, 155). This unrealized desire for real power in a semirepressive regime can thus become a contributor to collective violence. As a result, a semirepressive regime is not only a facilitating factor but also a direct cause of political violence.

CONCLUSION

My analysis indicates that Muller and Seligson's previous findings are suspect due to the inappropriate statistical methods employed. The new evidence suggests that inequality in the forms of income or land maldistribution does not have effects on political violence. Rather, regime structures and intensity of separatism are more important in explaining the occurrence of political violence. Semirepressive regimes are relevant not only because they allow individuals to get organized and engage in violent behavior but also because they are causes of violence.

The fact that none of the inequality variables is statistically significant does not mean that the relative deprivation thesis is theoretically meaningless. In Gurr's (1970) conceptualization of relative deprivation, discontent is not generated from inequality per se, but rather from the gap between an individual's expected and achieved well-being. Indices based on measures of objective inequality cannot, by themselves, purport to assess this crucial part of Gurr's hypothesis. Inequality in the form of income or land maldistribution will not be translated into widespread discontent if there is no perceived discrepancy between what people actually get and what they expect to get. It is necessary, therefore, to assess individuals' perception of this gap through survey data before the relative deprivation hypothesis can be tested—an argument made more than a decade ago (Eckstein 1980; Weede 1981).

The methodological implication of this analysis, however, goes well beyond the study of political violence. There are many other studies in the discipline that use event count data as the measurements of dependent variables. The number of coups, coup attempts, presidential vetoes, and presidential appointments of Supreme Court justices are only a few examples. Far from being only a problem of technical differences, the widely used conventional method can lead to incorrect or misleading substantive find-

ings, as demonstrated in the reanalysis of Muller and Seligson's data. Least squares solutions provide inefficient and biased estimates when the dependent variable uses event count data. An event count analysis is the appropriate technique for statistical estimation, since the underlying data generation process of the dependent variable is a Poisson or other count-based distribution.

T. Y. WANG

Illinois State University

RESPONSE

T. Y. Wang attacks cross-national studies of political violence on grounds that they routinely rely on faulty estimation methods. Similarly, Brockett has claimed that the conventional dependent variable in such studies, deaths from political violence, is "mortally flawed" by measurement error (1992, 169). In regard to estimation methods, Wang directs two key criticisms at contemporary conflict studies, which, if valid, undermine the empirical foundation for the proposition that income inequality promotes political violence. He asserts (1) that the common practice of logging dependent variables derived from counts leads to capricious and unreliable results and (2) that when estimated by appropriate maximum likelihood techniques, the relationship between inequality and political violence disappears. Based on documentary evidence for Central American countries, Brockett demonstrates that counts of deaths from political conflict events from the *World Handbook of Political and Social Indicators* (Taylor and Jodice 1983) are grossly underreported for several countries. He infers that this kind of measurement error is likely to be prevalent for the Third World in general and, if so, that the *World Handbook* data are too inaccurate to permit meaningful cross-national testing of hypotheses about determinants of political violence.

These criticisms are more than methodological quibbles about the validity of the finding of a positive effect of income inequality on political violence. They seriously challenge the empirical foundation underpinning our substantive understanding of the macrosocietal causes of political violence in general. We shall address each criticism. As we shall show, these criticisms fail to overturn previous results.

LOGARITHMIC TRANSFORMATION OF CONFLICT MEASURES

Wang's initial criticism is leveled at the near universal practice of transforming the dependent variable into logarithms prior to estimation by ordinary least squares (OLS). He argues two separate briefs against this procedure. The first hinges on the observation that violent deaths, death rates, and similar

count-based conflict measures commonly used as dependent variables cannot possibly assume negative values. Least squares techniques are therefore conceptually inappropriate because they can provide no assurance that predicted values will be positive. The second argument arises from the fact that zero is a typical value of these dependent variables and since the logarithm of zero is undefined, some small arbitrary constant must be added to these measures. He presents results showing that the choice of a constant entails serious consequences for subsequent estimates and for substantive conclusions drawn from them. We address each of these issues in turn.

The objection to negative predicted values with count-based measures is a special case of a more general conceptual problem with ordinary least squares that can occur whenever a dependent variable is theoretically constrained to a particular range or set of values. The problem is most often discussed in the context of dichotomous dependent variables, though it obviously applies with equal force to event counts. We certainly have no quarrel with Wang's position here. We, too, prefer models that impose some constraint on predictions of negative deaths or death rates. However, we do quarrel with Wang's implication that OLS models are just as susceptible to this problem as standard ordinary least squares. In fact, OLS always imposes some level of constraint on negative predictions.

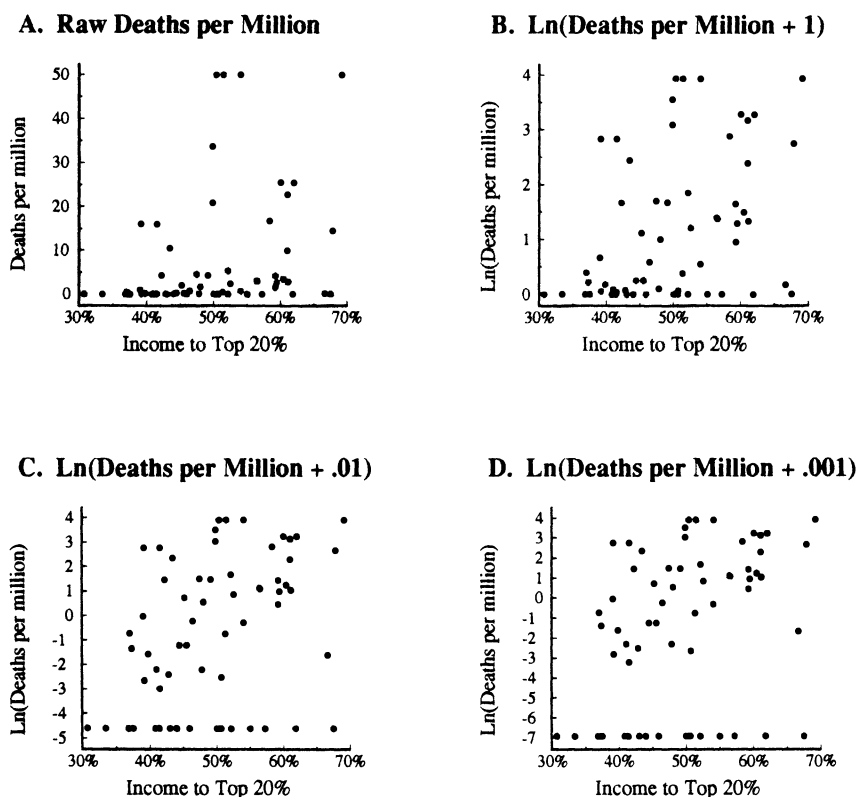
It is important to keep in mind that the dependent variable of an OLS model is expressed in logarithms and therefore its predicted values are also in logarithms. It is probably fairly common for OLS estimates on counts to produce negative predictions but these values are not themselves counts, they are *logged* counts.¹⁸ Since negative logarithms map within the range of zero to one, the problem of negative counts or rates is averted. One might still object to a prediction of, say, half a death; but fractions routinely occur elsewhere as well and are not really at issue.

It is true that under some conditions the OLS constraint on negative predictions may be only a bounded constraint. This will occur whenever a positive constant is added to an original measurement so that logarithms of zero counts can be taken. In this case, OLS can produce negative predictions, but only within a fixed range determined by the magnitude of the constant. For example, suppose an OLS model on raw death counts produces a predicted value of $-.69$ which, after exponentiation, is equivalent to one-half a death. If the original data had been incremented by a constant larger than this predicted count, then removal of the constant would yield a negative prediction. Notice that a negative result can only occur from subtraction of a constant and that the range of negativity is strictly bounded by the constant's size. In our view this is nothing more than a minor nuisance.

The addition of a constant prior to taking logarithms is at the heart of Wang's second line of argument against OLS estimation. The purpose of

FIGURE 1

Inequality and Deaths per Million Logged with Different Increments



Note: Deaths per million are totals for 1972-77 divided by 1975 population. Inequality Data is ca. 1970. There are 62 valid observations.

this operation is to transform zero counts or rates into positive values, and since any positive constant will do the job, Wang maintains that the choice of constant is essentially arbitrary. This is a critical point because he goes on to show in his Table 1 that the choice of constant affects the magnitudes of subsequent parameter estimates and their standard errors in such a way that substantive conclusions are called into question.¹⁹ Most importantly, with the addition of .001 instead of 1.0—the value conventionally used in conflict studies²⁰—the estimated effect of income inequality on political violence fails to be significant at the .05 level even by a one-tailed test. He concludes that one can have little confidence in substantive conclusions based on OLS estimates that are so drastically influenced by an arbitrary constant.

Wang is obviously correct in claiming that any positive constant added to zero will permit transformation to logarithms. In this sense, the choice of constant is arbitrary. We take issue, however, with his unstated assumption that this choice is also free of any substantive implications: it is not. In particular, we shall argue that under certain circumstances a constant can introduce step-level scaling changes that may or may not correspond with prior theoretical assumptions. Thus, while the need for a constant

may be purely methodological, the particular value chosen must be consistent with prior theory.

We can best illustrate these theoretical implications by focusing on the relationship Wang finds most seriously affected by the choice of constant—that between income inequality and death rates (per million population) from political violence. Figure 1 displays four scatterplots for the 62 countries with valid observations on both measures.²¹ Plot A in the upper left shows the relationship between inequality and raw death rates. Plots B, C, and D use logarithms of death rates with increments of 1.0, .01, and .001, respectively. In the unlogged case (plot A) the relationship is evidently exponential. Specifically, the death rate from political violence is a positively accelerated function of income inequality. This is consistent with the hypothesis that structural conditions of discontent must reach relatively high levels for mobilization in the form of rebellious collective action to occur (Muller 1985, 53).

When deaths per million are logged with a constant of 1.0 in plot B, the transformed relationship approaches linearity with minimal changes to the scaling at low-rate levels, which is consistent with the hypothesis of a positively accelerated function. Plot C with a constant of .01, however, is quite different.

Here we observe a gap beginning to open between observations having the lowest observed nonzero death rates and those with zero deaths. This gap becomes even more pronounced with an increment of .001 in plot D. The plots in Figure 1 demonstrate, with some clarity, why Wang's OLS results produced parameter estimates for income inequality that became smaller in relation to their standard errors as the constant adjustment became smaller. As the interval between zero and very small death rates widens, observations are moved away from any least squares line and standard errors become progressively inflated.

The distance between zero and nonzero rates or counts respond in this way because the logarithmic function is symmetrical around one. Thus, as fractions between zero and one become smaller, the magnitude of their logarithms (expressed in negative numbers) increases at an accelerating pace, paralleling the acceleration of positive values above unity. This has an obvious but strangely counterintuitive implication for the choice of a constant value for taking logarithms. Intuition might suggest a preference for .001, rather than 1.0, on grounds that the smaller the value added to a measure the less distortion it will introduce. But in fact, as fractional constants become smaller their logarithms become progressively larger negative numbers.

A table of logs (or, to be more contemporary, a moment with a hand calculator) will immediately reveal how a very small fractional constant like .001 added to an integer greatly expands the distance between zero and one in comparison to the nearby intervals between one and two, two and three, and so on. Moreover, by the symmetry of the logarithmic function, a constant adjustment of .001 necessarily represents the logged interval between zero and one as equivalent to the logged distance between one and one thousand. Substantively, this implies that a single death from political violence constitutes a huge step-level change from zero deaths. This step-level representation does not conform to our own hypothesis of a smoothly accelerating function, nor does it conform to any hypotheses that have been expressed by other conflict theorists. Indeed, we cannot conceive of any meaningful theoretical justification for such a representation.²²

The problems caused by small fractional constants are even more severe with rates. At least with counts the distortion is confined to a single interval. But rates less than one are affected even more perversely since distances between logged rates grow progressively larger as the observed rates themselves become smaller. Frankly, we cannot begin to imagine what sort of conflict theory could justify such a result. It strikes us as pure nonsense.

In contrast, addition of a constant 1.0 before taking logarithms progressively compresses all observed counts and rates as they become larger, both within and beyond the zero-one interval. This is a theoretically appealing result for conflict studies because it appropriately linearizes the hypothesis that political

violence is a smoothly accelerated positive function of income inequality. Of course, fractional constants yield the same linearizing result for observed counts and rates greater than one. Because fractional constants carry only a small and diminishing effect for values greater than one, it is only in the zero-one range that serious problems arise, which is possibly why the implications discussed here were overlooked by Wang.

Thus for reasons having to do with both our theoretical assumptions about political violence and the symmetry of the logarithmic function, a constant adjustment of 1.0 is anything but arbitrary.²³ In fact, the addition of 1.0 to counts and rates is the only way to preserve what we regard as theoretically meaningful distances in the zero-one range. Furthermore, in our view the theoretical utility associated with a constant of 1.0 substantially outweighs the minor nuisance of bounded negative predictions noted earlier. Accordingly, we do not regard Wang's criticism of OLS estimation to be sufficiently sound to undermine previous findings of the relationship between inequality and political violence. Yet such a relationship would at least come under suspicion if it failed to stand up under alternative estimation methods. Wang claims he has shown this to be the case. It is to this question that we now turn.

EXPONENTIAL POISSON REGRESSION

Wang's second main criticism of conflict studies is equally severe. Because he regards OLS estimates with an added constant as capricious and unreliable, he argues for an alternative estimation method based on the EPR model. He then presents maximum likelihood EPR estimations on raw death counts that diverge sharply from previously published OLS results using logged death rates. His most important finding—that inequality has no impact whatsoever on political violence—has consequences beyond our study (Muller and Seligson 1987), for it appears to undermine a conclusion of several other cross-national analyses as well. Or does it?

The EPR analysis in Wang's Table 2 is termed an "indirect replication" of ours (Muller and Seligson 1987) because it differs on numerous operational and design issues. The most important of these differences is Wang's decision to base his conclusions entirely on a single specification, including two land distribution measures along with income inequality. We believe that this is an inappropriate decision for several reasons. In the first place, the distribution of land and income are not of equal theoretical status. For reasons specified elsewhere, income inequality is theoretically a more proximate cause of political violence than land distribution; therefore, it is the former that merits closest scrutiny (Boswell and Dixon 1990; Muller and Seligson 1987). Wang is right, however, in pointing out that an observed relationship between income inequality and violence could still be the spurious result of land maldistribution affecting both.

TABLE 3

Regression of Political Violence on Income Inequality and Other Explanatory Variables

VARIABLE	MODEL 1 (N = 62)		MODEL 2 (N = 61)	
	NEGBIN	LOLS	NEGBIN	LOLS
Upper 20% income share, ca. 1970	.030* (.014)	.042* (.017)	.036* (.015)	.051* (.019)
Intensity of separatism, 1975	.959* (.230)	1.454* (.393)	1.054* (.389)	1.525* (.397)
Semirepressive regime, 1973–77	.767* (.237)	.874* (.278)	1.087* (.465)	1.136* (.355)
Negative sanctions per 1m, 1975 ^a	.018 (.012)	.475* (.218)	.015 (.016)	.361 (.227)
ln (Population, 1975)	.261* (.069)	—	.213* (.109)	—
Energy consumption per capita, 1970 ^a	—	—	.011 (.007)	.188 (.125)
Constant	.752* (.398)	-1.906* (.826)	.490 (.628)	-3.627* (1.501)
Dispersion parameter $\hat{\sigma}^2$	1,011.3*	—	979.5*	—
Log-likelihood	115,029	—	115,033	—
Adjusted R ²		.45		.46

Note: Main entries are negative binomial regression estimates on raw death counts (NEGBIN) or LOLS estimates on death rates plus 1; entries in parenthesis are standard errors.

^aVariable is logged in LOLS estimation.

* $p \leq .05$, one-tailed test.

Furthermore, while Wang's specification is not without some substantive justification, its inclusion of manifestly irrelevant variables pays a high price in lost efficiency and degrees of freedom. The latter problem is particularly severe in this context since the pattern of missing observations across the income and land measures has the deleterious consequence of reducing the sample size by approximately 25% from what would be available to models employing income inequality alone. Inexplicably, Wang elected not to follow the standard practice of "testing down" or "trimming"—that is, of reestimating his model with the irrelevant variables excluded. We shall show that had he done so, he would have arrived at different conclusions.

Thus both theoretical and methodological considerations lead us to suspect that Wang's EPR estimates may be more an artifact of his specification than results with revolutionary substantive implications as he claims. We therefore set out to replicate Wang's "replication" with a few improvements of our own. Our analyses follow the conventional procedure of beginning with the most causally proximate and theoretically parsimonious specification prior to testing for spurious effects. We omit considerations of negative sanctions for 1968–72 because it has failed to produce significant results in any of our LOLS or EPR estimations (Muller and Seligson 1987, 437). We also depart from Wang's reliance on two-tailed significance tests in favor of the one-tailed tests we originally used (though this is a minor point that turns out to have no bearing on our conclusions). Finally,

because the method of estimation is a point at issue, we provide LOLS estimates alongside EPR results to permit valid comparisons across equivalent specifications.

Our dependent variable for EPR estimation is cumulative deaths from political violence during 1973–77, controlled for population size by the log of 1975 population on the right-hand side. We follow Wang's use of negative binomial regression, since we would expect such deaths to be correlated with one another—an expectation that implies a positive and significant estimate for the dispersion parameter.²⁴ For LOLS estimation, we used cumulative deaths per million 1975 population as the dependent variable and dropped population from the right-hand side.²⁵ Death rates were transformed to logarithms after receiving a constant increment of one. Two other variables on the right-hand side—negative sanctions per million and (in model 2) energy consumption per capita—were also transformed to logarithms for LOLS estimation.

Results from our estimations are presented in Table 3. The initial specification labeled model 1 not only reflects our stated theoretical emphasis on income inequality but also affords the broadest possible cross-national comparison over a sample of 62 countries.²⁶ Examination of negative binomial estimates on the left and LOLS on the right reveals a high degree of convergence in the signs and robustness of estimated effects. Of particular interest is the fact that both estimation methods produce positively signed estimates for the income inequality measure. Because

TABLE 4

Regression of Political Violence on Income Inequality, Land Inequality, and Other Explanatory Variables

VARIABLE	MODEL 3 (N = 83)		MODEL 4 (N = 63)		MODEL 5 (N = 46)	
	NEGBIN	LOLS	NEGBIN	LOLS	NEGBIN	LOLS
Agrarian inequality	.002 (.008)	.009 (.009)	—	—	-.019 (.001)	-.017 (.016)
Landlessness	—	—	.011 (.008)	.011 (.016)	.001 (.374)	-.006 (.021)
Upper 20% income share, ca. 1970	—	—	—	—	.040* (.014)	.064* (.030)
Intensity of separatism, 1975	.704* (.152)	1.720* (.379)	.723* (.046)	1.161* (.347)	.816* (.005)	1.264* (.442)
Semirepressive regime, 1973–77	.462* (.308)	.538* (.323)	.521* (.043)	.516 (.344)	.896* (.005)	.887* (.392)
Negative sanctions per 1m, 1975 ^a	.025* (.015)	.674* (.195)	.018 (.380)	.475* (.202)	.013 (.260)	.329 (.228)
ln (Population, 1975)	.269* (.027)	—	.266* (.113)	—	.256* (.005)	—
Constant	3.638* (.203)	-.380 (.470)	2.057* (.051)	.112 (.291)	1.117* (.005)	-2.149* (1.094)
Dispersion parameter ($\hat{\sigma}^2$)	6,003.9*	—	739.8*	—	897.1*	—
Log-likelihood	694,192	—	87,700	—	82,221	—
Adjusted R ²	—	.35	—	.26	—	.35

Note: Main entries are negative binomial regression estimates on raw death counts (NEGBIN) or LOLS estimates on death rates plus 1; entries in parenthesis are standard errors.

^aVariable is logged in LOLS estimation.

* $p \leq .05$, one-tailed test.

we regard the sign of the coefficient as vital to its interpretation we employ a one-tailed test, yet these inequality estimates would remain significant even by more stringent two-tailed standards. Although both estimates are substantively interpretable as non-linear effects, they are expressed for differently scaled dependent variables and cannot be directly translated to a fully comparable common metric.²⁷

Results from our model 1 specification show that income inequality does contribute to political violence and, contrary to Wang, that this conclusion is unaffected by the method of estimation. However, these estimates alone are not sufficient to identify the source of the discrepancy between Wang's results and our own, nor do they address the possible causal role of land maldistribution. We explore these issues with some additional estimations beginning with model 2 in Table 3. First, we consider Wang's inclusion of energy consumption per capita as a control for economic development. Note that this specification forfeits two degrees of freedom, since the introduction of energy consumption results in the loss of one observation. Results listed in the two columns for model 2 reveal that model 1 estimates are little affected by the inclusion of energy consumption. By most accounts, economic development should be negatively related to political violence; yet both methods reveal positively signed coefficients, though nei-

ther reaches significance even for a two-tailed test.²⁸ Furthermore, Wang's own results show that the effect of energy consumption is virtually nil. Accordingly, energy consumption can be safely ignored in further analyses.

We have yet to examine the role of land distribution. We undertake this task in Table 4 with specifications that include each land measure individually and a specification similar to Wang's that encompasses both land variables and inequality. Model 3 is identical to model 1 except that it substitutes agrarian inequality for income inequality. Agrarian inequality is a composite indicator consisting of the Gini index of land concentration weighted by the size of the agricultural labor force.²⁹ Our estimations are conducted over the broadest possible sample consisting of all 83 countries for which this measure is available. It is evident from columns 1–2 that neither estimation method is able to detect any discernable effect of agrarian inequality on political violence.

Model 4 follows the same procedures, substituting landlessness for agrarian inequality. Landlessness is measured by agricultural households without land as a proportion of the total labor force. While this is not an ideal indicator of land-based discontent, it is nevertheless preferable to measures based on the Gini index (Brockett 1992; Prosterman and Riedinger 1987). Even so, estimations over all 63 countries for which data

are available once again reveal no effects attributable to the maldistribution of land. And once again, the two estimation methods yield comparable results.

We believe that the evidence adduced so far would lead most analysts to conclude that income inequality does promote political violence but that agrarian inequality and landlessness do not. Nevertheless, we examine one final specification that includes all three variables. Because the pattern of missing data across the three measures leaves a sample of only 46 countries, our basis for comparison is unavoidably diminished. Both sets of estimates for model 5 displayed in Table 4 columns 5–6 again show that income inequality contributes to political violence and that agrarian inequality and landlessness are immaterial. This is an especially interesting result for two reasons. First, even though the specification differs only marginally from Wang's by excluding two manifestly irrelevant variables (energy consumption and lagged negative sanctions), the resulting EPR estimates lead to strikingly different conclusions. And second, this estimation corroborates the pattern of results found for the individual estimations in models 1, 3, and 4. Certainly there is no evidence from any of our estimations to support the continued inclusion of land distribution measures. And excluding them takes us back to model 1, which we hypothesized as the proper specification in the first place.

To this point we have focused our attention on Wang's most provocative claim, namely, that EPR estimation overturns the finding of previously published studies showing income inequality to impart a positively accelerated effect on deaths from political violence. We have demonstrated that about the only evidence for this claim is to be found in Wang's lone estimation of what is apparently an idiosyncratic specification encompassing numerous irrelevant variables over a particular subset of cases. Indeed, the weight of present evidence, irrespective of estimation method, clearly supports the original conclusion that income inequality does promote violence. But to say that the two estimation methods concur in the identification of *some* positive inequality effect is not necessarily to say that they identify the *same* effect. Despite the coincidental similarity of the EPR and LOLS estimates displayed in our tables, the size of the effects are not directly comparable since they apply to differently scaled dependent variables under differently conditioned specifications. In order to undertake such a comparison we must depart from the issue as originally posed by Wang to consider directly parallel specifications.

An appropriately parallel specification can be constructed from the elements of our model 1 estimations in Table 3, though this new formulation is not exactly identical to either specification in its original form. Comparability of dependent variables is achieved by redirecting the LOLS specification to an estimation on raw death counts (logged with an increment of 1.0) and adding the logarithm of population to the righthand side.³⁰ With this modification the only remaining difference is in the representation

TABLE 5

Regression of Politically Violent Death Counts on Income Inequality and Other Explanatory Variables (N = 62)

VARIABLE	NEGBIN	LOLS
Upper 20% income share, ca. 1970	.037* (.012)	.070* (.026)
Intensity of separatism, 1975	.892* (.236)	2.004* (.593)
Semirepressive regime, 1973–77	.644* (.225)	1.391* (.484)
ln (Negative sanctions per 1m, 1975)	.360* (.126)	.758* (.322)
ln (Population, 1975)	.332* (.075)	.685* (.160)
Constant	-.474 (.394)	-8.684* (2.201)
Dispersion parameter ($\hat{\sigma}^2$)	964.7*	—
Log-likelihood	115,031	—
Adjusted R ²		.59

Note: Main entries are negative binomial regression estimates on raw death counts. (NEGBIN) or LOLS estimates on death counts plus 1, entries in parenthesis are standard errors.

* $p \leq .05$, one-tailed test.

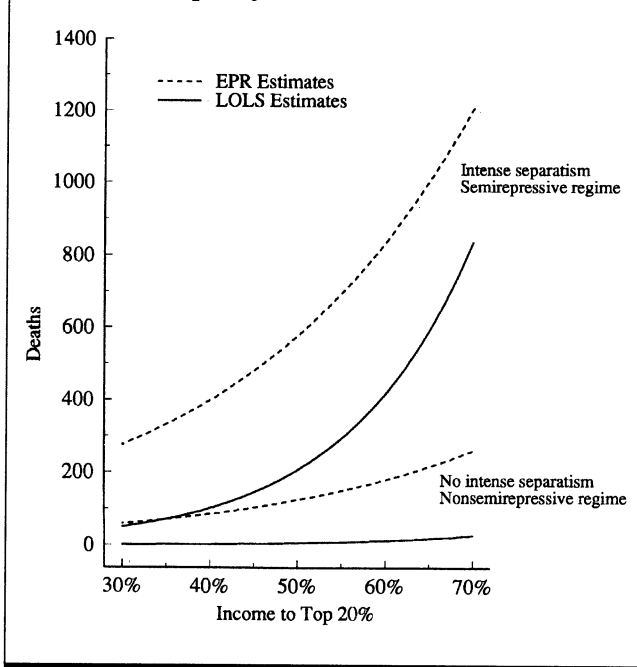
of negative sanctions, which is logged in LOLS and unlogged in EPR. Until now we have followed Wang's procedure of using unlogged sanctions for EPR estimation. Here, however, we depart from Wang by following the more common practice of logging sanctions in both specifications (e.g., Hibbs 1973; Muller and Seligson 1987).

Empirical estimates for this new parallel specification are listed in the two columns of Table 5. There are, of course, both similarities and differences depending on one's point of view. Lacking both fully specified theories and precise measurements over complete sets of observations, contemporary cross-national social scientists seldom hypothesize exact magnitudes of effects. Ordinarily, about all we can ask of our nascent theories and meager data is to identify that there is some effect in the predicted direction within a specified standard of certainty. Sometimes, as in the case of inequality and violence, we can also stipulate an expected functional form. From this conventional perspective the two sets of estimates in Table 5 bear a remarkable similarity. Both methods reveal a positively-accelerated inequality effect on deaths while holding population constant; both identify the important role of separatism and regime repressiveness; and both now show that contemporaneous negative sanctions carry a positive effect under a logarithmic specification.

Beneath the similarities highlighted by conventional social scientific practice are some important differences as well. Most notably, it is apparent that the LOLS estimates achieve systematically larger magnitudes than their EPR counterparts by a ratio of

FIGURE 2

Estimated Inequality Effects on Raw Deaths



roughly two-to-one (except for the constant term, which is larger still). The substantive implications of this size differential are not readily apparent from the reported values themselves since both methods summarize nonlinear functional forms. We can gain at least some limited perspective on these differences by graphing both estimated relationships between income inequality and deaths while holding other variables in the model at some constant value. The mean value is the usual choice of constant for continuous measures and is the value we use for negative sanctions and population. However, the mean is neither informative nor realistic for dichotomous attributes like intense separatism and regime repressiveness, which by our procedures must be either present or absent. Accordingly, we depict each estimated inequality relationship twice, first under conditions that are least likely to produce violence (that is, where intense separatism and semirepressive regime are both absent), and again under the most violent prone conditions (that is, where both are present).³¹

The four graphical depictions are displayed together in Figure 2 with estimated EPR functions represented by broken lines and LOLS functions by solid lines. Two observations seem particularly noteworthy. First, it is evident that EPR and LOLS estimations expose approximately the same functional form under each set of conditions. To be sure, they are not identical—LOLS accelerates somewhat more slowly than EPR under less violent prone conditions and somewhat more rapidly than EPR when violence is more likely. Nevertheless, within the limits of our data they do seem reasonably similar. Second, it is equally evident that LOLS produces systematically

smaller estimates of predicted deaths than EPR.³² The size of this discrepancy varies depending on the level of inequality and the conditions established by separatism and semirepressiveness (though it can be quite large, reaching a high of 420). Clearly this is a difference—the first we have encountered in any of our comparisons—that carries important substantive implications. Given the fact that EPR is specifically tailored to estimations on raw counts, it appears likely that LOLS is underestimating predicted deaths even while delivering a reasonably faithful portrait of the underlying theoretical process.

The results of our analyses have several important substantive and methodological implications for cross-national studies of internal conflict. First and foremost—and contrary to Wang—is the reaffirmation of income inequality as a condition promoting political violence. In addition, our analyses have shown that this inequality effect is not the misspecified result of excluded land maldistribution measures. Indeed, on the evidence at hand, these variables warrant no valid claim to systematic causal relevance. These results are especially impressive because they emerge from quite different estimation methods on different dependent variables. Furthermore, EPR estimation provides some external validation for certain design decisions commonly taken in LOLS estimation. In particular, the EPR results demonstrate that adding a constant 1.0 prior to taking logarithms does not in any way “skew the conclusions,” as Wang would have us believe.

More generally, we found a high degree of convergence in the results from our EPR and LOLS estimations across a variety of specifications. Most encouraging is the fact that each set of results taken alone points to exactly the same conclusions regarding the sources of political violence. We must not generalize too far: the two estimation methods do differ in numerous respects and there is no guarantee that they will deliver such similar results in other contexts. Moreover, when examined over comparable specifications on raw counts, we found substantial discrepancies in the predicted values produced by the two techniques. One advantage often attributed to EPR methods is their less-restrictive assumptions. However, for studies of political violence there is at least one respect in which estimation on raw counts may be somewhat more restrictive in practical applications than estimations on rates. We refer to problems of data collection that may impair accurate accounting of deaths and other violent events.

THE ACCURACY OF POLITICAL VIOLENCE MEASURES

We follow the now-standard practice of measuring political violence based on the number of deaths associated with riots, armed attacks, assassinations, and other politically motivated conflict events. Deaths are preferred to other indicators for a variety

of reasons itemized elsewhere (Muller and Seligson 1987, 448). Like others, we rely on the death reports compiled in the *World Handbook of Political and Social Indicators* (Taylor and Jodice 1983). Recently, however, *World Handbook* records have been criticized by Brockett (1992) as inaccurate due to underreporting of deaths. Specifically, he provides documentary evidence of significant underreporting for three nations in Central America: Guatemala, El Salvador, and Nicaragua. Although blanket charges of inaccuracy without supporting evidence can do little to advance inquiry, we believe that closer inspection of data such as these by country and area specialists is a constructive development. We are therefore grateful for Brockett's efforts and hope that his careful scrutiny of Central America is followed by others just as familiar with other regions of the world.

We concur with Gurr's (1974) observation that minor inaccuracies in political violence data are unlikely to have much affect on the general inferences drawn from cross-national studies. We also concur that "gross and systematic" errors are another matter entirely. All the more troublesome is the fact that errors are most likely to be gross and systematic during periods of large-scale insurgency, revolt, and civil war. Our approach to this problem is a bit more sanguine than Brockett's, though we share his intolerance for faulty data. We believe that knowledge of serious underreporting is better used as a remedy than an indictment. One such remedy would have us eliminate from future analyses any cases plagued by documented irregularities in the reporting of political violence. But to do so would surely introduce equally serious distortions, since these are cases where political violence not only is known to occur but does so at relatively extreme levels. An alternative—indeed, in our view optimal—solution would be to correct inaccurate data points with new and more reliable measurements whenever they become available. But there are drawbacks here, as well, for documented evidence of underreporting does not necessarily imply a precise accounting. Brockett's investigation of deaths in Central America is a case in point. While he does supply convincing evidence of underreporting of deaths, he fails to provide enough detailed information to permit an accurate correction.

There is a third course that will allow us to take account of Brockett's information but not require an exact accounting of cumulative deaths. This is made possible by establishment of a censoring point at some value representing moderately but not excessively high levels of violence. In our previous work we have used 50 deaths per million for this purpose. Because we can reasonably infer from Brockett that the underreported cases exceed this threshold, we can adjust our indicator for these cases by setting it to the maximum level of 50 per million.³³ This will permit us to investigate in at least a provisional way how these particular underreported cases may have affected our earlier estimates. Again we must emphasize that this is not our preferred solution, but it does at least permit us a more informed judgment than

TABLE 6

Regression of Corrected Political Violence on Inequality and Other Explanatory Variables

VARIABLE	MODEL 1A	MODEL 2A	MODEL 6
Upper 20% income share, ca. 1970	.030* (.016)	.032* (.018)	.042* (.016)
ln (Energy consumption per capita, 1970)	—	.102 (.123)	.239 (.109)
Intensity of separatism, 1975	1.045* (.363)	1.063* (.367)	.618* (.332)
Semirepressive regime, 1973–77	1.027* (.298)	1.218* (.340)	.927* (.301)
ln (Negative sanctions per 1m, 1975)	.424* (.210)	.348 (.220)	.357* (.190)
ln (Deaths from pol. viol. per 1m, 1968–72)	—	—	.517* (.113)
Constant	-1.314* (.787)	-2.128* (1.450)	-3.876* (1.310)
Adjusted R ²	.40	.40	.59
N	66	65	65

Note: Main entries are OLS estimates on death rates corrected with information from Brockett, 1992; entries in parenthesis are standard errors.

* $p \leq .05$, one-tailed test.

would otherwise be possible at the present time. Note, too, that because our censoring threshold is expressed as a rate, the adjusted indicator requires estimation by a procedure like OLS. This is not to say that similar solutions could not be implemented for estimating on raw counts, only that in our view the task becomes a much more formidable one.

For this new estimation we employ a slightly larger sample than before, made possible by the inclusion of four additional countries for which we now have reliable measures of income inequality.³⁴ Estimates using these threshold-adjusted death rates are presented in Table 6. Column 1 (model 1A) uses the specification of model 1 in Table 4 while the second (model 2A) includes energy consumption per capita and is comparable to model 2. Although the larger sample and death rate adjustments do bring some slight changes to the estimates, along with a marginal decline in overall fit, the general pattern of results remains very close to those in Table 4. We also present a new estimation (model 4) that includes death rates from 1968–72 as a control, thus transforming this specification into a change model corresponding to our original estimation 1.3 (Muller and Seligson 1987). Change specifications of political vio-

lence have the advantage of embedding models in historical context while providing even more stringent controls than the level specifications examined so far.

Our arguments are summarized by three main conclusions. First, logarithmic transformations on count variables are not arbitrary when accompanied by a proper theoretical understanding of the process being modeled. Second, the method of estimation makes virtually no difference in the general conclusions derived from properly specified model. Finally, making allowances for underreporting of deaths by using the best information available does not significantly alter the results reported in previous studies. As a consequence of this exercise, we are now even more confident that high levels of income inequality promote high levels of political violence.

WILLIAM J. DIXON

EDWARD N. MULLER

University of Arizona

MITCHELL A. SELIGSON

University of Pittsburgh

Notes

Wang would like to thank Chi Huang and Gary King for their guidance and Denis Thornton for his helpful comments. This paper was funded by a grant from the World Society Foundation located in Zurich.

1. For a review of this argument, see also Snyder 1978.
2. For a review of this literature, see Lichbach 1989.
3. "Event counts are variables that have for observation i ($i = 1, \dots, N$) the number of occurrences of an event in a fixed domain" (King 1988, 838). Examples include the numbers of coups, coup attempts, deaths from violence, and so on.
4. For a discussion of these distributions, see Cameron and Trivedi 1986; Hausman, Hall, and Griliches 1984; King 1989b.
5. The most frequently used value for c is 1. For instance, Muller and Seligson indicated that "an increment of one is added to each death score before logging because the log of zero is undefined" (1987, 448, n. 8). Many other studies also added one to the dependent variable before taking a logarithmic transformation (e.g., Boswell and Dixon 1990; Hartman and Hsiao 1988; Muller 1985, 1988b; Park 1986; Weede 1981, 1986).
6. The variables included in this analysis are deaths rate of political violence of 1973–77, upper 20% income share (as a measure of income inequality), energy consumption (as a measure of economic development), intensity of separatism, semirepressive regime, and negative sanctions per one million in 1973–77 and in 1968–72. In addition, there are two variables of land inequality: agrarian inequality and landlessness. For details, see Muller and Seligson 1987. Data are provided by Muller.
7. Least squares estimation often results in predicted event counts that are less than zero. A predicted value that is negative is empirically meaningless in event count analysis. For example, an assertion that the number of deaths in country A between 1973 and 1977 is (-1) is empirically and theoretically impossible.
8. If a prior occurrence of an event actually increases the probability of subsequent occurrence of similar or same event, the data are said to be *overdispersed*. When a prior occurrence

of an event decreases the probability of subsequent events, the data are considered *underdispersed* (King 1989b).

9. In fact, when Gini land concentration is included in the analysis, it is not statistically significant either with or without outliers.

10. Muller and Seligson adopt the same argument (1987, 431). Meanwhile, there are many criticisms toward the use of Gini index as a single measure of land inequality. See, e.g., Brockett 1992; Midlarsky 1988, 1989; Prosterman and Riedinger 1987.

11. These countries are Argentina, Philippine, and Zimbabwe. The accumulated number of deaths from political violence for these three countries between 1973 and 1977 are 4,476, 3,236, and 3,491. The death rates per one million before adjustment are 176.33, 72.82, and 545.12, respectively.

12. The robust standardized residual (ϵ_i^*) is defined as the standardized residual (ϵ_i) divided by the squared root of the estimated dispersion parameter ($\hat{\sigma}^2$), that is, $\epsilon_i^* = \epsilon_i/\hat{\sigma}$. The ϵ_i is defined as the Poisson residual (u_i) divided by its estimated standard deviation: $\epsilon_i = u_i/\sqrt{\hat{\lambda}_i}$, where $u_i = y_i - \hat{\lambda}_i$, and $\hat{\sigma}^2 = (n - p)^{-1} \sum [(y_i - \hat{\lambda}_i)^2/\hat{\lambda}_i]$, where n is the total number of observations and p is the number of independent variables plus one. See McCullagh and Nelder 1989, 200–206; Winkelmann and Zimmermann 1992.

13. Zimbabwe is not included in the following analysis due to missing values on landlessness and agrarian inequality.

14. Cameron and Trivedi suggest that a one-tailed test is appropriate, because it is testing whether α is greater than zero (1986, 43).

15. The results of Table 2, columns 1–2, are generated by the computer program COUNT (King 1991), which employs the NEGBIN I parameterization.

16. The robust standardized residuals for the three countries are 2.58, 3.08, and 2.79, respectively. None of the rest of the robust standardized residuals is greater than 2 in absolute value.

17. For a review of these arguments, see Lichbach 1989; Snyder 1978.

18. The calculation of predicted values will depend on certain assumptions that should be made explicit. For an arbitrarily selected observation, the OLS model can be written as $\ln(Y + c) = X\beta + \epsilon$, where c is a constant added to the dependent variable prior to taking logarithms. Solved for the original dependent variable, Y , this becomes $Y = \exp(X\beta + \epsilon) - c$, where $\exp(\cdot)$ indicates exponentiation. Since predicted values are just estimates of the expected value of the dependent variable, given the independent variables, X , we examine

$$E(Y|X) = E[(\exp(X\beta) \exp(\epsilon) - c)|X].$$

By the usual assumption that X is independent of the error, ϵ , and by properties of conditional expectations, this can be further simplified to

$$E(Y|X) = \exp(X\beta)E(\exp(\epsilon)) - c. \quad (10)$$

Because an exponentiated normal random variable with a zero mean has an expected value of one, if ϵ is assumed to be normally distributed with a zero mean, this expression can be simplified still further to $E(Y|X) = \exp(X\beta) - c$. Thus, whether predicted values are estimated by this last expression or equation 10 depends on assumptions about the error term in the original model. For simplicity of exposition here; and for our OLS estimations on death rates later, we assume a normally distributed error with zero mean. Although this is a common assumption in OLS, it is not always an appropriate one and may be particularly questionable for some models on raw counts.

19. Note that in his Table 1, column 1, Wang correctly reports an OLS estimate of .98 for semirepressive regime. This value was misprinted in Muller and Seligson 1987, equation 1.3: the coefficient of 1.23 on intensity of separatism was repeated for regime repressiveness, the variable below it in the equation.

20. The addition of an increment of 1.0 before logging has been the standard procedure in analysis of political conflict

event counts since such data began to be collected for large numbers of countries in the mid-1960s. (A few prominent examples prior to those cited by Wang for the 1980s are Russett et al. 1964; Gurr 1968; Hibbs 1973; and Snyder and Kelly 1977). Tufté describes the addition of 1.0 prior to logging as "the most common procedure" for overcoming the problem of zero values (1974, 109).

21. In the interest of space, we examine only this single bivariate relationship, though it should be understood that similar effects will occur for other variables in a multivariate model. Thus, a full picture of these effects would encompass additional bivariate plots, as well as a series of partial plots controlling for other variables appearing in the model.

22. This step-level representation may appear vaguely similar to what are sometimes called sample selection or hurdle models (King 1989a). In fact, they are quite different. A hurdle model permits the causal process accounting for a first death to differ from the process responsible for the accumulation of additional deaths. A step-level representation, on the other hand, would require a single causal process to account for a very large interval at the first death, a relatively small interval at the second, and progressively larger intervals thereafter.

23. Of course, the amount of distortion introduced by a fractional constant is necessarily a matter of degree. There is little difference between a constant of 1.0 and, say, .999. Thus, in a strictly technical sense there is still some arbitrariness in choosing among all the infinitesimally different values in the near vicinity of 1.0. However this amounts to picking extremely fine nits, since no one would claim these to make any difference in estimation outcomes.

24. To ensure that our procedures followed Wang as closely as possible, we used the same data and the same computational software. We are reasonably confident of these procedures, since we were able to reproduce all parameter estimates in his Table 2, column 1, to at least the second significant digit. Residual differences may be due to different treatments of the energy consumption per capita measure. We found its magnitudes too large in relation to other regressors to permit estimation, so we moved its decimal two places to the left. We are uncertain how Wang treated this problem.

25. Our previous estimations have censored death rates above 50 per million at 50 in acknowledgement of the inaccuracies of death reports in highly violent societies and also to reduce the influence of extreme cases (Boswell and Dixon 1990; Muller and Seligson 1987). Although we followed this practice in constructing the Figure 1 plots, we do not do so here in order to enhance the comparability between EPR and LOLS estimations. We shall return to this procedure and its justification.

26. The sample size is limited by the availability of information for income inequality. We shall use new information to increase the sample size to 66. Here, however, we rely only on the 62 countries available to Wang.

27. We consider a more directly comparable specification shortly.

28. Our collective experience with different measures of political violence and of economic development suggests that significant positive effects emerge only in change models that include lagged measures of violence (see Table 6 below, and Boswell and Dixon 1990; Muller and Seligson 1987). For an explanation of this rather perplexing result and an alternative formulation that eliminates it, see Boswell and Dixon 1993.

29. Wang correctly points out that the Gini index has been widely criticized and we are not persuaded that this weighting operation adequately corrects its flaws (Brockett 1992; Midlarsky 1988, 1989; Muller, Seligson, and Fu 1989; Prosterman and Riedinger 1987). Accordingly, our analysis of this variable should not be interpreted as an endorsement of its validity.

30. It would be equally revealing to compare EPR and LOLS using death rates, our preferred dependent variable. However, all of the EPR algorithms with which we are familiar are constrained to dependent variables expressed only in integer values.

31. The estimated EPR relationships are defined by

$$\hat{Y} = \exp(.892S + .644R + .360N + .322P - .474 + .037I),$$

where S and R are both valued at 1 or 0 to symbolize the presence or absence of separatism and semirepressiveness, respectively; N and P are the means of log transformed negative sanctions and population; and I varies between 30 and 70 which is the observed range of inequality. Using the same notation, the LOLS functions are given by

$$\hat{Y} = \exp(2.004S + .1391R + .758N + .685P - 8.684 + .071I) - 1,$$

where 1 is an adjustment for the constant increment.

32. This may appear puzzling given that LOLS was found to return systematically larger estimates than EPR. The explanation is in the large *negative* LOLS constant which has the effect of adjusting the predicted values downward.

33. The choice of 50 is somewhat arbitrary. Our experiments with different censoring points shows that the value chosen makes little difference in the results. However, the implementation of a censoring point necessarily introduces some bias that will reduce the size of subsequent parameter estimates.

34. These four countries are Austria, Guatemala, Iran, and Senegal. A summary measure of inequality for Austria was obtained from Simpson (1990); measures for the others are from Muller 1988a. All other inequality measures used are reported in Muller and Seligson 1987.

References

- Boswell, T., and W. J. Dixon. 1990. "Dependency and Rebellion: A Cross-National Analysis." *American Sociological Review* 55:540-559.
- Boswell, Terry, and William J. Dixon. 1993. "Marx's Theory of Rebellion: A Cross-National Analysis of Class Exploitation, Economic Development, and Violent Revolt." *American Sociological Review* 58:forthcoming.
- Brockett, Charles D. 1992. "Measuring Political Violence and Land Inequality in Central America." *American Political Science Review* 86:169-176.
- Cameron, A. Colin, and Pravin K. Trivedi. 1986. "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests." *Journal of Applied Econometrics* 1:29-53.
- Eckstein, Harry. 1980. "Theoretical Approaches to Explaining Collective Political Violence." In Ted R. Gurr, ed. *Handbook of Political Conflict*. New York: The Free Press:135-166.
- Green, Jerrold D. 1982. *Revolution in Iran*. New York: Praeger.
- Green, Jerrold D. 1984. "Countermobilization as a Revolutionary Form." *Comparative Politics* 16:153-169.
- Gurr, Ted Robert. 1968. "A Causal Model of Civil Strife: A Comparative Analysis Using New Indices." *American Political Science Review* 62:1104-1124.
- Gurr, Ted R. 1970. *Why Men Rebel*. Princeton: Princeton University Press.
- Gurr, Ted Robert. 1974. "The Neo-Alexandrians: A Review Essay on Data Handbooks in Political Science." *American Political Science Review* 68:243-52.
- Hall, B. H., Z. Griliches, and J. A. Hausman. 1986. "Patents and R and D: Is There a Lag?" *International Economic Review* 27:265-283.
- Hartman, John, and Wey Hsiao. 1988. "Inequality and Violence: Issues of Theory and Measurement in Muller." *American Sociological Review* 53:794-799.
- Hausman, J., B. H. Hall, and Z. Griliches. 1984. "Econometric Models for Count Data with an Application to the Patents-R & D Relationship." *Econometrica* 52:909-938.
- Hibbs, Douglas A. 1973. *Mass Political Violence*. New York: Wiley.
- King, Gary. 1987. "Presidential Appointments to the Supreme Court: Adding Systematic Explanation to Probabilistic Description." *American Politics Quarterly* 15:373-386.

- King, Gary. 1988. "Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for the Exponential Poisson Regression Model." *American Journal of Political Science* 32:838-863.
- King, Gary. 1989a. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. N.Y.: Cambridge University Press.
- King, Gary. 1989b. "Variance Specification in Event Count Models: From Restrictive Assumptions to a Generalized Estimator." *American Journal of Political Science* 33:762-784.
- King, Gary. 1991. *COUNT: A Program for Estimating Event Count and Duration Regressions*. Version 3.0. Cambridge: Harvard University.
- Lichbach, M. I. 1989. "An Evaluation of 'Does Economic Inequality Breed Political Conflict' Studies." *World Politics* 41:431-470.
- McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*. 2nd edition. New York: Chapman and Hall.
- Midlarsky, Manus I. 1988. "Rulers and the Ruled: Patterned Inequality and the Onset of Mass Political Violence." *American Political Science Review* 82:491-509.
- Midlarsky, Manus I. 1989. "Land Inequality and Political Violence." *American Political Science Review* 83:587-595.
- Muller, Edward N. 1985. "Income Inequality, Regime Repressiveness and Political Violence." *American Sociological Review* 50:47-61.
- Muller, Edward N. 1988a. "Democracy, Economic Development, and Income Inequality." *American Sociological Review* 53:50-68.
- Muller, Edward N. 1988b. "Inequality, Repression, and Violence: Issues of Theory and Research Design." *American Sociological Review* 53:799-806.
- Muller, Edward N., and Mitchell A. Seligson. 1987. "Inequality and Insurgency." *American Political Science Review* 81:425-451.
- Muller, Edward N., Mitchell A. Seligson, and Hung-Der Fu. 1989. "Land Inequality and Political Violence." *American Political Science Review* 83:577-86.
- Park, Kang. 1986. "Reexamination of the Linkage between Income Inequality and Political Violence." *Journal of Political and Military Sociology* 14:185-197.
- Pedhazur, Elazar J. 1982. *Multiple Regression in Behavioral Research*. 2nd edition. New York: Holt, Rinehart and Winston.
- Prosterman, Roy L. 1976. "IRI: A Simplified Predictive Index of Rural Instability." *Comparative Politics* 8:339-354.
- Prosterman, Roy L., and Jeffrey M. Riedinger. 1987. *Land Reform and Democratic Development*. Baltimore: The John Hopkins University Press.
- Russett, Bruce M., Hayward R. Alker, Jr., Karl W. Deutsch, and Harold D. Lasswell. 1964. *World Handbook of Political and Social Indicators*. New Haven: Yale University Press.
- Simpson, Miles. 1990. "Political Rights and Income Inequality: A Cross-National Test." *American Sociological Review* 55:682-93.
- Snyder, David. 1978. "Collective Violence: A Research Agenda and Some Strategic Considerations." *Journal of Conflict Resolution* 22:499-534.
- Snyder, David, and William R. Kelly. 1977. "Conflict Intensity, Media Sensitivity, and the Validity of Newspaper Data." *American Sociological Review* 42:105-123.
- Taylor, Charles L. and David A. Jodice. 1983. *World Handbook of Political and Social Indicators* 3rd. ed. Vols. 1 and 2. New Haven: Yale University Press.
- Tufte, Edward R. 1974. *Data Analysis for Politics and Policy*. Englewood Cliffs, NJ: Prentice-Hall.
- Weede, Erich. 1981. "Income Inequality, Average Income, and Domestic Violence." *Journal of Conflict Resolution* 25:639-654.
- Weede, Erich. 1986. "Income Inequality and Political Violence Reconsidered." *American Sociological Review* 51:438-441.
- Winkelmann, R., and Klaus F. Zimmermann. 1992. "Robust Poisson Regression." In L. Fahrmeir, B. Francis, R. Gilchrist, and G. Tutz. eds. *Advances in GLIM and Statistical Modelling*. New York: Springer-Verlag.

NOTES FROM THE MANAGING EDITOR

FORTHCOMING IN MARCH, 1994

The following articles, research notes and controversies have been scheduled for publication in a forthcoming issue.

Lucius J. Barker. "1993 Presidential Address. Limits of Political Strategy: A Systemic View of the African American Experience."

Barry Ames. "The Reverse Coattails Effect: Local Party Organization in the 1989 Brazilian Presidential Election."

David P. Baron. "Electoral Competition with Informed and Uninformed Voters."

Robert C. Bartlett. "Aristotle's Science of the Best Regime."

William J. Dixon. "Democracy and the Peaceful Settlement of International Conflict."

Michael W. Giles and Kaenan Hertz. "Racial Threat and Partisan Identification."

Jeffrey C. Isaac. "Oases in the Desert: Hannah Arendt on Democratic Politics."

David C. King. "The Nature of Congressional Committee Jurisdictions."

Arthur Lupia. "Shortcuts versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections."

Walter R. Mebane, Jr. "Fiscal Constraints and Electoral Manipulation in American Social Welfare."

Richard Nadeau, Richard G. Niemi, and Timothy Amato. "Expectations and Preferences in British General Elections."

Joshua Parens. "Multiculturalism and the Problem of Particularism."

Kaare Strøm. "The Presthus Debacle: Intra Party Politics and Bargaining Failure in Norway."

George Tsebelis. "The Power of the European Parliament as a Conditional Agenda Setter."

Jean-Pierre Benoit and Lewis A. Kornhauser. "Social Choice in a Representative Democracy." A Research Note.

Wendy Rahn, John H. Aldrich, and Eugene Borgida. "Individual and Contextual Variations in Political Candidate Appraisal." A Research Note.

ERRATUM

In "Messages Received: The Political Impact of Media Exposure" by Larry M. Bartels, June 1993, page 270, the graphs contained in Figure 1, Summary of Estimated Effects of Television News Exposure, and Figure 2, Summary of Estimated Effects of Newspaper Exposure, were reversed.