

Effect of Punishment Regimes on Crime Rates

Keywords: multiple regression, best subsets, causality.

Description

Criminologists are interested in the effect of punishment regimes on crime rates. This has been studied using aggregate data on 47 states of the USA for 1960. The data set contains the following columns:

Variable	Description
M	percentage of males aged 14–24 in total state population
So	indicator variable for a southern state
Ed	mean years of schooling of the population aged 25 years or over
Po1	per capita expenditure on police protection in 1960
Po2	per capita expenditure on police protection in 1959
LF	labour force participation rate of civilian urban males in the age-group 14-24
M.F	number of males per 100 females
Pop	state population in 1960 in hundred thousands
NW	percentage of nonwhites in the population
U1	unemployment rate of urban males 14–24
U2	unemployment rate of urban males 35–39
Wealth	wealth: median value of transferable assets or family income
Ineq	income inequality: percentage of families earning below half the median income
Prob	probability of imprisonment: ratio of number of commitments to number of offenses
Time	average time in months served by offenders in state prisons before their first release
Crime	crime rate: number of offenses per 100,000 population in 1960

Download

[Data File](#) (tab-delimited text)

Source

Ehrlich, I. (1973) Participation in illegitimate activities: a theoretical and empirical investigation. *Journal of Political Economy* 81, 521–565.

Vandaele, W. (1978) Participation in illegitimate activities: Ehrlich revisited. In *Deterrence and Incapacitation*, eds A. Blumstein, J. Cohen and D. Nagin, National Academy of Sciences, Washington DC, pp. 270–335.

Venables, W., and Ripley, B. (1998). *Modern Applied Statistics with S-Plus, Second Edition*. Springer-Verlag.

The data given here is rounded data taken from Vandaele (1978). The column scales differ somewhat from Venables and Ripley (1998). The data was originally collected by Ehrlich from the *Uniform Crime Report* of the FBI and other US government sources.

Analysis

- Only one of Po1 and Po2, and only one of U1 and U2, remain in the final regression, because of high collinearity.
- Data gives association not causal relationships. For example, does crime really increase with police expenditure?
- Crime is negatively associated with probability of imprisonment.
- Crime is slightly better modelled on a log scale.

Example Analysis

This analysis uses R 2.6.0 <http://www.r-project.org>

```
> UScrime <- read.delim("http://www.statsci.org/data/general/uscrime.txt")
> lm.crime <- lm(Crime~., data=UScrime)
> summary(lm.crime,cor=F)
> library(leaps)
> leaps.crime <- leaps(UScrime[,1:15],UScrime$Crime,nbest=2)
> leaps.tab <- data.frame(p=leaps.crime$size,Cp=leaps.crime$Cp)
> round(leaps.tab,2)
> lm.crime <- lm(Crime~M+Ed+Po1+U2+Ineq+Prob,data=UScrime)
> summary(lm.crime)
```