ANALYSING SEXUAL OFFENCE SENTENCES: AN EMPIRICAL APPROACH
A P* and D I** Simester

The judiciary have frequently refused to compare sentences, stating that the circumstances of the offender and offence vary so greatly that there is no workable basis upon which to do so. This article presents an empirical model that seeks to identify and quantify the relative significance of various mitigating and aggravating factors which affect sexual offence sentences. The results suggest that it is possible to develop such a model. In doing so, they offer evidence as to the current approach to sexual offence sentencing of the New Zealand Court of Appeal.

Introduction

The requirement for a consistent application of sentencing principles has been judicially recognised on many occasions.1 Although such consistency is necessary to preserve public confidence in the administration of justice, there is, however, a reluctance to formalise these principles.2

It is only by allowing the sentencing authorities a wide discretion that they are enabled to take account of the innumerable factors affecting the nature of the offence, the circumstances of the offence and the circumstances of the offender, all of which should ordinarily be weighed in determining the appropriate sentence in the particular case.

For this reason the courts have expressed a reluctance to compare sentences from different cases, and have preferred to treat sentencing as a subjective matter; thereby securing the flexibility to tailor a sentence according to the nature of the offence itself and to the circumstances of the particular offender.3 Nevertheless, the twin goals of flexibility and consistency need not conflict: one may be flexible in a consistent manner. Given that sentencing is dependent upon the circumstances of the offence and of the offender, individual sentences may be expected to be consistent according to the presence or absence of the various factors that constitute such circumstances.

This article develops the above statement, and investigates the evidence supporting a “flexible tariff” approach — so roundly dismissed by Lord Reid in DPP v Ottewell4 — to sentencing for crimes of a sexual nature5 in the New Zealand Court of Appeal. The following propositions are developed:
(i) It is possible to extract from the cases a model of consistent sentencing, founded on a “flexible tariff” basis, which will demonstrate in a systematic manner the effect of surrounding features upon individual cases.
(ii) As a corollary, it is possible to compare sentences, provided the presence or absence of mitigating and aggravating circumstances is taken into account.

If a model of the type suggested in proposition (i) can be established, then it is submitted that proposition (ii) follows. Additionally, it is implicit in the first proposition that the judicial approach to sentencing is intuitively, even if not explicitly, to vary individual sentences from average levels in a systematic fashion, according to the mitigating and aggravating features of each case.

Our article describes the development of such an empirical model. By doing so, it contends that propositions (i) and (ii) are valid. Additionally, the specific

---

* Assistant Lecturer, Faculty of Commerce, University of Auckland NZ.
** Research Fellow, Department of Marketing, University of Auckland NZ.
instantiation of the model that is developed will offer some interesting insights into the existing judicial approach to sentencing for sexual offences.

**Legislative Influence on Sexual Offence Sentencing**

In New Zealand, pursuant to the Criminal Justice Act 1985 and the Crimes Act 1961, sentences for sexual offending are generally custodial. Section 5 of the Criminal Justice Act lays down a general sentencing principle that where, in the course of committing an offence punishable by imprisonment for a term of five years or more, an offender has used serious violence against or caused serious danger to the safety of any other person, a full-time custodial sentence must be imposed (in the absence of special circumstances). This principle is repeated specifically in the context of sexual violence offences by s 128B of the Crimes Act,³ which supplements the judicial perception that “even if the degree of force is not great, the violation of the victim’s body can only be described, in any ordinary sense, as in itself serious violence”.⁷

Within the context of the above mandate there is a general discretion granted by s 72 of the Criminal Justice Act, empowering a court to impose less than the statutory maximum penalty.⁸ It is this section that allows the flexibility to “tailor” penalties according to the relevant mitigating or aggravating features of each case. The maximum penalty thus serves “... as an indicator to the Courts of the seriousness with which that class of offence is regarded by the legislature or regulation making authority”;⁹ it is reserved only for the most serious offences of each type.

The flexibility granted by s 72 is supplemented by ss 7(1) and 7(2), which require the sentencing judge to exercise his/her discretion with regard to the desirability of keeping offenders in the community and of keeping custodial sentences as short as possible (so far as is practicable and consonant with promoting the safety of the community).

An additional statutory intrusion into the sentencing process is provided by s 75 of the Criminal Justice Act. That section empowers the High Court to impose an indeterminate sentence (commonly known as preventative detention) upon sexual offenders over the age of 25 who have been previously convicted of a sexual offence¹⁰, where the court is satisfied that such a sentence is necessary for the protection of the public. The Court of Appeal has on several occasions emphasised that preventative detention should be reserved only for exceptional cases.¹¹

**The Nature of the Model**

In contrast to the principles espoused in *DPP v Ottewell*,¹² it has been suggested in New Zealand that the court’s approach is to begin with an intuitive sentencing level, and to add or subtract from that level according to the presence of mitigating or aggravating factors:¹³

... for a rape committed by an adult without any aggravating or mitigating features, a figure of five years should be taken as a starting point in a contested case. Aggravating features can include additional violence or indignities, acting in concert with other offenders, the youth or the age of the victim, intrusion into a home...

The judicial statements to this effect form the theoretical rationale for our model, which may be represented as a multi-variate linear regression equation:

\[ S = B + W_1F_1 + W_2F_2 + \ldots + W_nF_n \]

Where \( S \), the dependent variable, is the actual sentence length imposed (in years) as a proportion of the maximum penalty for the relevant offence.¹⁴ \( B \), the intercept, is the “base” level proportion of the maximum penalty for any sexual offence, which
one might expect in an unremarkable instance (where no particular mitigating or aggravating features are present). $F_1$ through $F_n$ are the independent variables, measuring the relevance of the various mitigating or aggravating factors to any particular case; and $W_1$ through $W_n$ are coefficients of sentence proportion to be added to or subtracted from $S$ according to whether the corresponding factor is present in that case.\textsuperscript{15}

Thus the actual sentence is determined by accumulated deviations from a base sentencing level; which deviations result from the presence, absence and relative importance of various mitigating and aggravating factors. We have thereby attempted, mathematically, to model the approach suggested in $R \times Clark$,\textsuperscript{16} beginning with $B$ as an intuitive standard sentence and adjusting that level according to the relevance of mitigating and aggravating features, $F_1$ through $F_n$.

It will be recognised that the dependent variable above is not bound by logical limits of 0 and 1 (which represent the range of possible custodial sentences). The model should, strictly, be defined piecewise: values less than 0 or greater than 1 are to be taken as 0 and 1 respectively, indicating that no term of imprisonment is necessary or that the maximum penalty is appropriate. It is conceivable that a sexual offence could be committed in such circumstances that a custodial sentence is inappropriate or, alternatively, that the maximum penalty is warranted. Cases of each type may not be identically severe.\textsuperscript{17}

Development of the model requires the following steps:

(i) Identification of the potential variables ($F_1$ through $F_n$) from the circumstances of the offence and offender that may influence the sentencing decision.

(ii) Collection of data to use for estimating the model from actual cases.

(iii) Analysis of the data collected to find the set of coefficients ($W_1$ through $W_n$) and intercept ($B$) which best fits the model to the analysed cases.

(i) Specification of Potential Factors

Twenty three variables describing the circumstances of an offence were identified (as being potentially relevant to sentence lengths) from a preliminary examination of criminological literature,\textsuperscript{18} sentencing decisions, and from responses to a series of questionnaires distributed to a sample of criminal barristers, the police and the public. The variables derived were then refined, following inspection of a pairwise correlation matrix, to reduce the level of collinearity in the data.\textsuperscript{19}

Thirteen variables describe the circumstances of the offence: Young Victim and Elderly Victim (which represent the number of years by which the victim was aged under 20 or over 60 respectively); Direct Violence (a composite measure derived from a combination of two individually linked variables, measuring physical injury and use of violence); Indirect Violence (an equally weighted agglomeration of four correlated factors: threat of violence, production of a weapon, threat of a weapon and housebreaking); Emotional Injury (suffered by the complainant as a result of the offence); Wife; De Facto; Breach of Trust (whether the opportunity for offending arose out of a position of authority or a supervisory role over the complainant); Acquainted (whether the offender was known to the complainant other than through family or spousal ties or a position of trust); Male Victim; Indecencies (the commitment of gross sexual indignities or perversions beyond those normally associated with the offence); Number of Offenders; and Premeditation.

Five variables were identified as describing the character of the offender: Offender's Age; Previous Convictions; Probation Report (whether the report was negative, neutral or positive); Remorse; and Psychiatric Problems.
The remaining five variables either fit into neither of the above categories, or are relevant to both: Year (the year in which the judgment was delivered); Maximum Penalty (for the particular offence charged); Appeal (whether appealed by the Crown or by the offender); Guilty Plea; and Intoxication.

Four other possible variables were considered for inclusion in the model: recommendations (if any) made by a jury; the offender's ignorance of the law; the elderly age of the offender; and the effect of a penalty upon the offender's family. Unfortunately, there were insufficient cases available where these factors were relevant to reliably calibrate coefficients for them. While this implies that their omission had little effect on the modelling of the decisions in the calibration dataset, if further research is undertaken using additional observations, these variables should be considered for potential inclusion in the model.

(ii) Collection of Data

The source data for the model was taken from 67 recent sexual offence sentencing decisions handed down by the Court of Appeal. Only Court of Appeal decisions were used, since with the exception of Her Majesty's Prerogative to review decisions (which to date has never been used in the context of a sentencing appeal from New Zealand), the Court of Appeal is New Zealand's highest appellate court. As such the Court of Appeal is expected to lay down sentencing guidelines and policy for all New Zealand courts, and to have a stabilising influence upon sentencing trends in the High and District Courts. It is anticipated that the consistency which is to be expected from the Court of Appeal may not be found in the lower courts.

The reliability of the data collection process was tested by repeating the measurement process for a sample of the observations using another researcher. No trouble was encountered where variables were binomial or were combinations of binomial measures; nor were the inherently quantitative variables (for instance, Year and Offender's Age) difficult to measure. There was, however, an instance of minor inconsistency found in the measurement of each of the Premeditated, Indecencies, Probation Report and Psychiatric Problems variables. These variables require a degree of subjectivity when assessing their severity in a given case: future research in this field might ensure the reliability of data collection by establishing explicit coding instructions. Any resulting improvements in the data base would be expected to improve the explanatory power of the model.

(iii) Analysis of the Data

Ordinary least squares regression was conducted using all 23 independent variables. From this preliminary investigation 18 variables were allocated weights that were consistent with expectations and were included in the final model. None of the five variables dropped from the model had coefficients which were significantly different from zero.

The eventual inclusion of 18 sentencing factors in the model contrasts with the somewhat smaller number found in models generated by other multivariate studies of this type (Wilkins et al, 1978; Konceni and Ebbesen, 1981, 1982; Sutton, 1978).

The Results

The calculated factor coefficients and estimated base sentence level are set out below, together with their corresponding impact on a potential 14 year sentence (Table A). The weights shown are a maximum influence: where a factor is only partially present or relevant, the sentence will be affected to a lesser extent. We
have included in the table T test statistics for each coefficient, together with initial Pearson correlation coefficients for each independent variable with the dependent variable.\textsuperscript{26}

<table>
<thead>
<tr>
<th>Circumstances of the offence</th>
<th>Coefficient weights</th>
<th>Years per 14 year maximum</th>
<th>T Value</th>
<th>Pairwise correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Victim</td>
<td>0.1517</td>
<td>2.12</td>
<td>1.8</td>
<td>0.065</td>
</tr>
<tr>
<td>Elderly Victim</td>
<td>0.2361</td>
<td>3.30</td>
<td>2.0</td>
<td>0.082</td>
</tr>
<tr>
<td>Direct Violence</td>
<td>0.1510</td>
<td>2.11</td>
<td>2.4</td>
<td>0.377</td>
</tr>
<tr>
<td>Indirect Violence</td>
<td>0.1672</td>
<td>2.34</td>
<td>2.5</td>
<td>0.341</td>
</tr>
<tr>
<td>Emotional Injury</td>
<td>0.0848</td>
<td>1.19</td>
<td>1.4</td>
<td>0.182</td>
</tr>
<tr>
<td>Wife</td>
<td>-0.2621</td>
<td>-3.67</td>
<td>-3.6</td>
<td>-0.298</td>
</tr>
<tr>
<td>De Facto</td>
<td>-0.2089</td>
<td>-2.92</td>
<td>-2.4</td>
<td>-0.198</td>
</tr>
<tr>
<td>Acquainted</td>
<td>-0.0778</td>
<td>-1.09</td>
<td>-1.8</td>
<td>-0.056</td>
</tr>
<tr>
<td>Male Victim</td>
<td>-0.2549</td>
<td>-3.57</td>
<td>-4.1</td>
<td>-0.341</td>
</tr>
<tr>
<td>Indecency</td>
<td>0.0934</td>
<td>1.31</td>
<td>2.3</td>
<td>0.179</td>
</tr>
<tr>
<td>Number of Offenders</td>
<td>0.1587</td>
<td>2.22</td>
<td>1.7</td>
<td>0.313</td>
</tr>
<tr>
<td>Premeditation</td>
<td>0.1210</td>
<td>1.69</td>
<td>2.5</td>
<td>0.261</td>
</tr>
</tbody>
</table>

| Character of the Offender   |                     |                           |        |                    |
| Offender's Age              | -0.0958             | -1.34                     | -1.8   | -0.052             |
| Previous Convictions        | 0.0138              | 0.19                      | 0.3    | 0.217               |
| Remorse                     | -0.0764             | -1.07                     | -2.1   | -0.304              |
| Probation Report            | -0.0701             | -0.98                     | -2.7   | -0.438              |

| Other                       |                     |                           |        |                    |
| Year                        | 0.0400              | 0.56                      | 1.4    | -0.025             |
| Appeal                      | 0.0322              | 0.45                      | 2.1    | 0.107               |
| Base Level (intercept)      | 0.3057              | 4.28                      | 8.1    |                    |

| Model Statistics            |                     |                           |        |                    |
| R²                          | 0.7621              |                           |        |                    |
| Adjusted R²                 | 0.6728              |                           |        |                    |
| F Statistic                 | 8.541               |                           |        |                    |

**Evaluation of Coefficients**

Collectively, it appears that the Court of Appeal has intuitively given more weight to the circumstances of the offence than to the character of the offender when determining appropriate sentence lengths for sexual offenders. The average absolute effect (upon a potential 14 year sentence) of factors describing the circumstances of the offence is 2.29 years. By contrast, the corresponding impact of factors describing the offender is only 0.90 years.

The victim's age variables were found to be amongst the most important aggravating features of sexual offences. Sexual attacks on children under the age of 10 have in the past increased sentences by over two years (for an offence with a 14 year maximum penalty), while a similar attack on an 80 year old victim could be expected to increase the sentence by up to three years (assuming other variables
remain the same).\textsuperscript{27} The extreme old age of the victim and the youth of the victim were noted as aggravating circumstances in the Court of Appeal cases of \textit{R v Te Pou}\textsuperscript{28} and \textit{R v Puru}\textsuperscript{29} respectively.

Surprisingly, the direct physical violence factor was found to have a smaller effect on sentence lengths than indirect violence. This perhaps suggests that the courts view all sexual offences as involving serious physical violence and are therefore less inclined to differentiate between cases on the basis of this feature. Alternatively, it may indicate that the offensive nature of sexual offences is derived as much from emotional shock as from physical damage occasioned to the complainant.\textsuperscript{30} Nevertheless, the results do indicate that violence is an important aggravating feature. Where an offence includes the most serious direct or indirect violence the expected sentence has tended to increase by at least two years (for a potential 14 year sentence).

The presence of emotional injury was found to have a relatively small influence upon sentence lengths. This finding may be attributable to the argument that emotional injury, like physical violence, is inherent in any conception of sexual offences. Additionally, the true significance of the emotional injury factor may be obscured in part by the absence of reliable information concerning emotional injuries sustained by victims of sexual offences. Until August 1987\textsuperscript{31} the courts had to rely on probation reports as the sole source of information regarding the physical or emotional harm suffered by the victim. These reports frequently do not consider the impact of the offence on the victim, focusing more often on the circumstances of the offender; perhaps not surprisingly, given the difficulty of evaluating the long term emotional harm likely to be suffered by the victim of a sexual offence. It follows that the lack of reliable information as to the extent of emotional injury sustained by a victim may explain the diminished impact of this factor (relative to the weights for direct and indirect violence). It remains to be seen whether the 1987 amendments, seeking to rectify the paucity of such data, will cause greater importance to be attached to the emotional harm inflicted upon complainants.\textsuperscript{32}

The Court of Appeal has rejected the notion that a different scale of penalties is applied for intra-marital rape in \textit{R v N}.\textsuperscript{33}

We are firmly of the view that in cases of this nature no separate regime of sentencing is called for simply because parties are married or have been in a continuing sexual relationship .... Parliament has made no distinction in the penalties between spousal and other kinds of rape, and the sense of outraged and violence experienced by a woman in that position can be equally severe.

The results indicate, however, that contrary to the statements of the Court of Appeal in \textit{R v N} and in \textit{R v M}\textsuperscript{34}, a lesser range of penalties have tended to be applied where the parties have been married or have otherwise been in a continuing sexual relationship. In such circumstances an average reduction in sentence of approximately three years may be expected for an offence carrying a maximum penalty of 14 years. It is acknowledged that these cases often may be accompanied by a number of mitigating factors — such as an absence of additional violence or the presence of remorse — which could be expected to lead to more lenient sentences.

The weighting afforded to Wife and De Facto contrasts with the relative insignificance of the acquaintance factor. Where the complainant and the victim were acquainted outside the family environment (and there was no breach of trust involved) a reduction of one year could be expected in a 14 year sentence — indicating that the mere fact that the parties are acquainted is, in itself, of lesser importance.

All male complainants found in the cases used to calibrate the model were the victims of incestuous offences perpetrated by the complainant's father, uncle or
Another relative. Thus the proposition that intra-family offending is treated as a separate category of offences is reinforced by the sizeable influence that the male victim feature has on sentence lengths.

These three intra-family variables (wife, de facto and male victim) account for three of the four most important influences on sentence length. In particular, the prominence of the male victim feature as a mitigating factor also suggests that the courts may treat the violation of a male as less serious than the violation of a female, notwithstanding the view of Cooke J in *R v Ngawhika*:

"... any suggestion of a hierarchy of sexual offences classifying violation of a female as prima facie more serious than violation of a male has to be avoided."

Where the facts of the case disclose gross sexual indecencies, the courts have been willing to add almost one and a half years to a potential 14 year sentence. Such a finding is consistent with the Court of Appeal's recognition of this variable as an aggravating factor.

The courts have previously noted the need to impose deterrent sentences for attacks on outnumbered and defenceless victims. In *R v Stoddart* Cooke J commented upon such a situation: "The Courts have not only to condemn this in the name of the community; they must also do what they can to deter the kind of group brutality involved."

In the present research the weight attached to the number of offenders factor suggests that the presence of five or more participants will add a year or more to a sentence for an offence which is punishable by up to 14 years imprisonment; a result that is consistent with the Court of Appeal's stated approach to this factor.

Each man who joins in... gang activity by any form of encouragement or help contributes to the ongoing wave of crimes. Each must know this very well. It is a process of group stimulation. There is mutual encouragement to reject civilised standards of human decency. The victim and the community are treated with arrogant contempt... The Courts should do what they fairly can, without totally jettisoning the need to judge each individual offender on his or her merits or lack of merits.

The weight accorded to premeditation may be influenced by the nature of sexual offences, which very commonly involve an offender luring, forcing or, in the case of incestuous offences, inducing the complainant to perform the desired acts. Thus sexual offences generally contain a preliminary act, although few have "long range" premeditation; moreover, where "long range" premeditation does exist in a particular case, its presence will normally be evidenced by other factors already accounted for in the model (for example, housebreaking).

The significant influence of the offender's age variable confirms the Court of Appeal's comments in *R v Puru*. The offender's age is considered an important mitigating factor when the offender is aged under 20: for instance, a 16 year old offender may expect a reduction of almost one and a half years for an offence carrying a maximum penalty of 14 years imprisonment. Legislative policy to this effect is found in the Children and Young Persons Act 1974, which requires the court to consider the welfare and needs of youthful offenders during the "vulnerable" years of their adolescence.

Previous convictions appear to have no practical impact upon sentence lengths. This result was obtained after omitting cases where preventative detention was imposed, as these decisions were inconsistent with sentencing patterns in the other cases. Preliminary analysis, including these cases, indicated that a record of previous convictions caused an average increase in sentence levels of up to one and a half years. It therefore seems that prior convictions will influence the decision as to the type of penalty required, rather than the length of a fixed custodial sentence. Such analysis does not establish whether the effect that previous convictions have upon
sentence levels arises as a result of a perception that offenders with extensive criminal records have little rehabilitative potential and are likely recidivists, or whether the courts are (in effect, rather than deliberately) merely repunishing an offender for an earlier offence, an issue which is beyond the scope of this article.

Remorse and a positive probation report have similar influences as mitigating factors, with potential reductions of around one year (from a 14 year sentence) accruing for a positive probation report or a fully repentant defendant. Conversely, a negative probation report is expected to result in sentence increases of a similar amount.

The weight calculated for the year variable is consistent with statistics presented by the Justice Department (1988). Departmental studies found that the average sentence length for rape rose from approximately 4 years in 1979 to 4.9 years in 1986; representing a mean increase of 0.12 years each year. By comparison, the coefficient of the year variable suggests that over five years the base sentencing level for rape (which carries a 14 year maximum penalty) is increased by an average 0.56 years; representing an average of 0.11 years per year. Such an increase (as calculated by the model) is consistent with the actual average yearly increase stated in the Department of Justice report, which was restricted to cases of rape alone. It should be noted that the model analysis does not indicate whether this increase is as a result of the passing of s 5 of the Criminal Justice Act 1985 (an explanation propounded in the Justice Department’s report, 1988); or of an ongoing trend, as was suggested by Woodhouse P in 1984 following an analysis of statistical evidence.

Although not conceptually a mitigating or aggravating factor, the identity of the appellant party was expected to indicate whether the sentencing level in a case is at the upper or lower end of the sentencing spectrum. The results are consistent with this reasoning: the positive weight indicates that where the offender was appealing the sentence, the sentence was at the upper end of the normal sentencing spectrum (and at the lower end where the Crown was appealing).

The value found for B indicates a base sentencing level in 1983 of 4.28 years for a 14 year sentence. This is not inconsistent with the stated base level of 5 years in R v Clark in 1987, nor with the increase in average rape sentences from around 4 years in 1979 to around 4.9 years in 1987 — particularly as the average sentence length increased by 17% in 1986.

The five potential factors not included in the model were Maximum Penalty; Intoxication; Psychiatric Problems; Breach of Trust and Guilty Plea. These factors were excluded as their calculated weights were very small and were liable to fluctuate upon minor changes in the overall model.

The sign of the weight initially calculated for the maximum penalty variable regularly fluctuated following minor changes to data or other variables in the model, indicating that in the decisions analysed, the size of the maximum penalty did not influence the “sentence length: maximum penalty” proportion. Such a result is especially meaningful as it shows that sentencing levels vary uniformly with the maximum penalties — suggesting that individual sentences are strongly dependent upon legislative guidelines as to the seriousness of that particular offence; and that the courts apply similar sentencing principles to each category of sexual offence.

The absence from the model of the intoxication and psychiatric problems variables is not unexpected. The omission of intoxication as a relevant factor is consistent with the recommendations of the Committee of Inquiry into Violence, which suggested that there should be no diminution of sentences solely on the grounds that the offender was at the time of the offence affected by drugs or alcohol. Any effect due to the presence of psychiatric problems may be neutralised
by two conflicting principles – the reduced culpability of such offenders (a mitigating factor) is countered by the need to protect the public from offenders with psychiatric disabilities, a consideration which supports longer prison sentences.

By contrast, omission of the breach of trust and guilty plea variables is of greater interest, since the courts have previously indicated that both features are relevant to the sentencing decision.51 However, like the psychiatric problems factor, a breach of trust appears to import two conflicting sentencing principles: the need to deter abuses of trust (a concern militating against the offender)52 is set against the general tendency to treat domestic offences more leniently (most offenders in a position of trust were found to have domestic or family ties to the complainant). To the extent that it is an aggravating factor, the effect of breaching a trust may well be absorbed by weights afforded to other factors within the model (for instance, Male Victim); thus the fact that a trust has been abused is of no additional significance.

In respect of guilty pleas, while they have repeatedly been identified as a mitigating factor by the courts, there are three potential explanations for their omission from the model.

First, as the cases analysed were all Court of Appeal decisions, the relevance of a guilty plea in reducing the pressure on the criminal justice system or saving a witness from having to give evidence is less apparent. The Court of Appeal is less exposed to the effect on the victims of having to give evidence in sexual offence trials, since such evidence would only be called where the Court of Appeal was hearing an appeal against both sentence and conviction (and even then the court is likely to rely on the transcript of evidence from the High Court or District Court hearing). Furthermore, the mere fact that an appeal has been made to the Court of Appeal (whether by the Crown or the offender) neutralises the time and cost savings achieved by avoiding a High Court or District Court trial due to a guilty plea.

Secondly, a guilty plea is a far less immediately emotive feature of a case than variables such as direct or indirect physical violence, indecencies, the number of offenders, the offender’s and complainant’s ages, or the presence of domestic or family ties between the offender and the complainant. As a consequence, a greater subjective weight seems to be afforded to these more emotive factors than is given to the offender’s plea.

Thirdly, the presence of remorse, one of Hall’s three reasons for regarding a guilty plea as a mitigating factor (1987: 42), is allowed for by the model as a separate factor; which could thus be expected to reduce the significance of the guilty plea variable. Moreover the mere fact of an appeal (if made by the offender) is evidence that the appellant’s remorse is limited.

The low weight placed on the guilty plea variable is supported by the minor reference or total omission of the details of the offender’s plea in many Court of Appeal decisions. The occasional explicit reference to the mitigating effect of the plea suggests that the courts recognise the importance of this factor, but in practice its influence may be less than consciously intended.

**Evaluation of the Overall Model**

**Quality of Fit**

Applying the T test, all but three of the parameters were found to be significantly different from zero at the 90% confidence level. Eleven of the variables have coefficients which are significant at the 95% level, and three variables and the
intercept were significant at the 99% level. Similarly, the F-test statistic of the joint distribution of the independent variables is 8.54, which is significant at the 99% level.

The coefficient of determination ($R^2$) for the model is 0.762, or 0.673 when adjusted for degrees of freedom, indicating that the model explains much of the variance of sentence lengths in the calibration dataset. This finding contrasts with those in other empirical studies examining the effect of offence and offender characteristics (Sutton, 1978; Wilkins et al, 1978. See also Vining and Dean, 1980; Lovegrove, 1989), which typically have accounted for only 35% to 50% of the sentencing variance.

The results demonstrate strong evidence of a systematic relationship between the aggravating and mitigating factors of a case and the sentence length imposed. They thereby offer support for our earlier propositions, (i) and (ii). In doing so, they contradict the opinion of Lovegrove (1989: 39) that “the multiple regression model is incompatible with the structure of judicial thought . . . .”

Lovegrove’s further comment, that “consequently, some of the complexities of judicial thought processes cannot be adequately represented”, is not addressed. Proposition (i), and the model itself, postulates a general structure for sentencing decisions. Within that structure, the specific aggregational steps taken by individual judges to assess particular cases are subjective, and differ for each judge.

Validation

A series of diagnostic tests were used to test the appropriateness of the application of ordinary least squares regression to the case data, by examining the dataset’s compliance with the following assumptions: additivity and linearity in the relationships between the independent variables and the dependent variable; a normally distributed error term with a mean of zero; no residual correlation or constant residual variance; and absence of multicollinearity.

The linearity of the relationships between the independent variables and the dependent variable was examined using two residual plotting techniques and by recalibrating the model with subsamples of the original dataset. While the results of these tests were not conclusive, it appeared from the partial residual plots that the relationships of all of the independent variables follow a linear trend. The subsamples test also indicated the presence of general linear trends, although the evidence was not as strong for the Direct Force, Indirect Force, Emotional Injury, Premeditation, Offender’s Age and Year variables.

Recalibration of the various subsamples was also used to check whether the relationships between the independent variables and the dependent variable are additive. There was no strong evidence of non-additivity. Again, however, these results were not fully conclusive; a larger dataset is required to properly investigate the validity of the additivity assumptions.

Normality of the residuals was evaluated using a number of statistical and graphical tests: a normal probability plot; standard normal deviates; the Shapiro-Wilk statistic; the Kolmogorov-Smirnov statistic; and both the kurtosis and skewness coefficients. The results of these tests make it clear that the residuals do satisfy an assumption of normality. Although the distribution is a little flatter and more heavily tailed than ideal, the curve is symmetrical and departures from true normality are not serious. The mean of the residuals is zero.

Standard tests of residual correlation or non-constant variance, such as inspection of residual plots or the Durbin Watson statistic, were rendered ineffective by the discrete nature of many of the independent variables and by the cross-sectional
nature of the study. Instead White’s (1980) consistent covariance matrix test was used to test the null hypothesis that the residuals have constant variance and are not dependent upon any of the independent variables. The test indicated that neither residual correlation nor non-constant variance is a problem. This result is hardly surprising given the relatively large number of statistically significant independent variables which influence the error term.

Multicollinearity in the model as a whole was examined by inspecting the variance inflation factors and the eigenvalues of the correlation matrix. Specific collinearities were then identified by evaluating the proportion of variance accounted for by each principal factor, by inspecting the pairwise correlation matrix, and by regressing each independent variable against the remaining independent variables.

The maximum variance inflation factor is 1.70, and the spread of the eigenvalues (maximum divided by the minimum) is 52.7; indicating that the overall level of multicollinearity in the data is low (Montgomery and Peck, 1982). The results of the specific collinearity tests were consistent with this conclusion, although there was a suggestion that Year and Young Victim are subject to low levels of correlation with other independent variables.

Stability and Predictive Ability

The stability of the coefficient estimates was checked by recalibrating the model after omitting random subsets of observations from the original dataset. Chow’s (1960) test was then used to evaluate the equivalence of the coefficients obtained from the smaller datasets with those obtained using all 67 observations. No evidence was found with which to reject the null hypothesis that the coefficients of the recalibrated models are equivalent to those derived when using the full dataset. This result suggests that the model is stable to the omission of observations from the calibration dataset. The stability of the estimated weights and the corroboration of these estimates by prior expectations offers grounds for confidence in the validity and reliability of the model. Although a number of alternative techniques for improving the efficiency of the model have been identified, the measures of adequacy of fit and statistical significance noted above indicate that the model is a useful representation of reality even without modifications.

Perhaps the best illustration of the validity of the model is its ability to predict the sentence lengths of cases which have not been used in the estimation of the model. The predictive ability of the model was tested against an additional 22 cases which had not been included in the original sample from which the model was calibrated.

The pairwise coefficient between actual and predicted sentence lengths was found to be 0.749. This finding was supported by the levels of Thiel’s inequality coefficients. The two coefficients calculated, 0.13 (for Thiel and Goldberger, 1961) and 0.26 (for Thiel, 1966), are both proximate to the common lower bound of 0 (representing perfect predictive accuracy). The average absolute error in the prediction of the dependent variable was 0.113.

These results provide further support for the approach to sentencing postulated by such a model, and suggest that predicted sentences are generally consistent with those arrived at by the Court of Appeal.

Conclusion

We conclude by restating our propositions:

(i) It is possible to extract from the cases a model of consistent sentencing, founded on a “flexible tariff” basis, which will demonstrate in a systematic manner the effect of surrounding features upon individual cases.
(ii) As a corollary, it is possible to compare sentences, provided the presence or absence of mitigating and aggravating circumstances is taken into account. A conceptual approach of adding or subtracting from a base sentencing level according to the presence and relative influence of the various mitigating or aggravating factors satisfies the structure of the empirical model contemplated in the first proposition. Evidence collected in this article, which demonstrates the validity and reliability of our estimated model, would therefore seem to confirm that it is possible to construct such a model using scientific and empirical principles.

Demonstration of proposition (i) above implies that, although perhaps not conscious or deliberate, the judicial approach to sentencing is to some extent structured. If it is possible to construct a valid scientific model of the sentencing decision, then the judiciary's approach to the sentencing process must at least in part be inherently systematic. Furthermore, the nature of the model developed in order to satisfy proposition (i) suggests that the approach in fact adopted by the courts is to vary sentences according to the mitigating and aggravating features present in each case.

At a general level, such conclusions could hardly be surprising. Sentencing may well be impressionistic in the minds of judges, but underlying those impressions must be the features of each case — those which are aggravating will (as a matter of common sense) tend to increase the sentence imposed, whilst those which are mitigating will have the opposite effect. The significance of proposition (i) lies rather in the systematic approach to sentencing that it postulates. Such factors as are relevant to each case can, it appears, be permitted to exacerbate or ameliorate the penalty imposed in a manner that is quantifiable with some degree of precision. Furthermore, by establishing that it is possible to develop a consistent and valid model of the sentencing process for sexual offences, direct evidence is made available as to the structure of that intuitive process.

As a corollary, proposition (ii) is also confirmed. Where the presence and relevance of the mitigating and aggravating circumstances for a category of cases can be ascertained with a degree of precision, the cases within that category may be usefully compared. The ability to do so could offer practitioners an additional tool for use with sentencing appeals, through the ability to demonstrate whether the original sentence was manifestly excessive or inadequate given the presence or absence of appropriate factors. Additionally a tool may assist the judiciary to preserve confidence in the administration of justice by facilitating the consistent application of sentencing principles; although it is recognised that the need for judicial assessment and discretion in the realm of sentencing cannot be displaced.

Further research aimed at assessing and improving the validity of scientific sentencing models would seem to be justified both in order to improve current knowledge of the judiciary's approach to sentencing in specific categories of offences, and in the hope of developing a precise and easily applied model for the assistance of both practitioners and the judiciary. There is potential to develop similar models for alternative categories of offences, or for individual offences within the broad category of sexual offences. Additionally, the specific content of the model yields an opportunity to assess the weights which perhaps ought to be accorded various features of each case, in the light of existing practice. A more accurate understanding of the criminal sentencing system as it presently works can only facilitate the process of legal reform.
NOTES

3 In R v Lawson [1982] 2 NZLR 219, the court expressly stated that sentencing is not to be treated as an exact science, and the circumstances of the offender can rarely be closely associated with those of another.
4 [1970] AC 642 at 647: "It was rather tentatively suggested by the learned Attorney-General that there is a 'tariff' for each kind of offence which is varied upwards or downwards according to the circumstances of the offence and the character of the accused. But offences of a particular kind vary so vastly in gravity that there cannot and should not be any 'normal' sentence and there is no workable standard by which to judge whether any particular sentence is extended beyond what is 'normal'." See also R v Williams [1983] Crim LR 693 and R v Ngawhika 3 July 1987 (CA 91/87) at p 5 of the decision.
5 Prescribed by ss 128 through 144 of the Crimes Act 1961 (NZ).
6 As amended by s 2 of the Crimes Amendment (No 3) Act 1985 (NZ).
8 Sexual offences do not provide for any mandatory sentences (cf for instance s 172 of the Crimes Act, which provides for an automatic sentence of life imprisonment in cases of murder).
9 Fisheries Inspector v Turner [1978] 2 NZLR 233 at 237, per Richardson J.
10 As defined in the section. The earlier offence must have been committed when the offender was over the age of 16.
11 Cf R v Bidwell 20 December 1985 (CA 249/85) and R v Sutton 24 November 1977 (CA 72/77), where the court referred to the sentence as being one of "last resort". As a result, between 1979 and 1987 only 14 offenders were sentenced to preventative detention. Of those paroled before 1985, the average sentence served was 12.7 years (Justice Department, 1983).
12 Supra, at note 4.
13 R v Clark [1987] 1 NZLR 380 at 383. See also R v Stanley 30 October 1984 (CA 202/84); R v Berrington 11 April 1986 (CA 308/86).
14 Approximately 15 individual offences fall within the general category of sexual offences. Since the maximum penalty varies amongst these offences, the dependent variable is defined as a proportion. For example, if a sentence of 7 years is imposed for rape (where the maximum sentence is 14 years) then \( S = \frac{7}{14} = 0.5 \).
   The dependent variable has not previously been so modified by other studies of this nature; it was, however, essential to do so in this study in order to ensure that the intercept \( B \) remained independent of the maximum penalty.
15 For instance, if factor \( F_i \) indicates whether the complainant and the offender are married, then \( F_i \) takes the value of 1 where the complainant in a particular case is the offender's wife and 0 otherwise; with a corresponding increment to the base sentence by \( W_i F_i \), being \( W_i \) (if married) or 0 (if unmarried).
16 Supra, at note 13.
17 In none of the cases used during the development and testing of the model did the estimated dependent variable breach the logical limits. Although the magnitudes of the coefficients eventually calculated, infra, allow a potential for extreme cases to breach these limits, the combination of factors required is so unlikely that the possibility is negligible.
18 In particular Hall (1987) and Thomas (1979).
19 A detailed discussion of the definition, quantitative measurement and scoring of the variables is available upon request from the authors.
20 Youthfulness was not recognised as a separate feature; where a breach of trust was involved (as to which see, infra).
21 Initial model estimation was performed using 98 cases; 27 of these were eventually omitted, as they did not disclose values for all 23 potential variables. A further four cases (where preventative detention was imposed) were also excluded from the final model. The above 98 cases, together with the additional 22 cases used to validate the model (see, infra) are as many as the authors were able to find: we estimate that they represent a major proportion of the relevant recent decisions.
22 Although the Court of Appeal has no jurisdiction to grant leave to appeal to the Privy Council in a criminal matter (see R v Kaitamaki [1981] 1 NZLR 527 at 528, per Richardson J), the Judicial Committee of the Privy Council may itself grant the leave to appeal by virtue of Her Majesty's Prerogative to review decisions (refer Arnold v King-Emperor [1914] AC 644; R v Naden [1926] AC 482).
That is, the coefficients had the expected sign: negative for mitigating factors, and positive for aggravating factors. See generally Lowegrove (1989).

Even at the 90% confidence level, the signs of these coefficients fluctuated regularly during development of the model.

Thus for instance an offence including less serious indecencies, where that factor might be given a normalised value of 0.5 rather than 1, would be expected to have its sentence increased by 0.65 years (where the maximum sentence was 14 years); since 14 (0.934) (0.5) = 0.654. A discussion of the normalisation of variable measurements is available from authors.

This data is included in order to illustrate the increased explanatory power of a joint distribution of the independent variables.

For an 80 year old victim, the factor Elderly Victim would be assigned a value of 20, normalised as 0.9091.

[1985] 2 NZLR 208.

Indeed, the long term physical harm resulting from violence associated with sexual offences is frequently less than that arising from other offences coming before the courts (for example, from the infliction of grievous bodily harm, aggravated assault, manslaughter or even murder).

Since November 1987 the courts have had access to Victim Impact Statements (pursuant to s 8 of the Victims of Offences Act 1987) which inform the court "about any physical or emotional harm, or any loss or damage to property, suffered by the victim of an offence". Additionally a 1987 amendment to the Criminal Justice Act (1987 No 168) amended ss 22 and 23 of the principal Act and empowered a court to call for an Emotional Impact Report when considering the possibility of ordering an offender to make reparation payments.

Any effect since 1987 is too recent to be incorporated adequately by this variable within the model.

24 November 1987 (CA 99/87) at p 6, per Casey J.
24 November 1987 (CA 114/87).
3 July 1987 (CA 91/87) at p 5.
See generally R v Paru [1984] 1 NZLR 248 at 251; R v Toko and Hutchinson 1 December 1982 (CA 122/82 and CA 123/82); and R v Kemp 24 September 1981 (CA 24/81).
R v Misimo 24 September 1987 (CA 163/87).
R v Stoddart [1986] 1 NZLR 264 at 257: "Most rapes probably involve at least some planning, even if it is not always long in advance of the crime".
See R v Brookes 19 August 1983 (CA 115/83).

In the case of a 14 year maximum sentence, sentences are expected on average to be 0.90 years higher if the offender is appealing than if the appeal is brought by the Attorney-General.

The expected base sentence level for other years would be estimated through adjustment by the weight allocated to the Year factor.

[1987] 1 NZLR 380; see also R v Villicky Clark 26 May 1987 (CA 50/87).
Justice Department (1988); average sentences increased by a further 17% in 1987. It is important to note that the average sentence length will not necessarily equal the base sentence level, which must be adjusted for the average effect of the various mitigating and aggravating features.

Both characteristics raised doubts as to whether the estimated weights were in fact significantly different from nil.


Except insofar as drug or alcohol abuse may be relevant to an offender’s rehabilitative potential. Intoxication may nevertheless be relevant to conviction. On this matter, see generally R v Kamipeli [1975] 2 NZLR 610 at 612-19.
R v Stanley 30 October 1984 (CA 202/84); R v B [1984] 1 NZLR 261.

It should be noted that the technical validity of the Kolmogorov-Smirnov and the Shapiro-Wilk tests has been questioned. The Shapiro-Wilk statistic assumes that the observations are derived independently, which is not necessarily the case when residuals are used. The critical values for the Kolmogorov-Smirnov test were shown to be unstable by Lilliefors (1967).
ANALYSING SEXUAL OFFENCE SENTENCES: AN EMPIRICAL APPROACH

REFERENCES


