AMERICAN SOCIOLOGICAL REVIEW

OFFICIAL JOURNAL OF THE AMERICAN SOCIOLOGICAL ASSOCIATION

ONLINE SUPPLEMENT
to article in

AMERICAN SOCIOLOGICAL REVIEW, 2006, VOL. 71 (FEBRUARY:95–122)

Deterring Delinquents: A Rational Choice Model of Theft and Violence

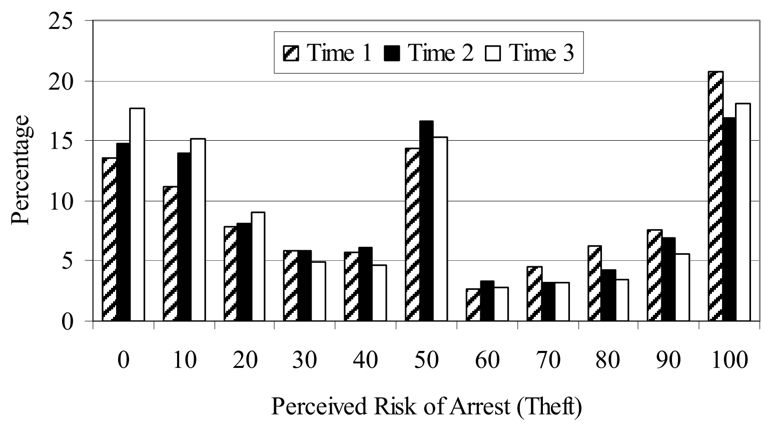
Ross L. Matsueda
University of Washington

Derek Kreager
University of Washington

David Huizinga
University of Colorado

A. TEMPORAL PATTERNS OF PERCEIVED CERTAINTY

Figure S1 graphs the univariate distributions of perceived certainty over time for theft and violence, respectively. Perceived risk appears on the x-axis, ranging from zero certainty to 100. Percentages of cases appear on the y-axis. The graph reveals floor and ceiling effects, as observations tend to clump at zero and 100 (suggesting the need for a censored regression model), and non-normal distributions over the range of risk (suggesting the need for sensitivity tests). The graphs also reveal that perceived risk tends to decline over time. The percentage of people scoring low on risk (e.g., 0–20) increases over time, moving from time 1 (striped bars) to time 3 (solid white bars). Conversely, the percentage of people scoring high on risk (e.g., 90–100) increases over time, again moving from time 1 to time 3. When we control for age, however, the decline is statistically significant between time 1 and time 2, but not between time 2 and time 3.



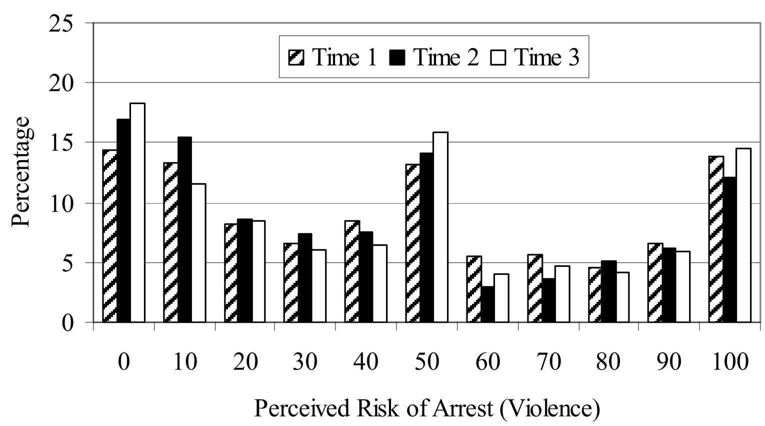


Figure S1. Perceived Risk of Arrest for Theft (*top*) and Violence (*bottom*) by Time Period

B. FIT OF POISSON VERSUS NEGATIVE BINOMIAL MODELS

Figure S2 graphs the proportions of cases by counts from the observed data along with mean predicted probabilities by counts for Poisson and negative binomial models. For both theft and violence, the Poisson regression model underestimates the observed number of zeros and overestimates the middle range, between one and four. In contrast, the negative binomial tracks the observed data relatively well, as reflected in likelihood-ratio test statistics.

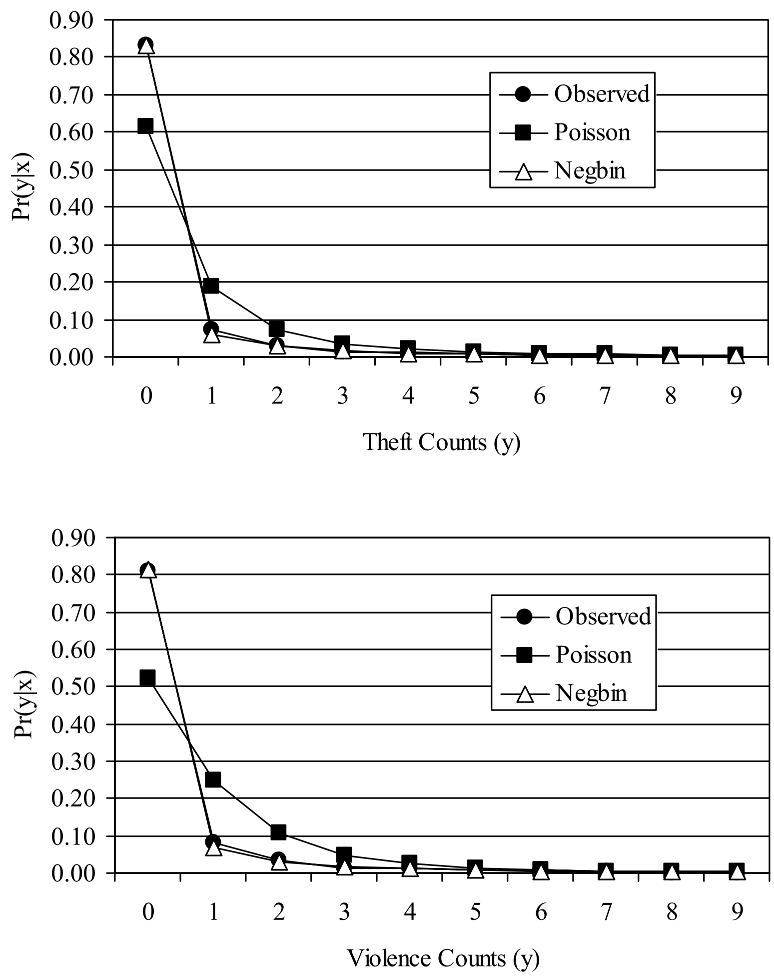


Figure S2. Mean Predicted Probabilities of Theft (*top*) and Violence (*bottom*) from Poisson and Negative Binomial Models (with Observed Values)

C. DETAILS OF THE RANDOM EFFECTS NEGATIVE BINOMIAL MODEL

If we let y_{it} be the count of self-reported acts of theft or violence for person i at time t, we can specify that y_{it} follows a Poisson distribution with parameter λ_{it}. Hausman et al. (1984) assume that λ_{it}, in turn, follows a gamma distribution with shape parameters (γ_{it}, δ_i), which produces the negative binomial distribution for y_{it}. We parameterize γ_{it} in the usual way as an exponential function of explanatory variables, $\gamma_{it} = e^{x_{it}\beta}$, where x_{it} is a vector of our predictor variables and β is a vector of coefficients. To introduce variation across individuals, Hausman et al. (1984) assume that each individual has its own δ_i, which are randomly distributed across individuals. They then assume that the ratio, $\delta_i/(1+\delta_i)$, is distributed as a beta random variable with shape parameters (a, b). Then the joint probability of the counts for the *i-th* individual is

$$\Pr(y_{i1},\ldots y_{iT} \mid x_{i1},\ldots,x_{iT}) = \frac{\Gamma(a+b)\Gamma(a+\sum_{t=1}^{T}\gamma_{it})\Gamma(b+\sum_{t=1}^{T}y_{it})}{\Gamma(a)\Gamma(b)\Gamma(a+b+\sum_{t=1}^{T}\gamma_{it}+\sum_{t=1}^{T}y_{it})}\prod_{t=1}^{T}\frac{\Gamma(\gamma_{it}+y_{it})}{\Gamma(\gamma_{it})\Gamma(y_{it}+1)}$$

In essence, this model allows for overdispersion in the Poisson model with the inclusion of δ_i, and then layers a random individual effect onto the negative binomial model by assuming $\delta_i/(1+\delta_i)$ follows a beta distribution with parameters a and b. The mean of $\delta_i/(1+\delta_i)$ is $E[\delta_i/(1+\delta_i)] = a/(a+b)$, while the variance is $V[(\delta_i/(1+\delta_i)] = ab/(a+b+1)(a+b)\cdot$ Finally, note that, unlike the random effects Poisson model, this model allows λ_{it} to vary across individuals and time (even if the x_{it} 's are constant) because it is a realization from a gamma distribution each year (for details, see Hausman et al. 1984).

D. SENSITIVITY ANALYSIS: EXAMINING ROBUSTNESS OF THE MODELS

Our statistical models make a number of strong assumptions about the population from which our data were drawn. Because violations of one or more assumptions could lead to misleading results, we examine the sensitivity of our results to such violations. Our strategy is to conduct, when possible, a statistical test on the assumption, and also examine whether estimates and conclusions change when the assumption is relaxed. In most cases, our large sample size results in powerful tests, and we can detect substantively trivial departures from assumptions, and the important question becomes how much of a substantive difference does a given departure make?

RANDOM EFFECTS TOBIT MODELS OF PERCEIVED CERTAINTY

In general, the tobit model is not robust to violations of assumptions, typically leading to biased and inconsistent estimates. We consider seven assumptions in turn. First, if the assumption of a homoscedastic error term is violated, estimates of coefficients can be biased and inconsistent (Greene 2003). To address this issue, we estimated a general heteroscedastic tobit model (without random effects), in which $\sigma_i^2 = \sigma^2 e^{\alpha' w_i}$, and used a likelihood ratio test of the null hypothesis of homoscedasticity ($\alpha = 0$) (see Greene 2003, p. 769; Greene 2002). The test rejects homoscedasticity for our perceived risk models for both theft ($\chi^2 = 178.62$; df = 24; p < .001) and violence ($\chi^2 = 106.06$; df = 23; p < .001).[1] The coefficient estimates of the heteroscedasticity model, however, were virtually identical to our reported results, leaving our substantive conclusions unchanged.

Second, if errors are not normally distributed, coefficient estimates will be biased and inconsistent. We used Pagan and Vella's (1989) conditional moment test, which tests whether third- or fourth-order moments (conditional on the model) depart from zero (Greene 2003, p. 773). The test clearly rejects the

[1] The heteroscedastic model for violence excludes the Hispanic variable, which has no effect and due to collinearity caused problems with estimation.

null hypothesis of normality for both theft (χ^2 = 6,411; df = 2; p < .001) and violence (χ^2 = 13,480; df = 2; p < .001). Therefore, we followed Greene's (2002) strategy for estimating a tobit model with non-normal disturbances by transforming the dependent variable and estimating a survival model, which has the identical likelihood function as the tobit model with censoring from below (where our censoring is acute), under alternate specifications of the disturbance distributions. We find that for certainty of arrest for theft, the gamma fits slightly better than the Wiebull and logistic, and each provides substantive results similar to our tobit model. For violence, the logistic fit better than the Wiebull, and again provides results similar to our tobit models. Thus, our substantive story appears robust to different specifications of the disturbance distribution.

Third, the tobit model assumes that the process producing censoring is identical to that producing the mean outcome. This assumption implies a proportionality constraint on coefficients of Cragg's (1971) specification of a probit model of censoring and a truncated regression model of the mean outcome (Greene 2003, p. 770). We use a Lagrangian multiplier test of this constraint and reject the tobit constraint for both theft (χ^2 = 114.35; df = 24; p < .001) and violence (χ^2 = 100.14; df = 24; p < .001). Inspection of the results reveals that for the probit models, neither the arrest ratios nor unsanctioned offenses influenced censoring, whereas in the truncated regression model of perceived risk, the coefficients resemble those of the tobit model. In fact, the major difference between the two is that the coefficients of the truncated regression are slightly larger and more significant. Thus our model appears robust to the non-proportionality, and our tobit coefficients appear to be conservative estimates.

Fourth, our tobit models of certainty of arrest include observed lagged certainty, y_{t-1}, as a right-hand side variable. It is well-known that the estimates of a lagged endogenous variable will be biased in the presence of serial correlation, and such biases could affect other estimates as well. Our random-effects models correct for a version of serial correlation in which all lags are equally-correlated, which could be the wrong autocorrelation structure. Therefore, we used the second-order lagged certainty, y_{t-2}, as an instrumental variable, and corrected for serial correlation in a two-wave panel model. Our substantive results remained the same, with estimated effects becoming slightly larger.

Fifth, our tobit models of certainty of arrest include observed lagged certainty, y_{t-1}, as a right-hand side variable. It may, however, make sense to control for Y^*_{t-1}, the uncensored latent lagged certainty, rather than y_{t-1}, as the latter may under-control for prior arrest (Wooldridge 2002, page 543). Therefore, we re-estimated our model first estimating the data-censoring mechanism for each time period to obtain $\hat{Y}^*_{t-1}$ and then introducing $\hat{Y}^*_{t-1}$ as a lagged endogenous predictor into the random effects tobit model. This entails using our wave three data to predict wave four and wave four to predict wave five, and therefore, leaves us with two waves. We compared the resulting two-wave model to the corresponding two-wave version of the model we report. We find that for both theft and violence arrest certainty, some coefficients are slightly attenuated in the $\hat{Y}^*_{t-1}$ model, but overall the pattern of results are similar across the two models.

Sixth, our random effects tobit models assume that individual heterogeneity is orthogonal to our time-varying regressors; violation of this assumption could lead to biased and inconsistent estimates. We estimated fixed effects tobit models censored from below (using LIMDEP 8.0) and OLS models (with and without lagged violence or theft), and found similar overall results. Such models have larger standard errors and are unable to estimate effects of stable characteristics.

RANDOM EFFECTS NEGATIVE BINOMIAL MODELS OF VIOLENCE AND THEFT

We examined two assumptions of our random effects negative binomial model. First, our model assumes it adequately accounts for the large number of zeros in our dependent variables of theft and violence counts. Although our random effects negative binomial models appear to do well in this regard, we nevertheless tried estimating a zero-inflated random effects model, which explicitly models the zero counts by changing the model's mean structure. For example, in our models, zero counts could be produced by an absence of structural opportunity to commit theft or violence, or by decisions to refrain from crime produced by our rational choice process. We were unable to obtain stable estimates of such a model. This is perhaps not surprising, given that such a model requires the data to parse individual heterogeneity into components reflecting stable covariates, lagged endogenous regressors, zero inflation, over-dispersion, and random intercepts. We are not strongly interested in disentangling each component, and therefore, report our random effects negative binomial results.

Second, our model assumes that the problem of serial correlated errors in the presence of a lagged dependent variable is adequately modeled by our random effects, which imply an equal-correlated lagged error structure (regardless of lag length). Again, we estimated a two-wave model using the second-order lagged theft or violence as an instrumental variable. We found that the overall story remained unchanged, although parameter estimates were slightly larger.

We were unable to examine the assumption that individual heterogeneity is orthogonal to our time-varying regressors. Because we lose those cases in which crime does not change appreciably, our dependent variables contain an abundance of zeros, and we include endogenous lagged predictors, the data were not up to providing stable estimates of coefficients in a fixed-effects version of our negative binomial model.