WHAT DOES THE WORLD SPEND ON CRIMINAL JUSTICE?

Graham Farrell and Ken Clark

WHAT DOES THE WORLD SPEND ON CRIMINAL JUSTICE?

Graham Farrell and Ken Clark

Graham Farrell is Professor of Criminology and Director of the Midlands Centre
for Criminology at Loughborough University.
Ken Clark is Senior Lecturer in the School of Economic Sciences at Manchester University.

http://www.heuni.fi

Contents

Acknowledgements

Summary

This study produces an estimate of the global costs of public expenditure on criminal justice. The global estimate is an extrapolation from data provided by the governments of seventy countries. At the country-level there is a strong relationship between the level of available public money and expenditure upon public policing, courts, prosecution and prisons. The relationship is explored using six regression models, and criminal justice expenditure in other countries is estimated using the best models. Global criminal justice expenditure in 1997 is estimated at $360 billion (the equivalent of $424 billion in 2004 prices), of which 62 percent was spent on policing, 3 percent on prosecutions, 18 percent on courts, and 17 percent on prisons. As the first systematic empirical estimate of global criminal justice expenditure, it is hoped that the present research may spur better data collection practices and further research.

Introduction

Knowledge relating to the overall well-being of the planet, and cross-national variation within it, is increasingly used to inform national and international policy. Some global indicators are counts, their attraction being that they were previously unknown: In 1999, the earth's human population exceeded six billion for the first time - a psychological watershed if quantity measures humanity's success. Countries are frequently ranked on key socio-economic indicators such as literacy, mortality and gross domestic product. Composite indicators such as the Human Development Index (see e.g. United Nations Development Programme 2002) and the Corruption Perceptions Index (Transparency International 2002) are derived. Such indicators serve as benchmarks against which to measure variations in time and place. Their political importance is such that they can provoke investigation of good and bad practice in countries that are statistically deviant. Yet, while the global village requires global indicators, they are relatively few and far between, and still often at a fledgling stage of development. Further, while there is a growing body of research highlighting the importance of estimating the costs of crime (see Cohen 2000 for a review), there is little, if any, on aspects of the global cost of criminal justice. It is in this context that the present study produces what we believe to be the first systematic empirical estimate of the direct cost of public expenditure on the global criminal justice system.

Previous research has identified a strong relationship between a country's economic welfare, measured as GDP, and its expenditure on criminal justice (Newman and Howard 1999). This study continues work begun in Farrell et al. (2001) which examined only expenditure on policing. Not surprisingly, on average, richer countries spend more per capita on criminal justice than poor countries. In this study, the relationship between GDP and spending on criminal justice is examined using data for seventy countries. To examine the extent to which GDP can be used to predict criminal justice expenditure, six regression models are developed. The 'best' model of the relationship is identified for each of policing, prosecution, courts and prisons.

These models are then used to predict criminal justice expenditure in other countries, which are summed to produce a global estimate. It should be noted at the outset that neither expenditure on privatised nor informal community aspects of criminal justice are included in the estimates.

The bulk of this study is concerned with ensuring transparency of method with regard to data preparation and statistical technique. A summary of the key finding is as follows: It is estimated that the world spent $360 billion on criminal justice in 1997. Of this total, 62 percent ($222.5 billion) was spent on public policing[1], 3 percent ($11.2 billion) on prosecutions, 18 percent ($63.5 billion) on courts, and 17 percent ($62.5 billion) on prisons.

What is the point of estimating global criminal justice expenditure? Primarily, the estimate is one component of the costs of crime that has not been estimated to date. Estimating the costs of crime is an increasingly important area of criminal justice research. While such an estimate may prove useful for future research and analyses, it is also necessary to be cautious about the potential abuse of such estimates. Mark Cohen notes some of the benefits while warning of the potential for abuse of aggregate cost estimates:

> "[W]hat are we to do with this information? If we are successful in fully estimating the cost of crime we can compare this total cost estimate with that of other social problems (e.g. cancer, auto crashes, homelessness). Whether one agrees that this is a useful exercise or not, various advocacy groups do compare 'costs of crime' estimates with the cost of other social ills in an effort to affect policy decisions. Unfortunately, misuses of these data occur on both sides of the political debate." (Cohen, 2000; 269)

While it is possible that any research findings can be wilfully abused, this should not become a reason to hinder the accumulation of knowledge. However, the possibility does highlight the need for methodological transparency, and this is one of the key facets of the present study. In addition, the findings of this study are presented with a range of cautions and caveats.

To our knowledge there is only one previous research study with a primary focus upon the expenditure data from the UN Crime Survey. Jon Spencer (1993) examined trends over time in expenditure between 1982 and 1986. He found expenditure had increased generally and that changing levels of expenditure tended to be correlated between stages of the criminal justice system. If spending on one stage increased then, generally speaking, it increased for the other stages. Spencer concluded that "This analysis suggests that governments allocate resources to criminal justice with little or no attention to outcomes" (p.1). Spencer's analysis was of trends and did not include any estimates of spending for countries where the data was not available.

The remainder of this study is structured as follows. The next section describes data sources and data preparation. The regression model analysis and findings are then presented. The discussion section suggests uses of the present research, makes suggestions for future research, and is followed by the conclusion.

2. Data Sources and Preparation

2.1 The UN Crime Survey

This section discusses survey methodology and data quality control. The primary data source was reports by national governments to the sixth United Nations Survey of Crime Trends and Criminal Justice Systems (hereafter the UN Crime Survey), for which the most recent data relates to 1997. The release of the survey used here was that of 27 June 2001. Survey responses had been received from eighty-three countries of which seventy provided full or partial information on criminal justice expenditures.

Methodological issues have been extensively explored in the literature relating to the UN Crime Survey. Cross-national differences reflect differences in the composition of criminal justice systems, variation in definitions and terms, as well as differences in the extent, capacity, logistical difficulty inherent to, and emphasis upon, national-level data reporting and recording practices.

Countries differ in purchasing power, so they can buy different amounts for the same dollar price. Differences in legal definitions may mean that, for example, country A might include military police as part of policing whereas country B does not. Such definitional differences may or may not be identifiable depending on whether they are reflected in the terminology of the report from a government. There are also differences in practice: Country A may prosecute all offences whereas country B does not. Procedural differences may produce spending differences if, for example, country A defined prosecution as only those cases that reach court, whereas country B included warnings issued by the prosecutor among its prosecutions. Counting rules and statistical classification are a source of variation as some countries may include minor local expenditures in national estimates whereas others do not. Certain factors remain unknown relating to informal and other justice procedures, such as whether the costs of Sharia courts are included where they exist.[2] Similarly, where there is a division between central and local government, it is not always clear whether local activities are included in national budgets. There is also the possibility of error during the various stages of reporting and recording the data: A country cannot accurately report expenditure levels if the data does not exist, if national data collection is not coordinated, if data are out of date, if the survey went to the wrong government department or respondent, or if clerical errors were made in the completion of the survey form. Finally, the data is fragmentary in part because some countries simply do not respond to the survey while others return it incomplete without explanation. It has been suggested that response rates and completion rates could reflect secrecy relating to issues of national security issues, or various aspects of international politics, and all of these issues have been extensively reviewed elsewhere (Nalla and Newman (1994), Newman and Howard (1999a, 199b), Neapolitan (1996), Kangaspunta et al. (1998a, 1998b), Marshall (1998), Spencer (1993), Pease and Hukkila (1990; Tseloni and Pease (1994), Joutsen (1998); Aromaa and Joutsen (2003).

The quality control measures incorporated in the UN Crime Survey data collection process are, like most research, far from perfect, but include a number of aspects. First, and perhaps key among these is the clear definitions of terms on the survey questionnaire sent to governments, plus a separate set of instructions for each component part (police, courts, prosecutions, prisons). The following are the instructions given on the survey for respondents completing the section on policing expenditures (but the wording is essentially the same for the other parts, substituting the relevant term for police):

> "Total police budget/financial resources should include all monies allocated to the civil police function at the national level, including salaries and fixed assets. When calculating salaries, please include total monies spent on every individual employed in the given area. When calculating fixed assets, please include all monies invested in non-personnel assets, such as buildings, automobiles and office equipment." (Centre for International Crime Prevention, 1999; 8)

A second aspect of quality control is the effort made to track each individual person who is responsible for completing the survey in a particular country. This reduces anonymity, generates accountability, and improves communication between respondents and the survey team. Respondents working for individual governments can contact the UN data coordinators with questions, and UN data coordinators can contact respondents to verify data where necessary. A third key control discussed by Newman and Howard (1999a; 8), is how political pressures relating to the release of national-level data promote accountability: Governments generally do not release data into the international arena without having first closely scrutinized it. Some countries may, of course, overtly abuse or misrepresent their statistics for political gain, but with increasing democracy in Eastern Europe and the ex-Soviet block this type of abuse has undoubtedly declined in recent years. Other aspects of quality control for the UN crime survey have improved over the quarter century that the survey has been conducted (see Burnham 1999 for an overview of the survey's origins and history). A

fourth quality control strategy is the maintenance of institutional memory through the continued involvement of key researchers. Even as new administrators or clerical staff are introduced into the process, this institutional memory reduces the likelihood of basic errors in data collection and processing. Fifth, the survey format has been revised to incorporate new and innovative approaches as part of a learning process.[3] Sixth, the small 'core' group working on the survey at the UN are informed by an informal 'friends of the survey' international network of survey researchers. Seventh, the survey process and speed of communication (facilitating the checking of data) has been helped by the Internet during the last decade. Eighth, the increasing amount of published academic research based on the survey, and its resultant gradual permeation of media and academic debates on cross-national issues, and its public availability on the Internet (via www.unodc.org) means that its utility is increasingly recognized - another form of institutional memory. It should also be borne in mind that, relative to many national-level surveys such as the US National Crime Victimization Survey or the British Crime Survey which are sponsored by national governments, the UN Crime Survey runs on a remarkably small budget. To our knowledge, the survey is typically undertaken largely by one professional researcher who later employs the skills of a data entry clerk. While other researchers are consulted as discussed above, no independent survey companies, data collectors or researchers are contracted. By most western research standards, the survey team are shockingly under-resourced. Subject to such constraints, the survey is arguably quite remarkable, while most of its critics derive from a better-resourced tradition.[4] This is not intended to suggest that the survey is perfect or might not benefit from improved data collection practice and further reliability and validity checks: Indeed, as discussed later, it is hoped that the present study might be used to stimulate better reporting of data by countries. However, it remains the case that, when approached with the necessary caution, the survey is the most comprehensive existing resource for many areas of cross-national research using criminal justice data.

The general quality control issues discussed above were supplemented with a further set undertaken specifically for the present study. The sixth survey requested information for 1995 to 1997 inclusive. Generally, those countries that reported data did so for each year. Data for 1997 was cross-checked for consistency with that for previous years. In a small number of cases where 1997 data was absent it was substituted with data for 1996 where available. Preliminary descriptive data analysis was undertaken to eyeball the expenditure data for obvious errors. Errors in reporting and recording as well as at the data entry stage were the most likely explanation for the small number of extreme values that were found. For example, an examination of expenditure upon Canadian prisons revealed that the source data set contained a value in Canadian dollars that was in error by two decimal places. This datum was corrected. However, to ensure a conservative approach, a small number of data items were excluded where their true value could not be verified even by cross-checking. Most of these were not from highly developed countries, so it is arguably more likely that such discrepancies reflect variation in the quality of data reporting and recording systems. A possible error source could be discrepancies introduced via exchange rates during the conversion of currencies into US dollars, particularly if local currencies were extremely volatile. The country-level specifics of the study's quality control process are detailed in a technical appendix.

Despite the various quality controls, the present analysis should be viewed with caution. Even though they are the best available, the data are imperfect. Indeed, if the data were perfect there would be no need for estimation of missing values: global criminal justice expenditure could be found by simply summing the values reported by individual countries. However, we proceed from the reasonable position that the data cup is half full rather than half empty.

2.2 Sampling and Representativeness of Countries

Richer countries can afford to spend more upon data collection efforts. They are consequently over-represented in this data set as in most cross-national analyses. The sample of countries is therefore not representative of the world. However, the key issue for present purposes is not whether the sample of countries is representative. The key issue is whether the *relationship* between GDP and criminal justice expenditure that derives from the data is representative. Since, as will become evident, the greatest divergence in this relationship occurs among richer countries, it is a positive attribute of the dataset to have an over-representation of richer countries (those toward the right-hand side of Figures 1 to 4). The majority of missing data points relate to poorer countries which, in absolute dollar terms, contribute far less to global expenditure totals.

A second aspect of the study overtly seeks to address the possibility that the relationship between GDP and criminal justice expenditure is different for different sets of countries. Specifically, the statistical models that is split or kinked (Model 4) is included to capture the fact that there may be a different GDP-spending relationship among poorer countries than richer countries. In short, once again we conclude that the present data set is imperfect but that there is good reason to believe this is not an overwhelming obstacle.

2.3 Other data sources

The database of the International Monetary Fund (IMF) was used to complete the list of world countries and populations, and was the source of the bulk of information on international exchange rates. [5] The specifics of this procedure are detailed in the technical appendix. Using the 1997 exchange rates, expenditure data for all countries was converted to 1997 US dollar equivalents. All analysis described below uses US dollars.

3. Analysis

As an analytical quality control, statistical outliers were excluded from the main analysis. They were identified as national expenditures where the standardized residual resulting from the regression of per capita GDP upon per capita expenditure (for each stage) was greater than the 5 percent value from a two-tailed Student's distribution.[6] Eight data points were excluded, three of which related to Northern Ireland for each stage of its criminal justice system except prosecutions. Bahrain and Malaysia were outliers for police and prison expenditures respectively. Scotland, Colombia and Panama were outliers for prosecutions. All of the outliers were far higher per capita expenditures, relative to the per capita GDP, than the remaining countries.

1) Six regression models of the relationship between GDP and criminal justice expenditure were developed. The aim of the modelling was to identify the model which yielded the most accurate predictions for each of police, prosecution, court and prison expenditure. Each model represents a modification to the basic relationship between per capita expenditure and per capita GDP. That model was selected, for each stage of the criminal justice system, which best satisfied the criteria of goodness of fit and statistical efficiency as detailed below. For each stage, that model was then used to predict expenditure for countries for which the data were unavailable.

2) The first model is the simple linear per capita model. It reflects the possibility that richer countries spend more on criminal justice. The precise form of the relationship may, however, be more complex than the simple linear per capita model. It may differ across stages of the criminal justice system, across time, across country type and so on. The subsequent models investigate whether the estimates of global criminal justice expenditure are sensitive to various changes in the functional form of the regression model.

3) Figures 1 to 4 show plots of the data points for per capita GDP and per capita criminal justice expenditure, with a linear best fit line (Model 1)

superimposed. The chart for policing includes the quadratic fitted curve (Model 2) for illustrative purposes. Countries are labelled to the extent possible on the charts. The clusters of countries at the bottom left of each chart are not labelled if it was not feasible. On the charts, readers should note the variation in the vertical axes, which reflect greater per capita spending on policing, followed by prisons, courts, and prosecutions. Countries that were statistical outliers as described above, are not shown in the charts. Expenditures for countries with missing values were estimated using the regression equation, and global expenditure estimated as the sum of national values. In the global total, the reported values were used for countries that had previously been excluded as outliers in the model's development. The 95 percent confidence intervals were calculated according to the following formula for the variance of the prediction of total expenditure, based on the method described by Judge et al. (1988: 251):

$$\text{var} = \sigma^2 \mathbf{w}'[\mathbf{X_0}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X_0'} + \mathbf{I_m}]\mathbf{w} \qquad (1)$$

where σ^2 is the regression mean square error, $\mathbf{w}$ is an m x 1 vector of population values for the m countries we are predicting, $\mathbf{X_0}$ is an m x 2 matrix of explanatory variables for the countries we are predicting, $\mathbf{X}$ is an n x 2 matrix of explanatory variables for the countries in the regression sample, and $\mathbf{I_m}$ is an m-dimensional identity matrix.

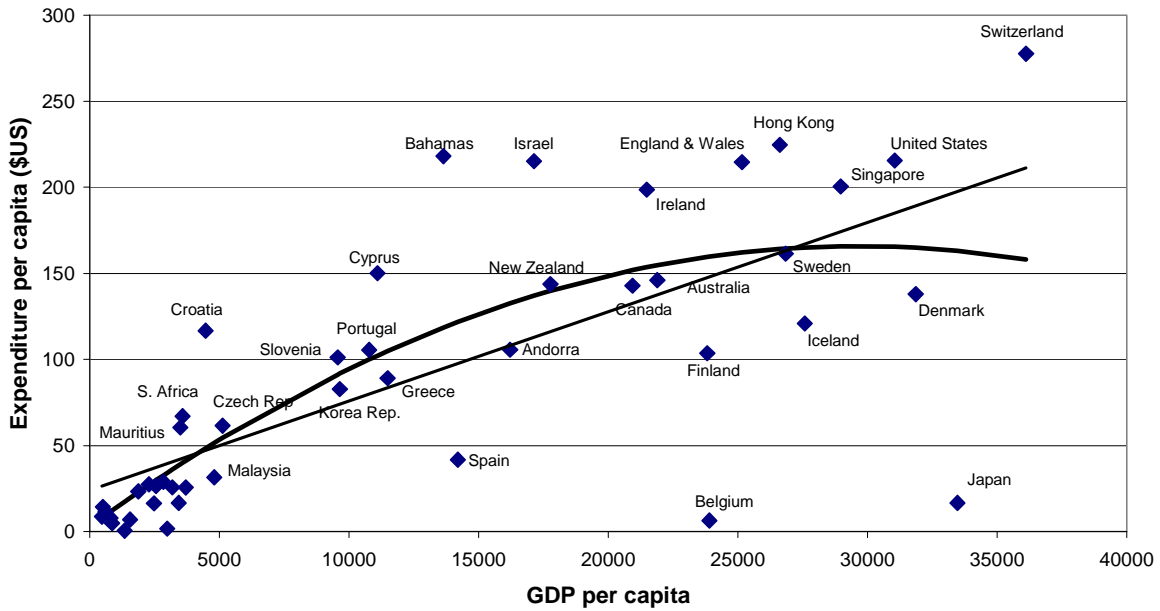## Figure 1: Policing Expenditure and GDP

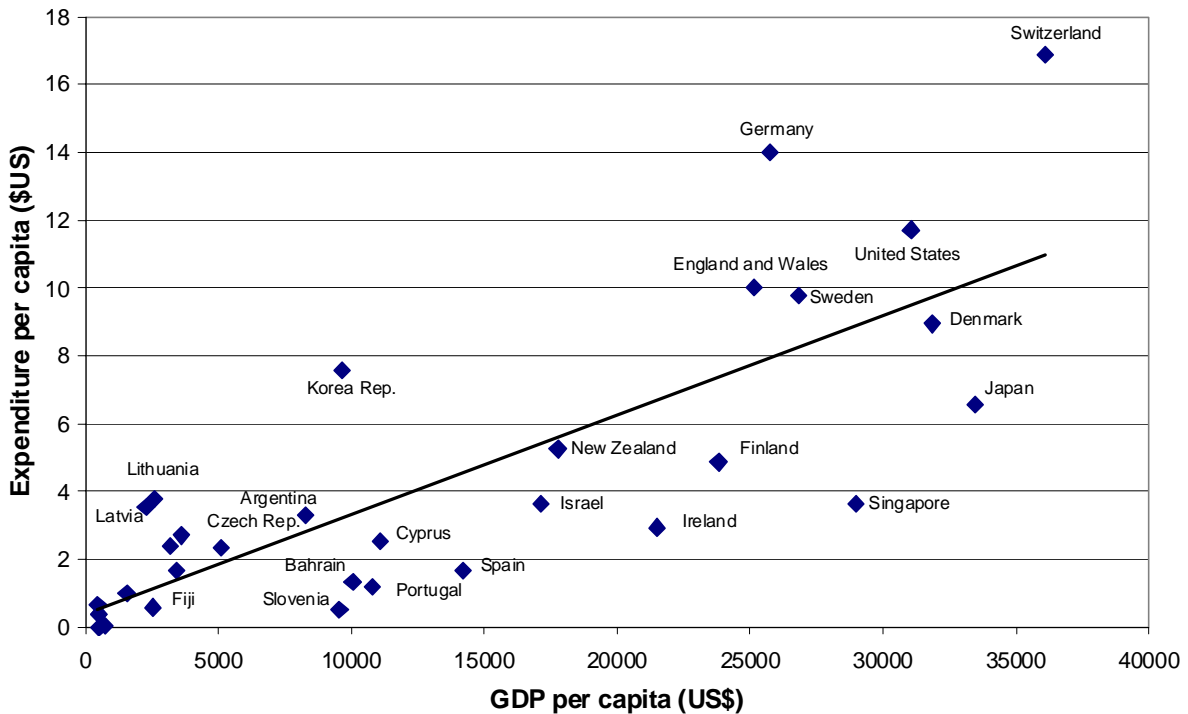

Figure 2: Prosecution Expenditure and GDP
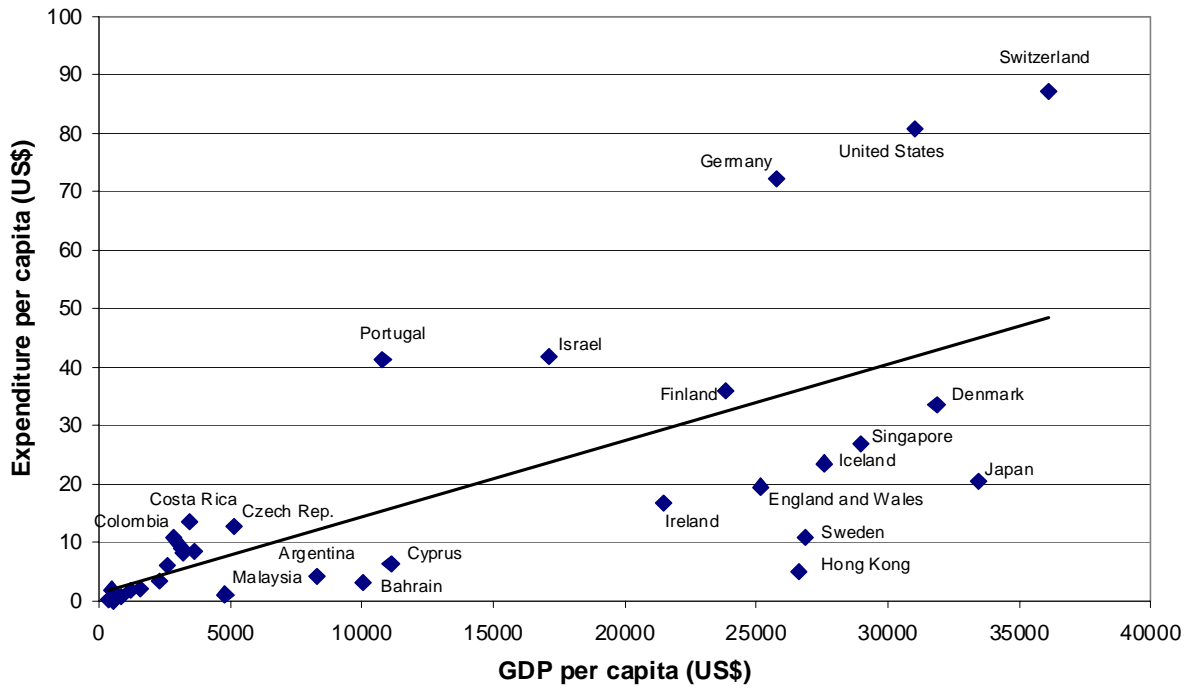
Figure 3: Court Expenditure and GDP

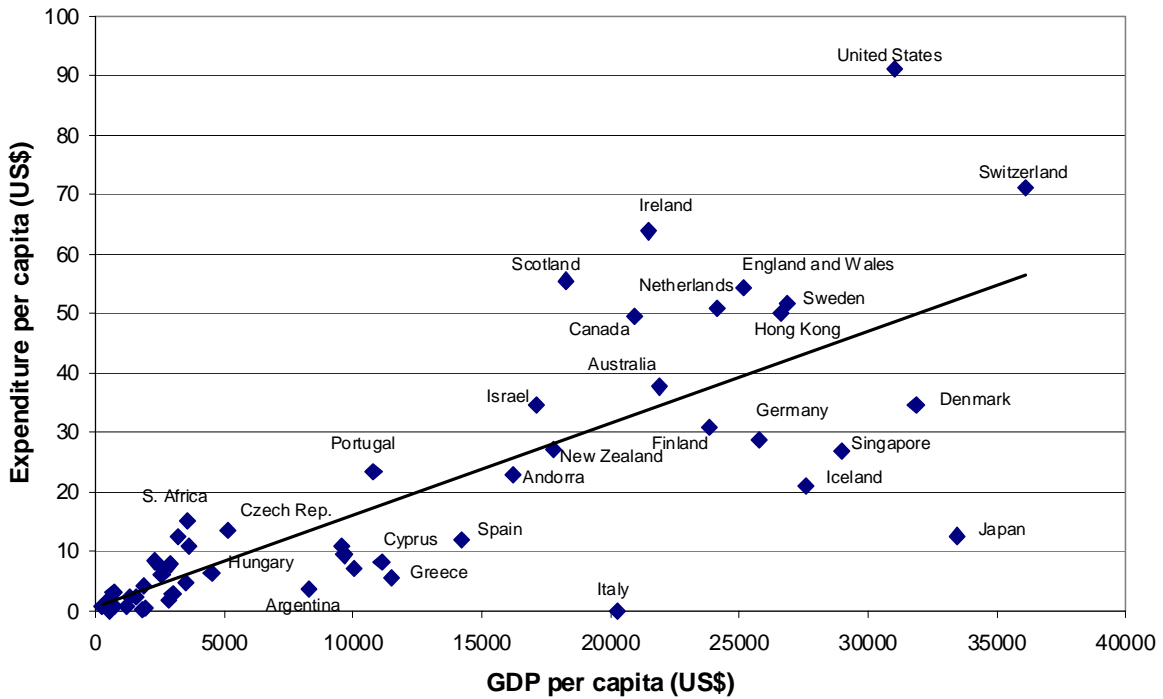# Figure 4: Prison Expenditure and GDP

TABLE 1: Results of Linear Per Capita Model

CJ Stage	N countries	# outliers excluded	R^2	Coeff.	Constant	Estimated Global Expenditure US$ Billion	Lower 95% CI	Upper 95% CI
Police	47	2	0.5405	0.0052	24.003	$286.560	80.220	492.901
Prosecution	35	3	0.6343	0.0003	0.3967	$12.055	4.447	19.663
Courts	34	1	0.483	0.0013	1.4604	$59.152	-10.073	128.376
Prisons	54	2	0.6181	0.0015	0.5946	$58.693	24.656	92.730
Total						$416.460	198.088	634.834

To illustrate the potential importance of modifications to the per capita model, consider again policing expenditure shown in Figure 1. Visual inspection of the scatter plot and fitted line suggest that there is some systematic pattern to the extent to which actual police expenditure per capita diverges from the fitted value (the value represented by the regression line)[7]. In particular, at low levels of per capita GDP (below $10,000), the linear model overpredicts per capita police expenditure for virtually all countries, while at higher levels of per capita GDP, the opposite is the case with most countries having lower fitted than actual values. It seems that, in the data, the slope of the relationship between per capita police expenditure and per capita GDP is higher at lower levels of per capita GDP and vice versa. Alternatively put, the marginal propensity to spend on policing is lower at higher levels of income. Such a relationship may fit with an explanation of criminal justice expenditure as a necessity whereby poorer countries spend disproportionately more on criminal justice than richer countries. This suggestion leads to the second statistical model of per capita police expenditure:

$$(X/P)_i = \alpha + \beta(G/P)_i + \gamma(G/P)_i^2 + \varepsilon_i \qquad (2)$$

$i = 1, \ldots, n.$

where X - police expenditure
P - population
G - GDP
ε - a random error term
i - country index
n - number of countries which report police expenditure (excluding outliers).

Through the quadratic term, Model 2 allows for a declining regression slope at higher levels of per capita GDP. The fitted regression line from this model shown in Figure 1 yields a noticeable improvement in fit upon the linear per capita model. There is a modest improvement in the model R^2 (the proportion of the variation in police expenditure explained by the model) from 0.54 to 0.59. In addition the coefficient on the quadratic term is statistically significant (two-tailed p-value = 0.030).

Clearly, introducing a quadratic term to the model is only one way whereby the per capita model could be modified. Other forms of non-linear relationship are possible, as are other transformations of the dependent and independent variables in the regression. Different stages of the criminal justice system might be best represented by quite different specifications. To investigate this issue fully, a series of regression models was estimated for each stage of the criminal justice system to find the one which fitted best. Three criteria were used to choose and refine the regression models. First, since the over-arching goal of the present study is to provide estimates of total world criminal justice expenditure, the class of regression model was selected using the criterion that it had the best in-sample forecasting performance. Second, in the interests of statistical efficiency, standard techniques of statistical inference were used to refine the resulting model. Third, we were at pains to ensure that the resulting predicted values passed what Hamermesh (1999) calls the "sniff test", that is, that they appeared to make economic sense.

Table 2 describes the various models that were evaluated and summarizes their in-sample forecasting performance. Models 1 and 2 were discussed above – the linear per capita model and quadratic per capita model, respectively. Models 3 and 4 are variants of the per capita model which attempt to capture differences in the slope of the relationship between per capita expenditure and GDP at different levels of per capita income. Model 3 is a double log model that allows for non-linearity through the specification of the dependent and independent variables in natural logarithmic (the ln function) form. This is a commonly used model in applied statistical research. Model 4 is effectively a piecewise linear model which estimates separate linear regression lines for 'poor' countries (defined as those with less than $10,000 per capita GDP) and 'rich' countries. Thus Model 4 also overcomes the issue of the over-representation of rich countries in the sample, a point discussed earlier. Close examination of Figures 1 to 4 suggests that the clump of countries at the bottom left (low per capita GDP, low criminal justice expenditure, many of which are not labelled due to lack of space in the charts) could be qualitatively different from other countries. By allowing for the possibility of a kinked expenditure curve, this model ensures the integrity of findings for 'poorer' countries if their relationship between GDP and criminal justice expenditure is somewhat different than that for 'richer' countries.

Models 5 and 6 are based around the notion that, since the aim is to predict total expenditure in each country, the appropriate dependent variable is the level of total criminal justice expenditure rather than per capita expenditure. Uncoupling population from the financial variables allows population itself to enter as a separate independent variable. Hence Models 5 and 6 are multiple regression models. Model 6 is a double log version of the relationship to account for potential non-linearities.

Table 2 reports the root mean square forecast error for each of the six models examined. This is defined as

$$\text{RMSFE} = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \hat{X}_i\right)^2}{n}} \qquad (3)$$

where X_i is expenditure on the relevant stage of the criminal justice system by country i and is the predicted or forecast level from the regression model under consideration for the same country. Note that $\hat{X}_i$ is a transformation of the regression fitted value for models 1-4 and 6; the per capita models require the multiplication of the fitted values by population while the double log models require the antilog transformation to be taken.[8]

$\hat{X}_i$

TABLE 2: The Model Selection Process

Model No.	Name	Equation	Root Mean Square Forecasting Error			
			Police	Pros.	Court	Prisons
1	Linear Per Capita	$(X/P)_i = \alpha_1 + \beta_1(G/P)_i + \varepsilon_i$	3695.03	205.84	1975.33	1759.40
2	Quadratic Per Capita	$(X/P)_i = \alpha_2 + \beta_2(G/P)_i + \gamma_2(G/P)_i^2 + \varepsilon_i$	3490.07	322.66	1930.18	1767.41
3	Double Log Per Capita	$\ln(X/P)_i = \alpha_3 + \beta_3\ln(G/P)_i + \varepsilon_i$	3740.37	187.70	2042.51	2098.39
4	Fully Interacted - Per Capita	$(X/P)_i = \alpha_4 + \beta_4(G/P)_i + \gamma_4D_i + \delta_4D_i(G/P)_i + \varepsilon_i$	3614.74	248.33	1964.19	1773.85
5	Multiple Regression	$X_i = \alpha_5 + \beta_5G_i + \gamma_5P_i + \varepsilon_i$	3659.00	148.39	1334.63	1425.55
6	Double Log Multiple Regression	$\ln(X_i) = \alpha_6 + \beta_6\ln(G_i) + \gamma_6\ln(P_i) + \varepsilon_i$	4638.97	239.11	2035.86	2466.01

Perusal of the columns of root mean squared forecast errors for each of the stages shows that, broadly speaking, the linear per capita and multiple regression models perform better than the double log models. For policing, RMSFE was minimized by the quadratic per capita model (Model 2) discussed above. This model also seemed to give reasonable predictions of the missing values. For the other three elements of criminal justice expenditure the linear multiple regression model had the best in-sample performance. However, this model had the unfortunate property of producing negative predictions of expenditure for a significant number of countries. One approach to solving this problem would be simply to set the predicted values to zero for such countries. As well as being *ad hoc* this is clearly unrealistic and so a different route was identified: since investigation suggested that statistically insignificant explanatory variables were the cause of the negative estimates, such variables were excluded from the regression models. Thus Model 5 was re-estimated for each of prosecution, court and prison expenditure with the restriction $a_5 = g_5 = 0$ imposed. This restriction was supported by the data in each case (the p-values were 0.3171 for prosecutions, 0.4575 for courts and 0.4134 for prisons).[9]

With a 'best' model in hand for each of the stages of the criminal justice system it was a simple matter to compute estimates of global expenditure for each category and to find the associated confidence interval. The results are displayed in Table 3. The 95 percent confidence intervals for police expenditure were computed as in equation 1 while for the other stages of the criminal justice system, the vector **w** was replaced by an m-dimensional unit vector. This is because, unlike models based on per capita expenditure, there was no need to weight the predicted values by the level of population.

4. Discussion

The use of several statistical models was more than just a changing room exercise of trying on various hats to see which fits best. Although it took place largely as an iterative and empirically-driven process, the models derive some theoretical justification from the literature on public budgeting. Lynch's classic text concludes that

> "*It is money.* The one common subject in any budget discussion is money. Other subjects are important, but they are mentioned in relationship to money or are translated into money. Budgeting involves dollars and cents often expressed in the millions of dollars." (Lynch 1979; 2, emphasis added)

TABLE 3: Results of Refined Estimation Procedure

CJ Stage	Estimated Global Expenditure US$ Billion	Percent of Total	Lower 95% CI	Upper 95% CI
Police	$222.530	61.87	16.990	428.070
Prosecution	$11.163	3.10	6.950	15.376
Courts	$63.472	17.65	26.216	100.727
Prisons	$62.528	17.35	26.171	98.886
Total	$359.693	100.00	149.550	569.836

Note to Table 3: Total in percentage column does not equal 100 due to rounding.

In practice, society's ideal of a welfare-maximization goal is transgressed by private interests from sources including pressure groups, bureaucrats and voter preferences as well as the influence of previous budget levels that produce path dependence (Steiner, 1971; Rubin, 1997; Peacock, 1978; Tullock, 1989). More specifically in relation to national spending on criminal justice, the major conclusion of the only other previous systematic cross-national study was that "governments allocate resources to criminal justice with little, or no, attention to outcomes." (Spencer 1993; 7). Even a critic of our parsimonious approach to variable selection cannot overlook its possibilities as a foundation upon which future research might build. Such future research would be testing additional hypotheses of the predictability of national-level criminal justice expenditure. We anticipate however, that such research may well experience negative or rapidly diminishing returns to the investment.

The introduction to this study began to address the utility of the present type of cost estimation exercise. It was suggested that global-level cost estimates may, in time, become increasingly useful - just as at one time it may not have been intuitively apparent that regional, national, city, or village-level estimates were of any utility. It was also suggested that the potential for wilful abuse of the estimates is insufficient reason to stop the pursuit of knowledge. However, some aspects of the potential utility of the present study will benefit from further elucidation. The present exercise is intended as a preliminary step towards better estimates in the hope that future research and data collection efforts can improve upon them. However, it may still be possible to develop some tentative comparisons to other global expenditures: If 1997 criminal justice expenditures were inflated to 2004 prices, global expenditure on criminal justice would be estimated as $424 billion.[10] This is around half the estimated $812 billion global expenditure on armaments, five times the estimated $80 billion spent on education, and roughly double the turnover of the global tobacco industry.[11] Other policy-relevant research questions could include: What are the main between-country and within-country variations in criminal justice expenditures? How do they vary across stages of the criminal justice system? What is the income elasticity of demand for criminal justice? Why do some countries spend proportionally more on some parts of the criminal justice system than others? Are there particular 'types' of countries that spend more on some stages of the criminal justice system than other? If prosecution and courts represent the stages of 'justice', are there some countries that pay a relatively greater attention to justice than others? How do these and other variables change over time? How do the income elasticities of demand for criminal justice compare to those for other public and non-public goods? This is not intended as an exhaustive list but, rather, is suggestive of the potentially broad and policy-relevant research agenda that could emerge.

As the first systematic empirical estimates of global spending, the current estimates may have the potential to instigate an improvement of international data collection. Improvements in data quality rather than refinements in statistical regression method are the preferable long-term source of improved confidence in global criminal justice expenditure estimates. The relevant United Nations bodies should consider using the present estimates to encourage countries to improve their data collection and reporting practices (a listing of actual and predicted expenditures by country and stage of the criminal justice system is available from the corresponding author upon request but, at the suggestion of several commentators, was excluded here for brevity). This is not meant to condescendingly suggest that poor countries will improve data collection when faced with the threat of estimates produced by foreigners, but it may stimulate greater effort in data collection and reporting (particularly a reduction in unnecessary non-responses and partial responses) in some marginal cases. In addition, the use of predictive techniques for the imputation of missing values, such as those used here, may have broader implications for the use of cross-national data other than that on criminal justice expenditure.

5. Conclusion

The study found that criminal justice expenditure levels are significantly tied to levels of available public monies as determined by the strength of a national economy. Despite acknowledged limitations of data and method, the study produced the first systematic empirical estimate of global expenditure on criminal justice. The confidence intervals are soberingly wide, the estimates should be interpreted with caution, and the study viewed as an incremental step towards improved data collection and estimation method. It is hoped that the present study provokes better data collection and encourages the investigation of methods for the estimation for missing data in this field.

Money makes the world go round, or so they say. However, while money does not necessarily take responsibility for the sun's gravitational pull upon the earth, it does facilitate the allocation of scarce resources in human society. Whether or not it proves to be a robust theory of global expenditure upon criminal justice remains to be tested by future research.

References

Aromaa, K. and M. Joutsen. 2003. 'Introduction' in K. Aromaa, S. Leppä, S. Nevala, and N. Ollus (Eds.) *Crime and Criminal Justice in Europe and North America 1995-1997*. HEUNI Report 40. Helsinki: United Nations European Institute for Crime Prevention and Control.

Berk, R. A. (1990) A primer on robust regression. In *Modern Methods of Data Analysis*, ed. J. Fox and J. S. Long, 292-324. Newbury Park, CA: Sage Publications.

Burnham, R. W. 1999. 'A short history of the collection of UN crime and justice statistics at the international level' in G. Newman (Ed.) *Global Report on Crime and Justice*. New York: Oxford University Press.

Centre for International Crime Prevention. (1999) *Questionnaire for the Sixth United Nations Survey of Crime Trends and Operations of Criminal Justice Systems, covering the period 1995-1997*. Vienna: United Nations Office for Drug Control and Crime Prevention.

Cohen, Mark. (2000) 'Measuring the costs and benefits of crime and justice' in Measurement and Analysis of Crime and Justice, volume 4 of Criminal Justice 2000. Washington D.C.: Department of Justice, National Institute of Justice.

Goldberger, Arthur S. (1968) The Interpretation and Estimation of Cobb-Douglas Functions. *Econometrica*, Vol. 36, Issue 3/4, 464-472.

Farrell, Graham, Erin Lane, Ken Clark and Andromachi Tseloni. (2001) 'What does the world spend on policing?' *International Journal of Comparative S*ociology, XLII, 1-2, 59-73.

Granger, Clive and Paul Newbold. (1977) *Forecasting Economic Time Series*. New York: Academic Press.

Hamermesh, Daniel S. (1999) The Art of Labormetrics. *National Bureau of Economic Research Working Paper 6927*. [http://www.nber.org]

Joutsen, Matti. (1998) 'Introduction', pp.1-19 in Kristiina Kangaspunta, Matti Joutsen and Natalia Ollus (Eds.) 1998a. *Crime and criminal justice systems in Europe and North America 1990-1994*. Monsey, NY: European Institute for Crime Prevention and Control; Criminal Justice Press.

Judge, G., R.C. Hill, W.E. Griffiths, H. Lutkepohl and T.C. Lee. (1988) *Introduction to the Theory and Practice of Econometrics*. New York: Wiley.

Kangaspunta, Kristiina; Matti Joutsen; Natalia Ollus (eds.) (1998a) *Crime and criminal justice systems in Europe and North America 1990-1994*. Monsey, NY: European Institute for Crime Prevention and Control; Criminal Justice Press.

Kangaspunta, Kristiina, Matti Joutsen, Natalia Ollus and Sami Nevala (eds.) (1998b) *Profiles of Criminal Justice Systems in Europe and North America 1990-1994*. Monsey, NY: European Institute for Crime Prevention and Control; Criminal Justice Press.

Lynch, Thomas D. (1979) *Public Budgeting in America*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Marshall, Ineke Haen. (1998) 'Operation of the Criminal Justice System', pp. 54-114 in Kristiina Kangaspunta, Matti Joutsen and Natalia Ollus (Eds.) 1998a. *Crime and criminal justice systems in Europe and North America 1990-1994*. Monsey, NY: European Institute for Crime Prevention and Control; Criminal Justice Press.

McDonalds. (2000) 'McDonalds Reports Record Global Results' *McDonalds Financial Press Release* 26, January 2000.

Nalla, Mahesh K. and Graeme R. Newman. (1994) Crime in the U.S and the Former U.S.S.R.: a comparison of crime trends from the third United Nations Survey. *International Journal of Comparative and Applied Criminal Justice* 18 (1), pp. 85-94.

Neapolitan, Jerome L. (1996) 'Cross-national crime data: some unaddressed problems' *Journal of Crime and Justice*, XIX, 1, 95-112.

Newman, Graeme (ed.). (1999) *Global Report on Crime and Justice*. Published for the United Nations Center for International Crime Prevention. Oxford and New York: Oxford University Press.

Newman, Graeme, and Greg Howard. (1999) 'Introduction: data sources and their use' in G. Newman (ed) *Global Report on Crime and Justice*, published for the United Nations Center for International Crime Prevention. Oxford and New York:.Oxford University Press.

Peacock, Alan. (1978) 'The Economics of Bureaucracy: An Inside View' in *The Economics of Politics*. IEA Readings 18. London: Institute for Economic Affairs.

Pease, Ken and Andromachi Tseloni. (1994) 'Juvenile-adult differences in criminal justice: evidence from the United Nations Crime Survey' Howard Journal of Criminal Justice, 35, (1): pp. 40-60.

Rubin, Irene S. (1997) The Politics of Public Budgeting: Getting and Spending, Borrowing and Balancing, 3rd. Chatham, New Jersey: Chatham House Publishers, Inc.

Spencer, Jon. (1993) 'Criminal Justice Expenditure: A Global Perspective' *Howard Journal of Criminal Justice*, 32, 1, 1-11.

Tullock, Gordon. (1986) 'Wanted: New Public Choice Theories' in Martin J. Anderson (Ed.) *The Unfinished Agenda: Essays on the Political Economy of Government in Honor of Arthur Seldon*. London: Institute of Economic Affairs, pp. 13-28.

Transparency International. (2002) *The Transparency International Corruption Perceptions Index 2002*. Berlin: Transparency International (at www.transparency.org).

Steiner, Peter O. (1971) 'The Public Sector and the Public Interest' in Robert H. Haveman and Julius Margolis eds. *Public Expenditures and Policy Analysis*. Chicago, Illinois: Markham Publishing Company.

Van Horne, Sheryl and Graham Farrell. (1999) *Drug Offenders in the Global Criminal Justice System.* HEUNI Paper #13. Helsinki, Finland: The European Institute for Crime Prevention and Control, affiliated with the United Nations.

United Nations Development Programme. (2002) *The Human Development Report 2002*. New York: Oxford University Press.

TECHNICAL APPENDIX

The first section of this appendix contains information relating to the UN Crime Survey data. The short second section relates to supplementary data sets.

The UN Crime Survey Data
The Bulgarian government reported expenditure data on courts, prosecution and prisons but not on policing, noting that "The data on the numerical strength of the police personnel and financial resources of the police in Bulgaria are a state secret." Partial data on national policing expenditure was provided by the Estonian government with the rider that "[The] Security police board is excluded as their personnel figures are not public." The Czech Republic reported that "The budget of the Czech Republic police covers total for current and capital data, including the Czech Republic Investigation Office. Data for the Czech Republic Investigation Offices within these years were not separately surveyed and investigated and that is why we cannot present them separately."

The response from the Israeli government noted that the police budget "[i]ncludes the budget of police prosecutors" while with respect to prosecution, "[t]he budget figures are for governmental prosecutors within the Ministry of Justice. " The report of the Romanian government noted in relation to prosecutions that "The budget of the Public Ministry is different from and independent of the budget of the Ministry of Justice and that of the Police. The structure of the annual budget comprises in principle the following sections: expenses for staff; maintenance (buildings, repairs, vehicles and other assets, fuel, office equipment, inventory items); investments (furniture, vehicles)."

In relation to courts in Bulgaria it was observed that "Included are financial resources for wages and salaries and for social insurance contributions of the magistrates and the personnel of the Ministry of Justice and Legal Eurointegration." A comment in relation to Danish courts noted that "Finances of Danish law courts, capital summary and edp in registration of title to land, etc. are included in this category"

The sixth UN Crime Survey was accompanied by a Crime Guide.[12] The Guide contained comments from responding governments, translated as appropriate. Based on the comments we made five adjustments to a specific data item. Each adjustment involved an assumption or extrapolation aimed at reducing the amount of error: The Danish government noted that "police and prosecution data could not be separated", and that the same value had been included for both in the database. This value was split proportionally between police and prosecutions

to reflect the proportional spending levels of countries which reported the data separately. For other countries, 6.08% of the combined police and prosecution budget was spent on prosecutions, and 93.92% on policing, and the Danish police and prosecution budget was split to reflect these proportions; A similar revision was made in relation to courts and prosecutions for Germany and Switzerland. In relation to Germany, the Crime Guide noted that a "[d]istinction between court and prosecution budget is not available", while the response of the Swiss government noted that the court budget "includes the prosecution budget". Countries that reported both prosecution and court data separately spent an average of 16.24% of the combined total on prosecutions and 83.76% on courts. The German and Swiss budget figures were split to reflect this average; The Finnish government noted that "Lay judges in circuit courts were financed by municipal resources until 1998. The cost, about 30 million marks is not included in [the] state budget [the figure given]." Consequently, the reported Finnish court budget of 931 million marks was adjusted to 961, an increase of 3.22%. The fifth change to the data was based upon a comment from the Canadian government that their prison budget included "[a]dult data only". Separate budget data for adult and juvenile prisons was not available for any country, and so the proportions going to each in other countries could not be calculated and applied to Canada. However, data on the number of beds in adult and juvenile prisons was available elsewhere in the UN survey, although not for Canada. On average, the number of juvenile beds was 5.53% of adult beds. On the assumption that proportional expenditures may mirror the proportion of beds, the Canadian prison budget was inflated by 5.53% as a crude estimate of the additional cost of juvenile institutions.

Twenty additional data items were excluded from the analysis. Four exclusions were made in relation to policing: Kazakhstan, Kyrgyzstan, Scotland and Tanzania. The lowest given value for policing expenditure in any of these countries was six times greater than expenditure in any other country. Six exclusions were made in relation to prosecutions: Bulgaria, Denmark, Georgia, Kyrgyzstan, Malaysia and Tajikistan, whose values were at least ten times

those for other countries. Three exclusions were made in relation to court expenditure: Georgia, Kazakhstan and the Republic of Korea, each of which had expenditures at least six times greater than all other countries. Three exclusions were made in relation to prison expenditure: Kyrgyzstan, Jordan and Panama, each of which had values that were several orders of magnitude greater than those of any other country. An effort was made to trace the values to determine the origins, but they were in the UN crime survey database. It is likely that some of the errors arose during the reporting stages – government officials who completed the questionnaire may have misinterpreted an aspect of the question, or perhaps may not have noticed that data were to be reported in millions of local currency units. The UN database is compiled in local currency units and so such discrepancies would be unlikely to be noticed if the currencies were unfamiliar. Hence it is reasonable to expect that expenditure data reported incorrectly by a responding country would go unnoticed prior to the release of the dataset. It is notable that seven of the exclusions were from the three countries Georgia, Kazakhstan, and Kyrgyzstan, suggesting that a general error was made in relation to those countries. It is possible that one or more of the criminal justice expenditures that were excluded as extreme outliers were not errors. For example, it could be that these three former Soviet states undertook massive capital expenditures on their criminal justice systems in 1997 - a point that would require further investigation. For present purposes however, such extremes of expenditure would have been excluded as statistical outliers by the method detailed in the Analysis section. With the exceptions of Scotland, Denmark and possibly Bulgaria, the other six exclusions related to countries that are not in the upper-echelons of development rankings and which are therefore arguably less likely to have high quality reporting and recording systems for their aggregate-level data relating to expenditures across the country.

Supplementary Data Sources

The list of world countries and their populations was compiled from the UN crime survey and the database of the International Monetary Fund (IMF). Where population data could not be found in these sources it was supplemented with data from, in order of preference, the 1998-based National Population Projections[13], the UN Population Division (World Population Prospects Population Database), and the CIA World Factbook.[14]

The primary source of data on national gross domestic product (GDP) was the database of the IMF, supplemented with other sources as necessary. Eight other sources were used, and in a few instances GDP data for 1996 was used when 1997 data was not obtained. The other sources of GDP data were the CIA World Factbook 2001, the Northern Ireland Financial Times, the UN Human Development Report 1999, The Scottish Parliament[15], UNCTAD[16], Geographic's Country Profiles[17], Yahoo! Finance[18], and Freiburg University[19]. There was no reason to believe that these sources contained bias in the reporting of GDP data, and every reason to believe that the inclusion of the data produced an incremental improvement in our database that was preferable to the exclusion of that data element.

Expenditure data was reported in local currency units. Exchange rates for 1997 that were not in the UN crime survey database were obtained from the database of the International Monetary Fund.

Where there was a difference in specific data items between the various possible sources, the differences were typically minor discrepancies of a few percentage points, almost certainly reflecting differences in the reporting and recording of national-level data. In such instances, the UN crime survey data was used if available. There were only a handful of cases where a decision had to be made regarding a non-trivial difference between sources. In such instances the two data items were compared and it would usually be apparent that one contained an obvious error (probably due to reporting, recording or data entry error at different stages and in different sources). Where it was not clear which

data item was preferable, additional qualitative information was sought to allow the identification of the most appropriate datum. If available, a third source of data was utilized as an arbiter. The end result was a database with complete GDP and population data for 204 countries.[20]

Notes

[1] Private policing and security expenditures are not included since, while they are substantial, they do not utilize government funds.

[2] We thank Kauko Aromaa for comments on this section in particular.

[3] It is, as with any change, possible to dispute some of the revisions: Recent streamlining of the survey has removed overlapping years of data from each iteration of the survey, leaving the survey open to suggestions that fewer validity checks can be undertaken.

[4] For discussion on such issues we thank Adam Bouloukos, formerly of the United Nations Office of Drugs and Crime.

[5] We thank Matthew Fleming, then of the International Monetary Fund, for assistance in locating the relevant exchange rates. The online currency converter available at www.oanda.com/convert/classic was used to complement the IMF data in only a handful of instances.

[6] An alternative approach to extreme observations would have been to use a robust regression technique (see, for example, Berk, 1990)

[7] Such a pattern in the residuals might reflect the influence of omitted explanatory variables, discussed further below.

[8] For the double log models there is an issue concerning bias correction when forming the prediction of total expenditure (see Goldberger, 1968). We applied the bias correction suggested in Granger and Newbold (1977) however this rarely improved the in-sample forecasting performance and even when it did, the improvement was not enough to significantly increase the ranking of the relevant model. We therefore do not pursue the issue further.

[9] The restriction increased the root mean square forecasting error but never by enough to prevent this model being 'best' in its column. The figures were 154.39 for prosecutions, 1369.87 for courts and 1451.48 for prisons.

[10] Inflation adjustment using the inflation calculator of the US Bureau of Labor Statistics at http://www.bls.gov/cpi/home.htm.

[11] These comparative global spending estimates were taken from sources of varying degrees of integrity, relevance and comparability, and are included only to be suggestive of possible future applications of the findings - the intention is to put criminal justice expenditure in a broad comparative perspective rather than to provide definitive estimates for other items. The source of the estimates is as follows: Global military spending of $812 billion in 2001 from the Center for Defense Information at www.cdi.org/issues/wme , retrieved on 2nd October 2002; Annual global spending on education from United Nations Children's Fund (UNICEF) report, The State of the World's Children, 1999 (see http://www.penpress.org/docs/PEN84.pdf); Annual sales turnover of the tobacco industry from www.didyouknow.cd/fastfacts/money.htm., page 1, retrieved 2nd October 2002.

[12] Our thanks are due to Sami Nevala for making the Crime Guide available.

[13] Series PP2 no22; retrieved Jan 31, 2002 from statbase: www.statistics.gov.uk/statbase/

[14] Available online at http://www.cia.gov/cia/publications/factbook/

[15] http://www.scottish.parliament.uk/

[16] www.unctad.org

[17] www.geographic.org

[18] http://biz.yahoo/

[19] www.zmk.uni-freiburg.de/Europe/yugoslavia.htm

[20] We note that 204 countries is not a definitive count or one that is necessarily approved by the United Nations. However, it is hoped that we did not exclude any major populations that would significantly influence the key findings of the study.

28