Article

Mind the step: A more insightful and robust analysis of the sentencing process in England and Wales under the new sentencing guidelines

Jose Pina-Sánchez
University of Leeds, UK

Ian Brunton-Smith
University of Surrey, UK

Guangquan Li
Northumbria University, UK

Abstract

The 'England and Wales Sentencing Guidelines' aim to promote consistency by organizing the sentencing process as a sequence of steps, with initial judicial assessments subsequently adjusted to reflect relevant case characteristics. Yet, existing evaluations of the guidelines have failed to incorporate this structure adequately, instead concentrating solely on sentence outcomes. We use multivariate multilevel models to offer new insights into the decisions made throughout the sentencing process. Focusing on cases of assault sentenced at the Crown Court we show that the level of compliance with the guidelines is high. However, we also show that some case characteristics are being unduly considered at more than one stage of the sentencing process, meaning existing studies may be underestimating their true influence.

Keywords
Compliance, consistency, Crown Court, mediating effects, multivariate multilevel modelling, sentencing guidelines

Corresponding author:
Jose Pina-Sánchez, School of Law, University of Leeds, Liberty Building, Leeds, LS2 9JT, UK.
Email: j.pinasanchez@leeds.ac.uk

Introduction

Sentencing is a highly complex process. Multiple preliminary decisions are considered before the final sentence is passed. The list of intermediate considerations varies by jurisdiction, yet some of them are quite common: for example, assessments of the offence seriousness (giving place to specific starting points or sentencing ranges); changes of disposal type on the basis of personal mitigating factors or individual risk assessments; or guilty pleas triggering sentence reductions. Such judgements are generally undertaken in a sequential process, where earlier decisions can affect those that follow. This creates a complex network of relationships between case characteristics and intermediate judgements.

To date, quantitative studies of sentencing practice have largely failed to reflect the complexity of the sentencing process. The excessive reliance on the univariate regression model (i.e. based on a single dependent variable), in its different forms, is partly responsible. Although it is true that models are always, by definition, a simplification of reality, there are two important assumptions made by univariate regression models that are inconsistent with the complexities of the sentencing thought process. First, they focus on a single outcome, usually the final imposed sentence, with little attention given to initial judicial assessments or subsequent sentence modifications. And second, they assume that all case characteristics are considered jointly in a single stage immediately prior to sentence, which ignores the sequential nature of the sentencing process.

To provide a more comprehensive understanding of the sentencing process, some researchers have advocated modelling more than one outcome. Specifically, the probability of receiving a custodial sentence, and the duration of such punishment (Johnson, 2006; Ulmer and Johnson, 2004; Ulmer et al., 2011). Yet, while the specification of two or more separate models can afford us greater insight than solely looking at final sentence outcome, this assumes that the series of sentence outcomes (or intermediate decisions carried out throughout the sentencing process) are independent. When that assumption is wrong, parameter estimates and measures of uncertainty will be affected. Furthermore, treating these measures independently assumes case characteristics (or any other explanatory variables) included in the model have only a direct effect. That is, by not accounting for the complex net of intermediate decisions in the sentencing process, the standard approach fails to acknowledge the existence of indirect effects of certain case characteristics on the final sentence outcome. This not only limits our understanding of the sentencing process, it can also give rise to misleading results. In particular, the total effect of case characteristics contemplated at different points of the process will be biased.

The identification of the intermediate stages of the sentencing process is nowhere more clearly defined than in the England and Wales sentencing guidelines. Historically, judges in England and Wales have enjoyed a substantial degree of sentencing discretion, restricted just by statutory offence maximums and minimums, and the Court of Appeal. However, this status has changed radically over the last decade, with a series of wide-reaching reforms intended to introduce greater structure into the sentencing process through the establishment of a system of guidelines relating to appropriate sentences and tariff adjustments.

Initially, the guidelines adopted a more discursive narrative form, laying out general principles for the sentencing of a particular type of offence (Dhami, 2013b; Hutton, 2013). The importance of the guidelines was enhanced in 2009, with judges no longer simply being recommended to consider the guidelines but actively being required to adhere to them. In addition, since 2011 the Sentencing Council for England and Wales has been progressively replacing the existing guidelines with new ones that are structured around a series of steps, with sentencers requested to consider specific offence factors to determine a set of preliminary outcomes. Each of these steps deals with different details of the case (e.g. the harm caused to the victim and the offender's culpability is considered in Step One, while the entering and timing of a guilty plea is considered in Step Four), and collectively they are meant to be comprehensive and non-overlapping. Preliminary decisions made at each of these steps modify the remaining stages in the sequence, with the final sentence outcome being determined at the last step.

In this study we implement a multivariate multilevel approach to provide a more accurate and comprehensive understanding of the sentencing practice in England and Wales under the new guidelines. Multivariate models (i.e. models examining more than one dependent variable simultaneously) have been implemented in the US sentencing literature to study the indirect effects of certain case characteristics on the final sentence through pre-trial decisions such as prosecutorial charging decisions, bail decisions or assessments of offence seriousness carried out by probation officers (Albonetti, 1998; Kautt, 2009; Spohn et al., 2014; Ward et al., 2016). Our focus is on the series of decisions strictly made by the judge during the sentencing process, and on the multiple effects that case characteristics might have through that process.

Drawing on assault data from the 2011 round of the Crown Court Sentencing Survey we examine judicial decisions surrounding the initial rating of the seriousness of the case, what – if any – reductions were made for a guilty plea and the final disposal that was imposed. Modelling these decisions using a multivariate approach allows us to identify the specific influence of each guideline factor across different steps in the sentencing process, as well as on the final sentence outcome. The multilevel part of the model allows us to estimate unwarranted disparities between courts at different steps as well as whether any court-level disparities are carried through the steps structure. Ultimately, the evidence obtained from this novel approach allows us to assess whether Crown Court judges comply with the specifications of the sentencing guidelines.

In the next section a more detailed explication of the England and Wales sentencing guidelines is offered, along with a discussion of the key differences with US guidelines. This is followed by a description of the dataset and the special characteristics of multivariate models. After presenting the main findings from the research, we conclude by considering how our results contribute to understanding the form in which decisions are being made, and the lessons that could be applied in the design of future guidelines.

Consistency of Approach through a Sequence of Steps

In 2003 England and Wales became the first jurisdiction outside the USA to design and implement a comprehensive system of sentencing guidelines (Ashworth and Roberts, 2013). Since then many others have taken – or are considering taking – steps in a similar

direction. In essence, guideline systems are all based on the same idea: wide judicial discretion is undesirable and should be constrained to promote consistency in sentencing; that is, like cases being treated alike (Sentencing Council for England and Wales, 2011; Tata and Hutton, 1998). However, guidelines vary widely in terms of the nature and degree of the restrictions imposed (Reitz, 2013; Roberts, 2012).

In the USA, sentencing guidelines are typically structured in the form of a grid, with the range of possible sentence outcomes prescribed as a function of the number of previous convictions and offence seriousness. The grid-based approach to promoting consistency has been heavily criticized for undermining the principles of proportionality and individualization by restraining judges' autonomy and limiting their ability to distinguish between factually different cases (Alschuler, 1991; Cooper, 2013; Lowenthal, 1993; Schulhofer, 1991). This apparent trade-off between consistency on the one hand, and proportionality and individualization on the other, was noted in 2008 by the Sentencing Commission Working Group, which firmly rejected the implementation of US-style sentencing grids in England and Wales (Ashworth and Roberts, 2013). Instead of emphasizing the 'consistency of outcome' sought by the US guidelines, the Working Group recommended the formulation of guidelines seeking to achieve 'consistency of approach', understood as the extent to which all judges follow the same principles throughout the sentencing process (Hola, 2012; Hutton, 2013; Krasnostein and Freiberg, 2013).

The Sentencing Council for England and Wales devised a system of offence-specific guidelines structured by a sequence of steps. At each of these steps sentencers must consider specific lists of case factors based on the rationale that 'if all courts follow the same methodical approach to considering characteristics of the offence and the offender, greater consistency and fairness will ensue' (Ashworth and Roberts, 2013: 9). To illustrate this step-structure, we will consider the guideline covering the most common offences of assault,[1] which has formed the template for most subsequent guidelines (Hutton, 2013; Roberts et al., 2018).

The guideline is structured in nine sequential and non-overlapping steps, of which the first two are most critical. At Step One sentencers must determine the seriousness of the case. One of three categories of seriousness is assigned, reflecting both the harm inflicted to the victim and the culpability of the offender. To determine the appropriate level of seriousness sentencers refer to an exhaustive list of 'principal factual elements' of the offence, including the extent of victim injury and vulnerability (harm factors), and evidence of premeditation and offender role (culpability factors). The central status of Step One in determining the final sentence is reflected in the fact that each category of seriousness is assigned a specific range of 'appropriate' sentence outcomes (Ashworth and Roberts, 2013).

In Step Two sentencers must choose a specific preliminary sentence for the offence, based on the range of appropriate sentence outcomes associated with each level of seriousness. Sentencers may also 'fine tune' the level of harm and culpability at this stage in light of further aggravating and mitigating factors. These factors do not form the principal factual elements of the offence, instead providing 'the context of the offence and the offender'. For example, sentencers may consider the location and timing of the offence, whether the offender was on bail, whether they show remorse and whether the crime was deemed to be an isolated incident. In some cases, this fine-tuning may make it

appropriate for the court to select a sentence that is outside the usual range for offences of that seriousness category. Importantly, while the list of Step One factors provided in the guideline is comprehensive, those in Step Two are not.

There are seven remaining steps to be followed. Step Three reminds sentencers about the possibility to reduce the severity of the sentence in cases where the offender has provided assistance to the prosecution or police. Step Four allows for reductions in sentence severity for offenders pleading guilty (the magnitude of the reduction should be directly proportional to the stage at which the guilty plea was entered). Step Five reminds sentencers about the possibility of extending sentences in accordance with the dangerousness of certain offences (such as grievous bodily harm or racially/religiously aggravated wounding). Step Six invokes the totality principle – relevant when the offender is being sentenced for more than a single offence, or where the offender is currently serving a sentence. Step Seven reminds sentencers to consider making a compensation order and/or any other ancillary orders. Step Eight requires sentencers to explain the effect of the sentence and Step Nine reminds sentencers to discount any remand time already served by the offender. Finally, based on the previous nine steps, sentencers come to a final determination of the appropriate sentence to be awarded.

At present there are 15 offence-specific guidelines in operation in England and Wales, structuring the sentencing process of more than 149 offence types,[2] and it is expected that by 2020 the Sentencing Council for England and Wales will have issued guidelines for all common offences. However, until recently, the almost complete absence of empirical assessments of these new guidelines cast doubts over what, if any, changes in sentencing practice had resulted from their introduction (Ashworth, 2013; Ashworth and Roberts, 2013; Padfield, 2013). The release of the Crown Court Sentencing Survey – a government dataset describing cases processed in the Crown Court – has reversed this trend, transforming the landscape from one dominated by theoretical commentaries to one that is more evidence based (see, for example, Belton, 2018; Fleetwood et al., 2015; Irwin-Rogers and Perry, 2015; Lightowlers, 2018; Lightowlers and Pina-Sánchez, 2017; Maslen, 2015; Maslen and Roberts, 2013; Pina-Sánchez, 2015; Pina-Sánchez and Grech, 2018; Pina-Sánchez and Linacre, 2013, 2014; Pina-Sánchez et al., 2016; Roberts, 2013a; Roberts and Bradford, 2015; Roberts and Pina-Sánchez, 2014; Roberts et al., 2018).

Most of the above cited studies have relied on regression modelling techniques to explore topics such as compliance with the guidelines, consistency in sentencing or the specific effect of certain guideline factors. For example, Pina-Sánchez and Linacre (2013) look at custodial sentence length for offences of assault, refuting the widely held proposition that there are extensive unwarranted disparities in sentencing between courts, while Pina-Sánchez and Linacre (2014) demonstrated that the application of the new assault guideline has led to an increase in consistency. Roberts and Bradford (2015) showed that there are aggravating and mitigating factors having an undue influence on the guilty plea discount, which raises questions about compliance with the guidelines. Irwin-Rogers and Perry (2015) showed that Step One factors (those used to determine the seriousness of the offence) are more important than other mitigating and aggravating factors, with the exception of previous convictions.

Yet, these same studies have restricted their focus to a single outcome, typically the final awarded sentence. As such, researchers are implicitly assuming that the sentencing

process can be represented as the product of a unique data generating mechanism. That is, the step structure present in the English sentencing guidelines with its sequence of intermediate decisions is ignored.

Data

Our analysis is based on the Crown Court Sentencing Survey (CCSS), a comprehensive survey conducted by the Sentencing Council for England and Wales from 2011 to 2015 to monitor the effect of the newly introduced guidelines. The CCSS records the factual elements of each case in unprecedented detail, covering the seriousness of the offence, the aggravating and mitigating factors present in the case, the stage at which a guilty plea was entered, the sentence reduction granted and the number of previous convictions taken into account by the sentencer. The CCSS therefore enables more accurate and informative research on sentencing than has previously been possible. Crucially, by capturing factors considered and decisions made at different stages of the process, the CCSS can be used to shed new light on the functioning of the step-structure of the England and Wales guidelines.

The specific court location in which each sentence was passed is also recorded, although this information is only available for offences sentenced at the Crown Court in 2011. At that time, the only offence that had been structured according to the new design of the Sentencing Council for England and Wales was the assault guideline, which came into force in June 2011. As a result, we focus on cases of assault recorded by the CCSS from June 2011 until the end of the year, a total of 2851 cases of assault sentenced within 74 court locations. The mean and range of all the variables used are shown in Appendix I.

Sentencing decisions

Decisions following three specific steps in the guidelines are fully covered by the CCSS. These are: the assessed level of seriousness (determined in Step One, and 'fine tuned' in Step Two); the magnitude of the guilty plea reduction (Step Four); and the final sentence imposed (determined after all steps have been considered). The level of seriousness has three levels, reverse coded with one indicating the least serious offence rating (representing the guidelines' category 3 seriousness), and three the most serious (category 1). Five categories are used to describe the size of the discount awarded as a result of entering a guilty plea: '1%–10%'; '11%–20%'; '21%–32%'; '33%'; and 'greater than 33%', with a higher score indicating a larger reduction. The final sentence outcome is composed of four disposal types coded from one to four: discharges and fines (grouped together because of their small sample sizes); community orders; suspended sentences; and custodial sentences. Operationalizing final sentence outcome in this way, as opposed to focusing on custodial sentence lengths, allows us to include all sentenced offences, irrespective of whether a custodial sentence was imposed.

Explanatory variables

The CCSS also covers most of the factors that sentencers are meant to consider when deciding each of the above outcomes. From Step One we identify a total of 18 principal

factual elements of assault cases.[3] All 28 factors listed in Step Two of the guidelines regarding the context of the offence and the characteristics and circumstances of the offender are also included. These cover both 'factors increasing seriousness', and 'factors reducing seriousness or reflecting personal mitigation'. At Step Four, we identify both whether the offender pleaded guilty at the first opportunity, and the specific stage at which the guilty plea was entered. These stages can be loosely classified from the soonest to the latest as follows: 'at magistrates' courts', 'prior to Plea and Case Management Hearing' (PCMH), 'at PCMH', 'after PCMH but before the day of the trial' and 'on, or after, the day of the trial'. Finally, we also include details of the specific type of assault case, distinguishing between common assault, grievous bodily harm, grievous bodily harm with intent and assault with bodily harm. These sets of explanatory variables were selected to reflect the guidelines' provisions as accurately as possible.

Limitations

In spite of its remarkable detail and coverage, the CCSS is also prone to problems of unit and item missing data affecting its validity. The average response rate in 2011 was an acceptable 61%, however, this rate varied widely across Crown Court locations, ranging from 95% to 20% (Sentencing Council for England and Wales, 2012). It is possible that judges who disregarded the data collection requirements of the Council's research team were also less compliant with the guidelines issued by the same institution, which would generate a problem of selection bias (Bushway et al., 2007), affecting the generalizability of our findings. The extent of this problem is hard to assess. Analyses carried out by the Sentencing Council for England and Wales found that the available auxiliary data (the date when the sentence was imposed, and the location of the court where the sentence was imposed) could not predict more than 20% of the missingness detected. Furthermore, the questionnaire format leaves open the potential for substantial misclassification errors that may bias regression estimates (Küchenhoff et al., 2006). This is because sentencers were required to tick whether each sentencing factor was present, in each case, with no available option for 'not present' or 'not applicable'. Given that the survey is self-completion,[4] if judges forget to tick a certain case characteristic, it will be counted as non-present, potentially leading analysts to underestimate the extent that these features may impact on sentencing decisions.

A number of scholars have also criticized the CCSS's mode of data collection on different grounds, arguing that the reported case characteristics can be manipulated by the judge when filling in the form to demonstrate compliance with the guidelines. While this may be possible, we believe that this problem may have been exaggerated, not least because if there was such a strong incentive to appear to be complying with the guidelines, we should expect it to be clearly evident in judges' sentencing statements. And since the identity of the judge cannot be retrieved directly from the questionnaire form, any incentives to misreport the factual elements of the case taken into consideration will be limited. Added to this, the computational burden involved in covering up any conscious biases or departures from the guidelines by systematically ticking different combinations of case characteristics in a form that covers more than 100 items is the sort of optimization task that can be expected to be performed by computers, not by 'rationally bounded agents'.

Multivariate Multilevel Models for the Analysis of Sentencing Decisions

To account for the step-based structure of the new sentencing guidelines we use a multilevel multivariate modelling approach (Goldstein et al., 2009; Kautt, 2009). This allows us to examine simultaneously three key stages in the sentencing process: the assessed level of seriousness (Step One); the magnitude of the guilty plea reduction (Step Four); and the final sentence outcome awarded. Ideally, we would have included a fourth outcome model to deal directly with Step Two of the guideline too. However, the outcome of that step, 'the preliminary within-range sentence', was not captured by the CCSS. We therefore include all Step Two factors as predictors of the final sentence outcome. The outcome values associated with the decisions made at each of these three stages are modelled using an ordinal logistic regression to reflect the ordered nature of these outcome values (see earlier section on 'Sentencing decisions'), with the variables discussed in the earlier section on 'Explanatory variables' as the set of explanatory variables. In addition, we have also included a court-level random intercept for each of the three outcomes to account for the hierarchical structure of the data whereby offences are grouped within specific courts. These court-level random intercepts can be used to explore the level of between court inconsistencies at each step. Furthermore, decisions made following Step One and Step Four are linked to the final sentence outcome by including them as regressors predicting the final sentence. A full technical explanation can be found in Appendix II.

This joint modelling approach accounts for the sequential organization of the sentencing steps as defined in the guidelines. To identify potential undue effects of certain factors outside the steps where the guidelines indicate that they must be considered, we also estimate a second model (Model 2) where all guideline factors are included in the specification of Step One, Step Four and the final sentence. In addition to detecting effects that do not comply with the specifications of the guidelines, this allows us to capture indirect effects of step-specific factors on the final sentence outcome, and provide more accurate estimates of the total effect of different guideline factors on the final sentence.

Our analysis is placed within the Bayesian framework, which allows us to develop complex multivariate and multilevel models for analysing the sentencing data arising from a complex, sequential process. Another benefit of using the Bayesian approach here is that we can straightforwardly estimate interpretable and relevant quantities (such as the intra-class correlation coefficients) that are non-linear transformations of some model parameters. We use vague priors (or so-called uninformative priors) for all the model parameters so that the prior specification plays a minimal role in the posterior estimates of the model parameters. In other words, we 'let the data speak for themselves'. Both models are estimated using WinBUGS (Lunn et al., 2000). Appendix III shows the code used to implement Model 1 as well as the prior specification for the parameters in that model. For fitting each model, two MCMC (Markov chain Monte Carlo) chains with dispersed starting values were run for 60,000 iterations for each chain. The first 10,000 iterations of each chain were discarded as burn-in. Thus, a total of 100,000 iterations (50,000 from each) were used to provide the posterior summaries presented next.

Results

The full results from Model 1 are presented in Appendix IV; a summary of those results – including main effects for factors associated with Step One (assessed seriousness), Step Four (guilty plea reduction) and final sentence outcome – is displayed in Figure 1. The size of the bars included in Figure 1 represent the effect size of each of the factors considered, measured in lo. Black bars indicate that their effects were found to be statistically significant,[5] grey bars indicate non-significant effects. Looking first at the assessed level of offence seriousness, all principal factual elements operate in the expected direction and almost all are significant. For example, cases tend to be rated as more serious when there was evidence of *serious injury*, a *weapon* was involved, the defendant was identified as playing a *leading role in a gang* or when the victim was particularly *vulnerable*. Conversely, cases are rated as less serious when there was no evidence of *premeditation*, it was judged to be *self-defence* or *injuries* were less serious. A consistent picture is also evident when the determinants of a guilty plea reduction are considered, with significant sentence reductions if the guilty plea was entered at the *first opportunity*. As specified in the guilty plea guidelines,[6] the specific timing of the guilty plea is also important, with larger reductions if it was entered *at the magistrates' court*, or *prior to a Plea and Case Management Hearing*.

More complex results are seen when the final sentence outcome is considered. Here, we allow the final sentence to be informed by the outcomes of decisions made at Step One and Step Four, as well as the type of offence and other aggravating and mitigating factors listed in Step Two of the guideline. Most factors are significant, and all operate in the expected direction. However, the level of guilty plea reduction does not appear to be closely related to the final sentence outcome.[7] This may reflect the reliance on percentage reductions in the phrasing of the guilty plea reduction, giving the impression it is only applicable to adjustments in custodial sentence length (Pina-Sánchez and Linacre, 2013, 2014), or fine amount.[8]

The effect of the initial assessed level of seriousness is also weaker than expected given the primacy given to this step in the sentencing guidelines (Hutton, 2013; Maslen and Roberts, 2013; Roberts, 2013b). For example, the odds ratio to be sentenced to prison rather than getting a suspended sentence is 4.48 if the seriousness is identified as 'medium' rather than 'low', which is smaller in magnitude than the effects of four Step Two aggravating factors, having *four or more previous convictions*, committing the offence while *on bail* or *on licence* and *failure to comply with court orders* (with odds ratios of 5, 5.21, 6.36 and 5.05, respectively). This is consistent with recent findings from Irwin-Rogers and Perry (2015) when considering cases of domestic burglary, suggesting that Step Two factors such as previous convictions are not only used to 'allow a tailoring' (Edwards, 2013: 84) or 'fine tune' the final sentence within the given category, but instead they are used to select what that category will be. We also find that aggravating factors from Step Two play a more substantial role in shaping the final sentence than the equivalent mitigating factors, a result anticipated by Cooper (2013) when he expressed concerns regarding the emphasis placed on harm and culpability factors by the guidelines.

Turning to the random part of the model, we find evidence of small differences between courts, with 4.6% of the variation in final sentence outcomes attributed to between court
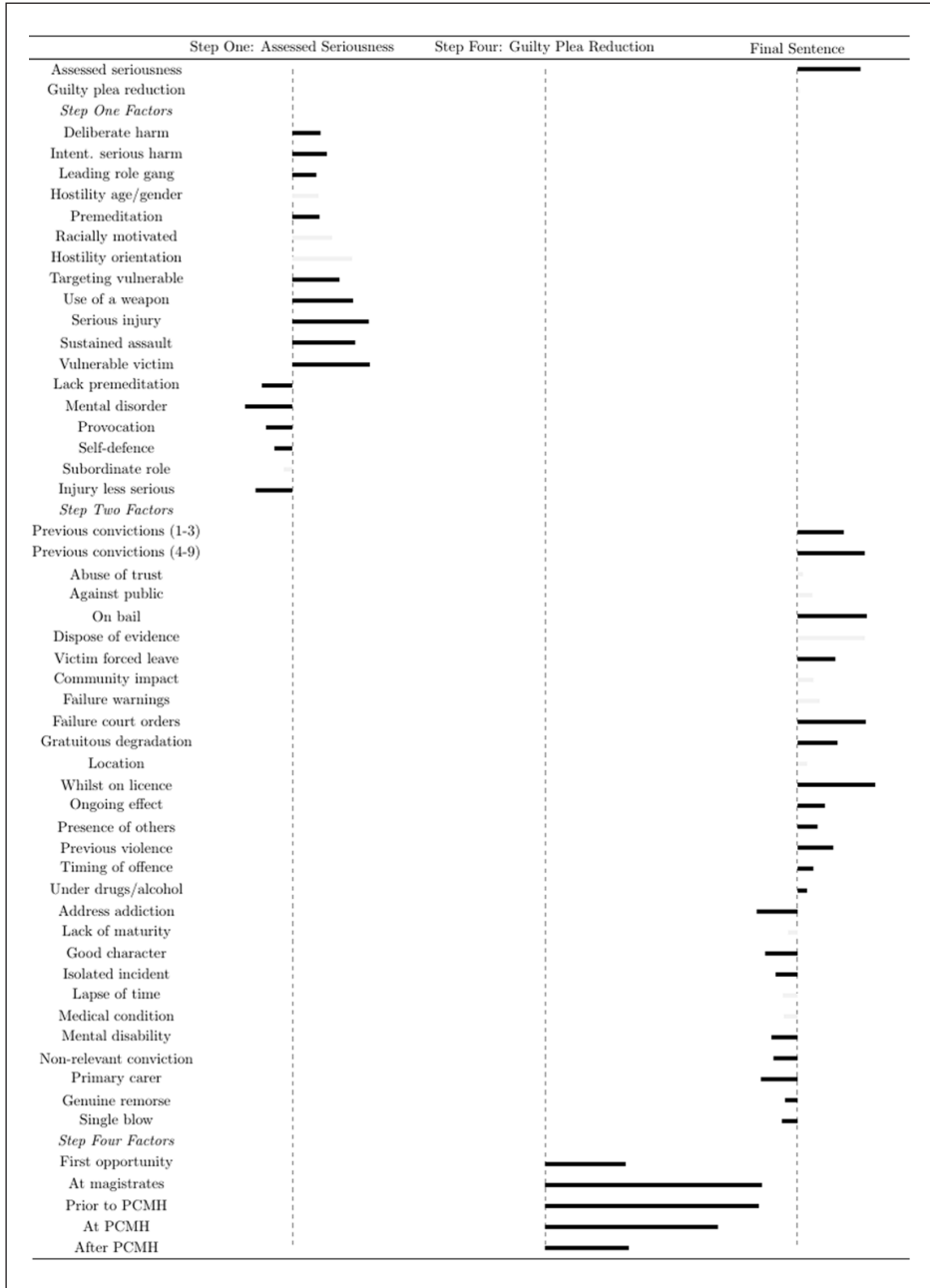
Figure 1. Visual summary of Model 1 (multilevel multivariate without mediating effects).[a]

[a]Black bars represent statistically significant factors (95% significance level), grey bars represent non-significant factors.

differences, with similar estimates for Step One and Step Four, 3.8% and 5.3%, respectively. The covariances between the random intercepts are found non-significant. The implication is that having accounted for various case-level explanatory variables, court disparities are not carried across the step structure. Two important new insights follow from this. First, there are no courts that show either a systematically harsh or systematically lenient approach throughout the sequence of steps. That is, courts that tend to judge offences at Step One as unduly more serious than the rest of the courts, are not necessarily making more restrictive guilty plea reductions at Step Four, or opting for a more severe final sentence outcome than would be expected based on the case characteristics. Second, courts are not trying to compensate for unduly harsher or lenient judgements through the sequence of steps – for example by providing generous reductions at the guilty plea stage to compensate for disproportionately harsh judgements on the seriousness of the case at Step One – if that was to be the case, the covariance between the random intercepts in Steps One and Four would be estimated to be positive and away from 0. This finding refutes, to a certain extent, the data quality criticism pointing at a gaming strategy conducted by judges when filling in survey questionnaires[9] (covered in the earlier section on 'Limitations').

Model 1 makes the following two assumptions. First, the level of seriousness and the guilty plea reduction are assumed to be defined strictly by Step One and Step Four factors. Second, the factors at Step One and those at Step Four do not directly influence the final sentence outcome but, instead, they operate via their influence on the judged level of offence seriousness and any guilty plea reduction. Model 2 relaxes both assumptions by allowing all guideline factors to influence the judicial decisions at each and every one of the three stages. The full results of this model are presented in Appendix V. A summary of these results is presented visually in Figure 2.

Here we find clear evidence that many Step One factors have a direct effect on the final sentence outcome; in addition to their indirect effects on sentence outcomes via modifications of the assessed level of offence seriousness. For ease of comparison the direct, indirect and total effects of Step One and Two factors on the final sentence outcome are included in Appendix VI. For example, for offences involving *the use of a weapon*, we observe a direct impact on the final sentence of .70, in addition to the indirect effect via an increased level of seriousness ($1.38 \times .90 = 1.24$). In fact, for nine out of the 13 Step One factors that were found to be significantly influencing the level of seriousness, we also identify an additional direct effect on the final sentence outcome. However, the relative hegemony of Step One over Step Two factors does not apply uniformly, with some Step Two factors such as *on bail* or *while on licence* having a stronger effect than three-quarters of the Step One factors. Step Two factors, then, seem to be used for more than just 'fine-tuning' the starting point. The timing of a guilty plea also appears to have a direct effect on final sentence outcomes, over and above the sentence reductions made at Step Four.

Model 2 also identifies an undue influence of some Step One and Step Two factors on guilty plea reductions (Step Four). This is despite the clear instruction not to take into consideration factors other than the timing of a guilty plea and whether the offender was caught red handed at this stage of the sentencing process. For example, those offences where the offender is deemed to be of *good character* and/or shows *genuine remorse*
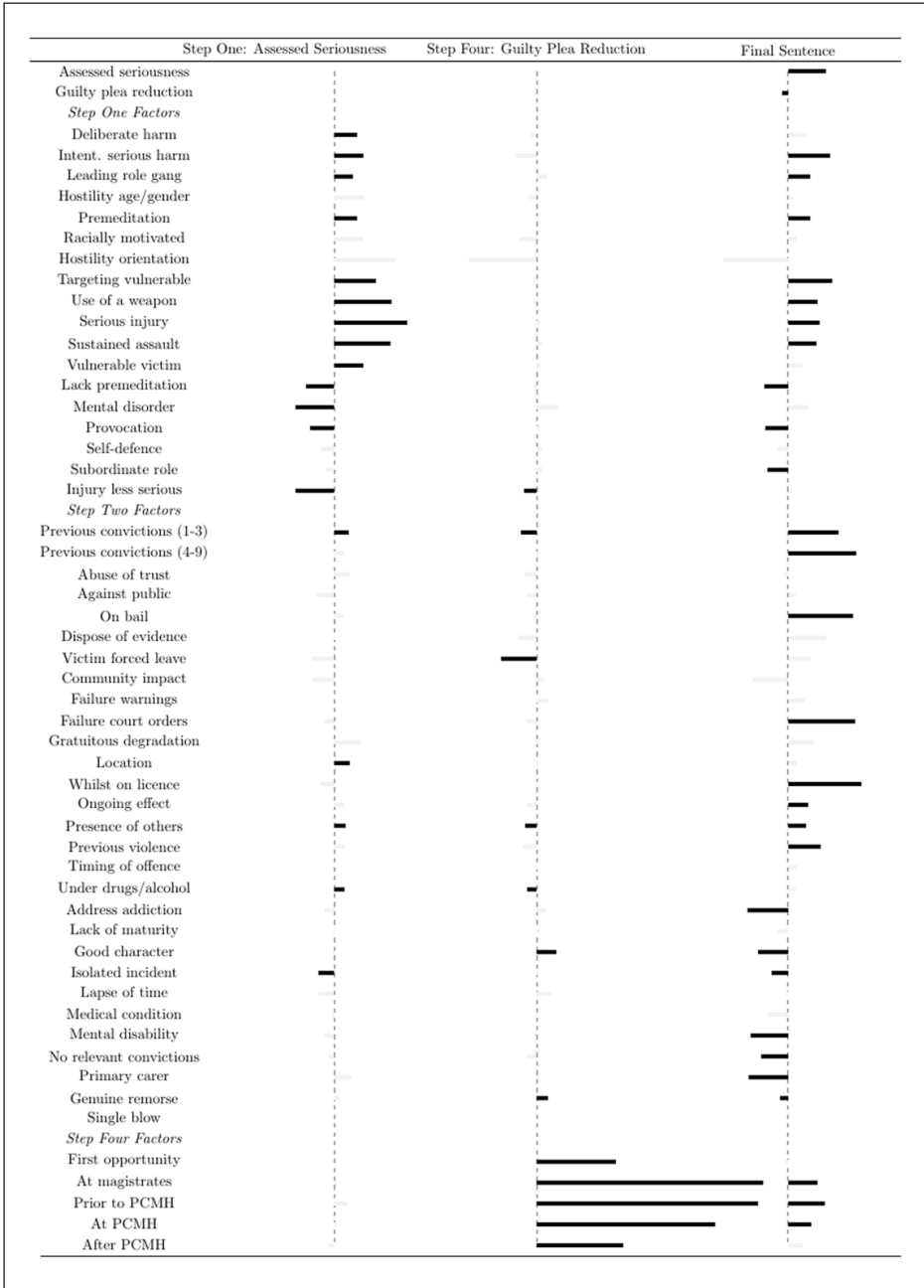
Figure 2. Visual summary of Model 2 (multilevel multivariate with mediating effects).[a]
[a]Black bars represent statistically significant factors (95% significance level), grey bars represent non-significant factors.

receive more generous discounts, while assaults caused in the *presence of others*, or where the *victim is forced to leave* their home receive more modest guilty plea reductions. However, these undue influences are small in magnitude, particularly in comparison to the large sentence adjustment associated with the timing of the plea. We also find evidence of weak – yet undue – effects of some Step Two factors on the judged level of offence seriousness. Importantly, three of these factors – *previous convictions*, in the *presence of others*, and under the *influence of drugs/alcohol* – were also found to be significant in the determination of the guilty plea reduction, illustrating the problems that judges are facing to apply them correctly.

Discussion

This article's contributions are twofold. First, it demonstrates the need to replace univariate regression models as the default approach for the study of sentencing with more realistic multivariate models. These models are capable of investigating the complex network of intermediate decisions made throughout the sentencing process, allowing researchers to identify the undue influence that some offence characteristics may be having, and estimate more accurately their total effect on the final sentence. Second, it provides new insights regarding the considerations made by Crown Court judges when sentencing cases of assault, establishing the level of compliance with the new guidelines, and identifying specific areas in need of attention.

In general, our findings demonstrate that sentencing in the England and Wales Crown Court is both highly complex, yet thoughtfully carried out and consistent across courts. We have studied the use of 47 different case characteristics relevant to assault offences, finding 35 of them to be statistically significant, and all of them operating in the expected direction. The timing of the guilty plea was the strongest predictor of guilty plea reductions. Step One factors indicating higher harm or culpability increased the level of seriousness, while factors indicating lower harm and culpability were associated with lower assessed seriousness. Furthermore, the guidelines' identification of a comprehensive list of factors to be considered at Step One, together with the availability of these factors in the CCSS, has allowed us to take account of all relevant sentencing factors in the determination of the level of seriousness, enabling us to distinguish with unprecedented accuracy between warranted and unwarranted disparities, and in so doing obtain more robust estimates of inconsistencies between courts.

We found no substantial evidence of unwarranted between court disparities in deciding the level of seriousness, guilty plea reductions or the final disposal type given with less than 5% of the unexplained variability at each sentencing step attributed to systematic differences between Crown Court locations. Furthermore, none of the detected court disparities at each step were related to each other, suggesting that extreme decisions made at earlier steps are not exacerbated – or ameliorated – across the step structure. Considering the variety of steps and factors listed, and the degree of judicial discretion offered in each of them – for example, no specific weights are attached to the harm and culpability factors used to decide the level of seriousness (Cooper, 2013; Dhami, 2013a) – these results suggest that consistency in sentencing and judicial discretion are not

mutually exclusive goals. In fact, it seems that England and Wales may be doing a reasonable job of reconciling them.

However, our results also show that contrary to the instructions in the guidelines – which call for each step to be considered sequentially and in a non-overlapping fashion – many guideline factors are being unduly considered outside their specified step. We identified substantial indirect effects of many Step One factors on the final sentence outcome after having controlled for the level of seriousness. This may, in part, be explained by a clause opening the possibility of Step One factors being considered at Step Two: '[a] case of particular gravity, reflected by multiple features of culpability in step one, could merit upward adjustment from the starting point before further adjustment for aggravating or mitigating features' (Sentencing Council for England and Wales, 2011: 5, 8, 12, 17, 21 and 24). Regardless of the cause, an important methodological implication follows from this finding. To assess the influence of relevant aggravating and mitigating circumstances on sentencing decisions correctly, account should be taken of both direct and indirect influences of each factor. For example, in our analysis the three most influential factors on final sentence outcomes are *serious injury, targeting* a *vulnerable* victim and *use of a weapon*, all Step One factors. The total effect (considering both direct and indirect influences) of each of these factors is bigger than that of *previous convictions*, suggesting a more complex picture than was evident in Irwin-Rogers and Perry (2015).

Our results also identify the presence of undue effects in the determination of the level of seriousness and guilty plea reduction. This is particularly problematic given the exhaustive lists of factors that judges are required to use in these steps. Consistent with recent work from Roberts and Bradford (2015), we show that decisions on the magnitude of a guilty plea reduction are being influenced by more than the timing of the plea. But we also extend this work by identifying the specific factors responsible for such undue effects, which we classify in two main groups, factors depicting personal mitigation and others suggesting domestic violence.

We propose four mechanisms which might be behind the undue influence that some factors are having across the whole sentencing process, although it is possible that there are many more. First, some of the factors included in the guideline are either not sufficiently clear or legitimized. For example, the *timing* or the *location* of the offence are applied to account for very diverse settings without clear guidance about which locations might represent an aggravating factor. And there is considerable debate over the extent that being under the influence of *drugs or alcohol* should be used as an aggravating factor (Dingwall, 2006; Lightowlers and Pina-Sánchez, 2017; Padfield, 2011). Second, the treatment of *previous convictions* as a Step Two factor in spite of its vast importance in deciding sentence severity (Roberts and Pina-Sánchez, 2014) seems to be something against which judges are rebelling. This is demonstrated by the fact that *previous convictions* was one of only two factors (out of the 47) to have a statistically significant effect on Step One, Four and on the final sentence outcome. Third, guilty plea reductions seem to be affected by redeeming characteristics observed in the offender such as expression of genuine *remorse* or *exemplary conduct/character*. Fourth, Step Two factors related to domestic violence – such as victims being forced to *leave their home* or assault committed in the *presence of others* including relatives – appear to be having an undue influence

on the determination of guilty plea reductions and the level of seriousness. This suggests frictions in the application of the assault guidelines for cases of domestic violence.

The potential presence of such mediated effects was already highlighted by Dhami (2013b) in her critique of the more loosely defined previous sentencing guidelines. But the fact that they were also detected under the more clearly structured and newly designed sentencing guidelines should be a cause of concern for the Sentencing Council for England and Wales. This research is, however, timely since the Council is currently undertaking a revision of the new assault sentencing guidelines. In light of our findings we suggest the following recommendations.

First, either remove disputed or unclear case characteristics (such as *timing, location* or under the *influence of drugs/alcohol*) from the list of factors to be considered by sentencers, or clarify how they are meant to be applied. This could expedite the decision-making process and improve consistency in sentencing. Second, include *previous convictions* as a Step One factor associated with an increase in culpability. *Previous convictions* were shown to play a substantial role in sentencing decisions among our sample, and moving this to Step One would bring the guideline more into line with current sentencing practice. Third, in 2017 a new guilty plea guideline was published highlighting the specific requirement not to consider personal mitigating factors such as showing *remorse*. We recommend expanding this request in future versions of the guilty plea guidelines to include a specific reference to offenders deemed of *good character* and/or exemplary conduct, given that this factor was shown to have an undue effect almost twice as large as that observed for *remorse*. Finally, the specific factors defining cases of domestic violence and the detected undue use of these factors, demonstrates the difficulty of sentencing these types of cases under the current assault guidelines. This problem could be solved by grouping offences of domestic violence under a distinct offence type, or perhaps even under distinct sentencing guidelines.

Further methodological discussions

The new and more robust insights into the sentencing process that we have reported here demonstrate the value of abandoning the standard framework in the field based on univariate regression models and opting for more encompassing, insightful and robust multivariate approaches. By incorporating as many of the preliminary decisions made in the sentencing process as possible, we have been able to explore the complexities of such process in more detail. This also results in more statically efficient estimates than if those same outcomes were estimated independently (Park et al., 2015). More importantly, by accounting for the possibility that some factors (case characteristics, or other extralegal factors) have indirect effects on the final sentence outcome via their influence on earlier considerations, we can obtain more robust estimates of their full effect on sentences. Kautt (2002) identified the introduction of multilevel modelling techniques as the '5th wave' of sentencing research – following Zatz's (1987) famous distinction of methodological breakthroughs in the field. Here we suggest pushing the methodological frontier once again, and give rise to a '6th wave' based on the adoption of multivariate modelling techniques.

Access to data on preliminary outcomes – such as the assessment of the offence seriousness, or the guilty plea discount – and the clear identification of those outcomes as

preliminary stages in the sentencing process, has facilitated the examination of sentencing in England and Wales. However, our approach could be replicated in other jurisdictions, particularly those where systems of guidelines specify the types of factors to be considered at different stages of the process. The number of such jurisdictions is growing steadily. In addition to the well-known federal guidelines in the USA and those operating in some US states, there are now similar systems introduced in South Korea and others under consideration in Canada, Scotland, Northern Ireland, New Zealand, some Australian states, the Gulf states and Israel (Roberts et al., 2018). Future quantitative studies of the sentencing process in those jurisdictions will benefit from the more comprehensive and robust analytical framework developed here.

Finally, it is worth highlighting the central role that the CCSS has played in advancing sentencing research in England and Wales. At the outset, we drew attention to the new wave of studies that resulted from the publication of the CCSS, transforming sentencing research in England and Wales into a more empirically based discipline. These recent contributions have shed new light on important topics such as proportionality (Fleetwood et al., 2015); consistency (Pina-Sánchez, 2015; Pina-Sanchez and Grech, 2018; Pina-Sánchez and Linacre, 2013, 2014), individualization (Roberts et al., 2018), severity (Pina-Sánchez et al., 2016), compliance with the guidelines (Roberts, 2013a; Roberts and Bradford, 2015) or the use of different case characteristics (Belton, 2018; Irwin-Rogers and Perry, 2015; Lightowlers and Pina-Sánchez, 2017; Maslen, 2015; Maslen and Roberts, 2013; Roberts and Pina-Sánchez, 2014). Added to this, we have demonstrated that unique features of the CCSS can also be exploited to estimate accurately: (1) the specific contribution of each case characteristic on the sentence outcome, something that was previously considered impossible by many sentencing scholars, and the CCSS founders at the Sentencing Council for England and Wales;[10] (2) the degree of unwarranted disparities at different stages of the sentencing process – for the case of Step One decisions those disparities were estimated after controlling for all relevant case characteristics; and (3) the undue influence that different case characteristics might have across the sequence of steps configured in the guidelines, which opens a new and more insightful approach to the study of compliance with the guidelines.

The broad range of new evidence and methodological breakthroughs afforded by the CCSS makes its discontinuation in 2015 regrettable and short-sighted. We believe that there is a clear need for the Council to facilitate access to new datasets so that the current momentum in sentencing research does not fade away. Since 2016 the Council has commissioned more ad hoc surveys seeking to assess the impact of new guidelines as they were progressively introduced. These surveys have focused on specific offence types, narrower time-frames and lower courts, while maintaining the same level of detailed that characterizes the CCSS. Dropping the goal of collecting census-level sentence data will no doubt limit the types of analyses that could be done (e.g. time-series or any kind of wide-reaching longitudinal approach will not be available). On the other hand, the change of focus of some of these new surveys to the magistrates' court does offer the possibility to move beyond analyses of the sentencing practice as carried out at the Crown Court to explore the main courts where sentencing takes place. Important questions regarding differences in severity, consistency, individualization, the use of different

case characteristcs, and compliance with the guidelines across the two courts remain to be answered. To do so it will be crucial that the Sentencing Council for England and Wales proceeds to publish these new datasets online or creates the necessary channels to share these new datasets to all interested researchers.

Acknowledgements

Declaration of Conflicting Interests

Funding

Notes

1. The assault definitive guideline is available at: https://www.sentencingcouncil.org.uk/publications/item/assault-definitive-guideline/.
2. Excluding the Magistrates' Courts Sentencing Guidelines.
3. One of the factors indicating higher culpability, 'offence motivated by hostility towards the victim based on the victim's disability', was dropped from the analysis because it was only noted in six of the 2851 cases captured in our sample.
4. Available at: https://www.sentencingcouncil.org.uk/wp-content/uploads/Assault_and_Public_Order_-_April_2014.pdf.
5. Since *p*-values do not make sense in a Bayesian framework, to assess the statistical significance for each of the estimates reported we use their 95% credible interval. Specifically, we check whether the 95% credible interval includes zero, which can be interpreted as a non-statistically significant estimate.
6. The guidelines structuring the guilty plea in operation in 2011 and the latest version published in 2017 are available at: https://www.sentencingcouncil.org.uk/publications/?type=publications&cat=definitive-guideline&s&topic=guilty-plea.
7. Additional models treating guilty plea reduction as categorical confirm this non-significant effect (available on request).
8. The latest guilty plea guidelines published in 2017 explicitly indicate that guilty plea reductions can also be applied to change the disposal type.
9. To investigate this hypothesis robustly the court ids used here should be complemented with judge ids.
10. In his examination of the assault guidelines Hutton (2013: 98–99) notes that: '[j]udges are not required to provide details of their calculations but simply to list those factors which they have taken into account. It will therefore be impossible to find out the degree of influence which any individual factor had on the judicial assessment of seriousness.' Similarly, in his analysis of compliance rates Roberts (2013a: 111) indicates that: '[t]he full impact of the guidelines, and the way that courts use them remains hidden from view. For example, a court considering a factor at Step One of the guidelines methodology which is not specified

by the guideline would be an example of a court failing to follow a guideline, yet this form of non-compliance cannot be detected by a survey such as the CCSS.' Lastly, in its review of the CCSS, the Sentencing Council for England and Wales (2012: 19) indicated that: '[o]n the Crown Court Sentencing Survey forms, judges are requested to indicate which aggravating and mitigating factors had an influence on the sentence imposed. However, they are not requested to record the relative importance of these factors' and proceeded to analyse the influence of aggravating and mitigating factors on the sentence outcome by counting their frequency.

References

Albonetti CA (1998) Direct and indirect effects of case complexity, guilty pleas, and offender characteristics on sentencing for offenders convicted of a white-collar offense prior to sentencing guidelines. *Journal of Quantitative Criminology* 14(4): 353–378.

Alschuler AW (1991) The failure of sentencing guidelines: A plea for less aggregation. *University of Chicago Law Review* 58(3): 901–951.

Ashworth A (2013) The struggle for supremacy in sentencing. In: Ashworth A and Roberts JV (eds) *Sentencing Guidelines: Exploring the English Model*. Oxford: Oxford University Press, 15–30.

Ashworth A and Roberts JV (2013) The origins and nature of the sentencing guidelines in England and Wales. In: Ashworth A and Roberts JV (eds) *Sentencing Guidelines: Exploring the English Model*. Oxford: Oxford University Press, 1–12.

Belton I (2018) *The role of personal mitigating factors in criminal sentencing judgments: An empirical investigation*. Doctoral dissertation, Middlesex University, UK.

Bushway SD, Johnson BD and Slocum LA (2007) Is the magic still there? The use of the Heckman two-step correction for selection bias in criminology. *Journal of Quantitative Criminology* 23(2): 151–178.

Cooper J (2013) Nothing personal: The impact of personal mitigation at sentencing the creation of the Council. In: Ashworth A and Roberts JV (eds) *Sentencing Guidelines: Exploring the English Model*. Oxford: Oxford University Press, 157–164.

Dhami MK (2013a) Sentencing guidelines in England and Wales: Missed opportunities? *Law and Contemporary Problems* 76(1): 289–312.

Dhami MK (2013b) A 'decision science' perspective on the old and new sentencing guidelines in England and Wales. In: Ashworth A and Roberts JV (eds) *Sentencing Guidelines: Exploring the English Model*. Oxford: Oxford University Press, 165–181.

Dingwall G (2006) *Alcohol and Crime*. Abingdon: Willan Publishing.

Edwards I (2013) Victims, sentencing guidelines, and the Sentencing Council. In: Ashworth A and Roberts JV (eds) *Sentencing Guidelines: Exploring the English Model*. Oxford: Oxford University Press, 71–85.

Fleetwood J, Radcliffe P and Stevens A (2015) Shorter sentences for drug mules: The early impact of the sentencing guidelines in England and Wales. *Drugs: Education, Prevention and Policy* 22(5): 428–436.

Goldstein H, Carpenter J, Kenward MG, et al. (2009) Multilevel models with multivariate mixed response types. *Statistical Modelling* 9(3): 173–197.

Hola B (2012) Sentencing of international crimes: Consistency of case law. *Amsterdam Law Forum* 4: 3–24.

Hutton N (2013) The definitive guideline on assault offences: The performance of justice. In: Ashworth A and Roberts JV (eds) *Sentencing Guidelines: Exploring the English Model*. Oxford: Oxford University Press, 86–103.

Irwin-Rogers K and Perry TH (2015) Exploring the impact of sentencing factors on sentencing domestic burglary. In: Roberts JV (ed.) *Sentencing Guidelines: Exploring Sentencing Practice in England and Wales*. Basingstoke: Palgrave, 213–239.

Johnson BD (2006) The multilevel context of criminal sentencing: Integrating judge- and county-level influences. *Criminology* 44: 259–298.

Kautt PM (2002) Location, location, location: Interdistrict and intercircuit variation in sentencing outcomes for federal drug-trafficking offenses. *Justice Quarterly* 19(4): 633–671.

Kautt PM (2009) Heuristics influences over offense seriousness calculations: A multilevel investigation of racial disparity under sentencing guidelines. *Punishment and Society* 11(2): 191–218.

Krasnostein S and Freiberg A (2013) Pursuing consistency in an individualistic sentencing framework: If you know where you're going, how do you know when you've got there. *Law and Contemporary Problems* 76: 265–288.

Küchenhoff H, Mwalili SM and Lesaffre E (2006) A general method for dealing with misclassification in regression: The misclassification SIMEX. *Biometrics* 62(1): 85–96.

Lightowlers C (2018) Drunk and Doubly Deviant? The Role of Gender and Intoxication in Sentencing Assault Offences. *The British Journal of Criminology*.

Lightowlers C and Pina-Sánchez J (2017) Intoxication and assault: An analysis of Crown Court sentencing practices in England and Wales. *British Journal of Criminology* 58(1): 132–154.

Lowenthal GT (1993) Mandatory sentencing laws: Undermining the effectiveness of determinate sentencing reform. *California Law Review* 81(1): 61–123.

Lunn DJ, Thomas A, Best N, et al. (2000) WinBUGS – a Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and computing* 10(4): 325–337.

Maslen H (2015) Penitence and persistence: How should sentencing factors interact? In: Roberts JV (ed.) *Sentencing Guidelines: Exploring Sentencing Practice in England and Wales*. Basingstoke: Palgrave, 173–193.

Maslen H and Roberts JV (2013) Remorse and sentencing: An analysis of sentencing guidelines and sentencing practice. In: Ashworth A and Roberts JV (eds) *Sentencing Guidelines: Exploring the English Model*. Oxford: Oxford University Press, 122–139.

Padfield N (2011) Intoxication as a sentencing factor: Mitigation or aggravation. In: Roberts JV (ed.) *Mitigation and Aggravating at Sentencing*. Cambridge: Cambridge University Press, 81–101.

Padfield N (2013) Exploring the success of sentencing guidelines. In: Ashworth A and Roberts JV (eds) *Sentencing Guidelines: Exploring the English Model*. Oxford: Oxford University Press, 31–51.

Park R, Pituch KA, Kim J, et al. (2015) Comparing the performance of multivariate multilevel modeling to traditional analyses with complete and incomplete data. *Methodology* 11(3): 100–109.

Pina-Sánchez J (2015) Defining and measuring consistency in sentencing. In: Roberts JV (ed.) *Sentencing Guidelines: Exploring Sentencing Practice in England and Wales*. Basingstoke: Palgrave, 76–92.

Pina-Sánchez J and Grech D (2018) Location and sentencing: To what extent do contextual factors explain between court disparities. *British Journal of Criminology* 58(3): 529–549.

Pina-Sánchez J and Linacre R (2013) Sentence consistency in England and Wales: Evidence from the Crown Court Sentencing Survey. *British Journal of Criminology* 53(6): 1118–1138.

Pina-Sánchez J and Linacre R (2014) Enhancing consistency in sentencing: Exploring the effects of guidelines in England and Wales. *Journal of Quantitative Criminology* 30(4): 731–748.

Pina-Sánchez J and Linacre R (2016) Refining the measurement of consistency in sentencing: A methodological review. *International Journal of Law, Crime and Justice* 44: 68–87.

Pina-Sánchez J, Lightowlers C and Roberts JV (2016) Exploring the punitive surge: Crown Court sentencing practices before and after the 2011 English riots. *Criminology and Criminal Justice* 17(3): 319–339.

Reitz K (2013) Comparing sentencing guidelines: Do US systems have anything worthwhile to offer England and Wales? In: Ashworth A and Roberts JV (eds) *Sentencing Guidelines: Exploring the English Model*. Oxford: Oxford University Press, 182–201.

Roberts JV (2012) Structured sentencing: Lessons from England and Wales for common law jurisdictions. *Punishment and Society* 14(3): 267–288.

Roberts JV (2013a) Complying with sentencing guidelines: Latest findings from the Crown Court Sentencing Survey. In: Ashworth A and Roberts JV (eds) *Sentencing Guidelines: Exploring the English Model*. Oxford: Oxford University Press, 104–121.

Roberts JV (2013b) Sentencing guidelines in England and Wales: Recent developments and emerging issues. *Law and Contemporary Problems* 76: 1–25.

Roberts JV and Bradford B (2015) Sentence reductions for a guilty plea in England and Wales: Exploring new empirical trends. *Journal of Empirical Legal Studies* 12(2): 187–210.

Roberts JV and Pina-Sánchez J (2014) The role of previous convictions at sentencing in the Crown Court: Some new answers to an old question. *Criminal Law Review* 8: 575–588.

Roberts JV, Pina-Sánchez J and Marder I (2018) Individualisation at sentencing: The effects of guidelines and 'preferred' numbers. *Criminal Law Review* 2: 123–136 (online).

Schulhofer SJ (1991) Assessing the federal sentencing process: The problem is uniformity, not disparity. *American Criminal Law Review* 29: 833–873.

Sentencing Council for England and Wales (2011) Assault: Definitive guideline. Available at: https://www.sentencingcouncil.org.uk/publications/item/assault-definitive-guideline/ (accessed 15 December 2017).

Sentencing Council for England and Wales (2012) Crown Court Sentencing Survey: Annual publication, 2011. Available at: https://www.sentencingcouncil.org.uk/publications/item/crown-court-sentencing-survey-annual-publication-2011-full-report/ (accessed 13 September 2018).

Spohn CC, Kim B, Belenko S, et al. (2014) The direct and indirect effects of offender drug use on federal sentencing outcomes. *Journal of Quantitative Criminology* 30: 549–576.

Tata C and Hutton N (1998) What 'rules' in sentencing? Consistency and disparity in the absence of 'rules'. *International Journal of the Sociology of Law* 26(3): 339–364.

Ulmer JT and Johnson B (2004) Sentencing in context: A multilevel analysis. *Criminology* 42: 137–178.

Ulmer J, Light MT and Kramer J (2011) The 'liberation' of federal judges' discretion in the wake of the Booker/Fanfan decision: Is there increased disparity and divergence between courts? *Justice Quarterly* 28: 799–837.

United States Sentencing Commission (2004) *Fifteen years of guidelines sentencing: An assessment of how well the federal criminal justice system is achieving the goals of sentencing reform*. Washington, DC: United States Sentencing Commission.

Ward JT, Hartley RD and Tillyer R (2016) Unpacking gender and racial/ethnic biases in the federal sentencing of drug offenders: A causal mediation approach. *Journal of Criminal Justice* 46: 196–206.

Zatz MS (1987) The changing forms of racial/ethnic bias in sentencing. *Journal of Research in Crime and Delinquency* 24: 69–92.

Author biographies

Jose Pina-Sánchez is a lecturer in Quantitative Criminology at the University of Leeds. Previously he worked at the London School of Economics as a Fellow in Statistics. His research interests span substantive and methodological areas such as sentencing, compliance with the law, survey research and statistical modelling.

Ian Brunton-Smith is a professor of Criminology at the University of Surrey. His research interests cover criminal statistics, neighbourhood effects, survey methodology, and advanced quantitative methods.

Guangquan Li is a senior lecturer in Statistics at Northumbria University. His research interests are in the development and application of Bayesian methods to analyse data arising from health and social sciences

Appendix 1. Descriptive statistics of the variables used

	Mean	SD dv	Min	Max
Dependent variables				
Level of seriousness	2.01	.66	1	3
Guilty plea reduction	3.31	1.11	1	5
Final sentence outcome	3.33	.80	1	4
Specific type of offence				
Grievous bodily harm	.26	.44	0	1
Grievous bodily harm with intent	.10	.30	0	1
Common assault	.09	.29	0	1
Reference category: actual bodily harm	.54	.50	0	1
Step One: Harm and culpability factors				
Deliberate cause of harm	.05	.23	0	1
Intention to commit more harm	.04	.19	0	1
Leading role in gang	.05	.23	0	1
Motivated by victim's age/gender	.01	.08	0	1
Premeditation	.09	.29	0	1
Motivated by victim's race/religion	.01	.08	0	1
Motivated by victim's sexual orientation	<.01	.05	0	1
Deliberate targeting of vulnerable victim	.07	.26	0	1
Use of weapon	.38	.49	0	1
Serious injury	.29	.45	0	1
Sustained assault on the same victim	.25	.43	0	1
Victim is particularly vulnerable	.13	.33	0	1
Lack of premeditation	.28	.45	0	1
Offender's mental disorder	.03	.16	0	1
Great degree of provocation	.09	.29	0	1
Excessive self-defence	.07	.25	0	1
Subordinate role in gang	.03	.18	0	1
No serious injury	.31	.46	0	1
Step Two: Additional aggravating & mitigating factors				
Abuse of a position of trust or power	.02	.15	0	1
Offence against public sector worker	.04	.21	0	1
Offence committed on bail	.03	.17	0	1
Attempt to conceal or dispose of evidence	.01	.08	0	1
Victim forced to leave their home	.01	.12	0	1
Evidence of community impact	<.01	.04	0	1
Failure to respond to warnings expressed by others	.01	.12	0	1
Failure to comply with current court orders	.06	.24	0	1
Gratuitous degradation of victim	.02	.15	0	1
Location of the offence	.28	.45	0	1
Offence committed while on licence	.02	.15	0	1
Ongoing effect upon the victim	.17	.38	0	1
Presence of others including relatives	.16	.36	0	1

Appendix I. (Continued)

	Mean	SD dv	Min	Max
Previous violence or threats to the same victim	.07	.26	0	1
Timing of the offence	.10	.31	0	1
Commission of offence under the influence of alcohol or drugs	.33	.47	0	1
Steps taken to address addiction or offending behaviour	.08	.28	0	1
Offender's age or lack of maturity	.10	.29	0	1
Offender's exemplary conduct	.17	.38	0	1
Isolated incident	.17	.38	0	1
Lapse of time since the offence	.03	.16	0	1
Suffering from serious medical conditions requiring treatment	.03	.16	0	1
Suffering from a mental disorder or mental disability	.04	.19	0	1
No previous relevant or recent convictions	.28	.45	0	1
Sole or primary carer for dependant relatives	.04	.19	0	1
Genuine remorse	.40	.49	0	1
Single blow	.23	.42	0	1
Previous convictions: 1 to 3	.11	.31	0	1
Previous convictions: 4 to 9	.07	.24	0	1
Step Two: Guilty plea timing				
Guilty plea entered at first opportunity	.39	.49	0	1
At magistrates' courts	.13	.34	0	1
Prior to Plea and Case Management Hearing	.08	.27	0	1
At Plea and Case Management Hearing	.44	.50	0	1
After the Plea and Case Management Hearing but before trial	.10	.30	0	1
Reference category: On, or after, the day of trial	.25	.43	0	1

Appendix II. Technical explanation of the modelling strategy

To reflect the ordinal measurement of the three outcomes studied (decisions made following Step One, Step Four and the final sentence), an ordered logistic regression specification is used. This is based on the assumption that the three outcomes can be expressed as latent variables, Y_1^*, Y_2^* and Y_3^*, representing continuous but unobserved levels of offence seriousness, guilty plea reduction and final sentence severity respectively. The three outcomes are linked to their respective observed variables, Y_1, Y_2 and Y_3, through the following measurement models:

$$Y_1 = \begin{cases} 1 \ if & Y_1^* \le \theta_{1,1} \\ 2 \ if \ \theta_{1,1} < Y_1^* \le \theta_{1,2} \\ 3 \ if \ \theta_{1,2} < Y_1^* \end{cases} ; \ Y_2 = \begin{cases} 1 \ if & Y_2^* \le \theta_{2,1} \\ 2 \ if \ \theta_{2,1} < Y_2^* \le \theta_{2,2} \\ 3 \ if \ \theta_{2,2} < Y_2^* \le \theta_{2,3} \\ 4 \ if \ \theta_{2,3} < Y_2^* \le \theta_{2,4} \\ 5 \ if \ \theta_{2,4} < Y_2^* \end{cases} ; \ Y_3 = \begin{cases} 1 \ if & Y_3^* \le \theta_{3,1} \\ 2 \ if \ \theta_{3,1} < Y_3^* \le \theta_{3,2} \\ 3 \ if \ \theta_{3,2} < Y_3^* \le \theta_{3,3} \\ 4 \ if \ \theta_{3,3} < Y_3^* \end{cases} \quad (1)$$

The terms, $\theta_{1,j}$, $\theta_{2,k}$ and $\theta_{3,l}$, represent the thresholds to be estimated in each of the three outcome models, with subscripts j, k and l used to indicate the ordering: from 1 to 2, 1 to 4 and 1 to 3.

Having defined the response variables and their respective thresholds, the three outcome models can be formally expressed as follows:

$$Y_1^* = \beta_{1,m}X_{1,m} + \mu_{1,q}$$
$$Y_2^* = \beta_{2,n}X_{2,n} + \mu_{2,q} \tag{2}$$
$$Y_3^* = \beta_{3,p}X_{3,p} + \alpha_{3,1}Y_1 + \alpha_{3,2}Y_2 + \mu_{3,q}$$

Where $X_{1,m}$, $X_{2,n}$ and $X_{3,p}$ represent the set of guideline factors included in each outcome model with regression coefficients $\beta_{1,m}$, $\beta_{2,n}$ and $\beta_{3,p}$. We also include the terms $\alpha_{3,1}Y_1$ and $\alpha_{3,2}Y_2$ in the model for the final sentence, allowing us to estimate the effects of Step One and Step Four decisions on the final sentence outcome. Finally, the random part of the three models is represented by $u_{1,q}$, $u_{2,q}$ and $u_{3,q}$, where the subscript q is used as an index for each of the 74 courts included, with variance structure:

$$\begin{bmatrix} \mu_{1q} \\ \mu_{2q} \\ \mu_{3q} \end{bmatrix} \sim MVN(0,\Omega_u); \qquad where \quad \Omega_u = \begin{bmatrix} \sigma_{\mu 1}^2 & & \\ \sigma_{\mu 12} & \sigma_{\mu 2}^2 & \\ \sigma_{\mu 13} & \sigma_{\mu 23} & \sigma_{\mu 3}^2 \end{bmatrix} \tag{3}$$

The variances $\sigma_{\mu 1}^2$, $\sigma_{\mu 2}^2$ and $\sigma_{\mu 3}^2$ provide estimates of the unexplained between court disparities at each of the included steps, and can also be used to derive the intra-class correlation (ICC) coefficients, which can then be used to measure the level of between court inconsistency (Pina-Sánchez and Linacre, 2016); United States Sentencing Commission, 2004). ICCs were estimated using the following formula, ICC = $\sigma_\mu^2 / \left(\sigma_\mu^2 + \pi^2/3\right)$. The covariances $\sigma_{\mu 12}$, $\sigma_{\mu 13}$ and $\sigma_{\mu 23}$, indicate the extent that those disparities are interrelated. As such, they can be used to explore the patterning of sentencing inconsistencies across each of the guideline steps. In common with standard ordinal logit models, for each outcome, the case level errors are assumed to follow a logistic distribution.

Appendix III. WinBUGS Code for Model I

```
model{
  for(i in 1:2851) {
##################################################################################
####Modelling the level of seriousness as an ordinal outcome (Y1 as defined in Eq. 1)############
##################################################################################

  Level.of.seriousness[i] ~ dcat(pl[i,1:3])
  pl[i,1] <- cpl[i,1]              #### level 1 probability
  pl[i,2] <- cpl[i,2] - cpl[i,1] #### level 2 probability
  pl[i,3] <- cpl[i,3] - cpl[i,2] #### level 3 probability
```

```
#### define cpl, the cumulative probabilities on the logit scale (see the expression for Y1* in Eq. 2)
  for (j in 1:2) {
    logit(cpl[i,j]) <- gl[j] + mul[i] + l[1]*Step1HigherCulpA[i] + l[2]*Step1HigherCulpC[i] +
                                  l[3]*Step1HigherCulpD[i] + l[4]*Step1HigherCulpE[i] +
                                  l[5]*Step1HigherCulpF[i] + l[6]*Step1HigherCulpG[i] +
                                  l[7]*Step1HigherCulpH[i]
                                  + l[8]*Step1HigherCulpI[i] + l[9]*Step1HigherCulpJ[i] +
                                  l[10]*Step1HigherHarmA[i] +
                                  l[11]*Step1HigherHarmB[i] + l[12]*Step1HigherHarmC[i] +
                                  l[13]*Step1LowerCulpA[i] + l[14]*Step1LowerCulpB[i] +
                                  l[15]*Step1LowerCulpC[i] + l[16]*Step1LowerCulpD[i] +
                                  l[17]*Step1LowerCulpE[i] + l[18]*Step1LowerHarmA[i]
  }
  cpl[i,3] <- 1
  mul[i] <- ul[Sentencing.court[i]] # court-level random effects


###############################################################################
####Modelling the guilty plea reduction as an ordinal outcome (Y2 as defined in Eq. 1)############
###############################################################################

  Reduction.for.guilty.plea[i] ~ dcat(pr[i,])
  pr[i,1] <- cpr[i,1]           ####  level 1 probability
  pr[i,2] <- cpr[i,2] - cpr[i,1] #### level 2 probability
  pr[i,3] <- cpr[i,3] - cpr[i,2] #### level 3 probability
  pr[i,4] <- cpr[i,4] - cpr[i,3] #### level 4 probability
  pr[i,5] <- cpr[i,5] - cpr[i,4] #### level 5 probability
#### define cpr, the cumulative probabilities on the logit scale (see the expression for Y2* in Eq. 2)
  for (j in 1:4) {
    logit(cpr[i,j]) <- gr[j] + mur[i] + r[1]*first[i] + r[2]*stage_magist[i] + r[3]*stage_priorPCMH[i] +
                               r[4]*stage_PCMH[i] + r[5]*stage_afterPCMH[i]
  }
  cpr[i,5] <- 1
  mur[i] <- ur[Sentencing.court[i]] # court-level random effects


###############################################################################
####Modelling the final sentence outcome as an ordinal outcome (Y3 as defined in Eq. 1)##########
###############################################################################

  Sentence.outcome[i] ~ dcat(ps[i,])
  ps[i,1] <- cps[i,1]           #### level 1 probability
  ps[i,2] <- cps[i,2] - cps[i,1] #### level 2 probability
  ps[i,3] <- cps[i,3] - cps[i,2] #### level 3 probability
  ps[i,4] <- cps[i,4] - cps[i,3] #### level 4 probability
#### define cps, the cumulative probabilities on the logit scale (see the expression for Y3* in Eq. 2)

  for (j in 1:3) {
    logit(cps[i,j]) <- gs[j] + mus[i] + s[1]*Reduction.for.guilty.plea[i] + s[2]*Level.of.seriousness[i] +
                               s[3]*GBH[i] + s[4]*intent[i] + s[5]*common[i] +
                               s[6]*PC_1to3[i] + s[7]*PC_4to9[i] + Aggr[i] + Miti[i]
  }
```

```
 Aggr[i] <- s[8]*Aggravating.3[i] + s[9]*Aggravating.5[i] + s[10]*Aggravating.9[i] +
            s[11]*Aggravating.11[i] + s[12]*Aggravating.19[i] + s[13]*Aggravating.22[i] +
            s[14]*Aggravating.25[i] + s[15]*Aggravating.26[i] + s[16]*Aggravating.28[i] +
            s[17]*Aggravating.37[i] + s[18]*Aggravating.44[i] + s[19]*Aggravating.45[i] +
            s[20]*Aggravating.54[i] + s[21]*Aggravating.55[i] + s[22]*Aggravating.70[i] +
            s[23]*Aggravating.72[i]
 Miti[i] <- s[24]*Mitigating.3[i] + s[25]*Mitigating.6[i] + s[26]*Mitigating.16[i] +
            s[27]*Mitigating.20[i] + s[28]*Mitigating.22[i] + s[29]*Mitigating.24[i] +
            s[30]*Mitigating.25[i] + s[31]*Mitigating.29[i] + s[32]*Mitigating.33[i] +
            s[33]*Mitigating.36[i] + s[34]*Mitigating.39[i]
 cps[i,4] <- 1
 mus[i] <- us[Sentencing.court[i]] # court-level random effects
 }


###############################################################################
####Prior specifications#######################################################
###############################################################################


#### diffuse (vague) priors for the overall thresholds for level of seriousness
 gl[1] ~ dnorm(0,0.0001)
 gl[2] <- gl[1] + deltal[2]
 for (i in 2:2) {deltal[i] ~ dgamma(0.01,0.01)}


#### diffuse (vague) priors for the overall thresholds for guilty plea reduction
 gr[1] ~ dnorm(0,0.0001)
 gr[2] <- gr[1] + deltar[2]
 gr[3] <- gr[2] + deltar[3]
 gr[4] <- gr[3] + deltar[4]
 for (i in 2:4) {deltar[i] ~ dgamma(0.01,0.01)}


#### diffuse (vague) priors for the overall thresholds for sentence outcome
 gs[1] ~ dnorm(0,0.0001)
 gs[2] <- gs[1] + deltas[2]
 gs[3] <- gs[2] + deltas[3]
 for (i in 2:3) {deltas[i] ~ dgamma(0.01,0.01)}


#### modelling the three sets of court-level random effects jointly through a multivariate normal (see Eq. 3)
 for (j in 1:74) {     #### 74 courts in total
  ul[j] <- U[j,1]     #### random effects on level of seriousness
  ur[j] <- U[j,2]    #### random effects on guilty plea reduction
  us[j] <- U[j,3] #### random effects on sentence outcome
#### jointly modelling the three sets of random effects
  U[j,1:3] ~ dmnorm(U.hat[1:3], Tau.U[1:3,1:3])
  }
#### means of the random effects are set to 0 so that the random effects are centred at the overall thresholds
 for (i in 1:3) {U.hat[i] <- 0}


#### a Wishart prior on the inverse of the random effect covariance matrix Tau.U
#### a Wishart distribution is a probability distribution with k, a scaler parameter, and R, a symmetric positive definite p-by-p matrix
```

```
#### where p=dimension of the MV normal
#### Here, we assign a vague Wishart prior with the following setting for k and R:
#### the scaler parameter k needs to be greater than p-1 and a smaller k value shows higher uncer-
tainty (hence less prior information) for Tau.U
#### hence we set k=3
#### For R, we set the off-diagonal terms to be 0, corresponding to a prior guess of no pairwise
correlations amongst the three sets of random effects;
#### the diagonal terms in R are set to be 1, corresponding to a prior guess at the variance for each
set of random effects to be 1/3 (i.e. 1/k from the
#### properties of a Wishart distribution)
    Tau.U[1:3,1:3] ~ dwish(R[1:3,1:3],k)
    k <- 3 # degrees of freedom in the Wishart
    for (i in 1:3) {R[i,i] <- 1} # diagonal entries
    #### off-diagonal entries
    R[1,2] <- 0
    R[2,1] <- 0
    R[1,3] <- 0
    R[2,3] <- 0
    R[3,1] <- 0
    R[3,2] <- 0

####       vague priors for the regression coefficients for level of seriousness
    for (k in 1:18) {
   l[k] ~ dnorm(0,0.001)
    }
####       vague priors for the regression coefficients for guilty plea reduction
    for (k in 1:5) {
   r[k] ~ dnorm(0,0.001)
    }
####       vague priors for the regression coefficients for sentence outcome
    for (k in 1:34) {
   s[k] ~ dnorm(0,0.001)
    }


##############################################################################
####calculating other quantities of interest#################################
##############################################################################


#### compute the covariance matrix of the random effects
    Sigma.U[1:3,1:3] <- inverse(Tau.U[1:3,1:3])

# random effect standard deviations (see end of Table 2 in the main text)
    sigma.L <- pow(Sigma.U[1,1],0.5) ### random effects SD on Level
    sigma.R <- pow(Sigma.U[2,2],0.5) ### random effects SD on Reduction
    sigma.S <- pow(Sigma.U[3,3],0.5) ### random effects SD on Sentence outcome

# random effect covariances (see end of Table 2 in the main text)
    sigma.LR <- Sigma.U[1,2] ### covariance between Level and Reduction
    sigma.LS <- Sigma.U[1,3] ### covariance between Level and Sentence outcome
    sigma.RS <- Sigma.U[2,3] ### covariance between Reduction and Sentence outcome
```

Appendix IV. Results from Model 1 (multilevel multivariate without mediating effects)

Step One: Assessed seriousness		Step Four: Guilty plea reduction		Final sentence	
Variables	Mean (95% CI)	Variables	Mean (95% CI)	Variables	Mean (95% CI)
Threshold 1/2	-.81 (-1.04, -.58)	Threshold 1/2	-.17 (-.38, .03)	Threshold 1/2	**-1.74** (-2.23, -1.27)
Threshold 2/3	**3.22** (2.94, 3.50)	Threshold 2/3	**1.41** (1.18, 1.64)	Threshold 2/3	**1.06** (.63, 1.49)
		Threshold 3/4	**3.09** (2.81, 3.37)	Threshold 3/4	**3.75** (3.28, 4.23)
		Threshold 4/5	**8.84** (8.40, 9.30)		
				Seriousness (step one)	**1.50** (1.33, 1.67)
				Reduction GP (step four)	.06 (-.02, .14)
				Offence (ref: ABH)	
				GBH	**1.45** (1.23, 1.66)
				GBH with intent	**5.15** (4.36, 6.02)
				Common assault	**-1.17** (-1.48, -.87)
Step One factors		*Step Four factors*		*Step Two factors*	
Deliberate harm	.68 (.24, 1.12)	First opportunity	**1.92** (1.65, 2.19)	Previous convictions (1–3)	**1.09** (.73, 1.45)
Intent. serious harm	.83 (.34, 1.31)	Plea (ref: on/ after trial)		Previous convictions (4–9)	**1.61** (.93, 2.43)
Leading role gang	.57 (.17, .98)	At magistrates	**5.17** (4.73, 5.62)	Abuse of trust	.13 (-.50, .77)
Hostility age/gender	.62 (-.71, 1.98)	Prior to PCMH	**5.09** (4.60, 5.59)	Against public	.35 (-.07, .79)
Premeditation	.64 (.30, .98)	At PCMH	**4.11** (3.82, 4.40)	On bail	**1.65** (.93, 2.43)
Racially motivated	.95 (-.15, 2.09)	After PCMH	**1.99** (1.70, 2.29)	Dispose of evidence	1.61 (-.71, 4.88)
Hostility orientation	1.43 (-.64, 3.99)			Victim forced leave	**.91** (.11, 1.76)
Targeting vulnerable	**1.12** (.74, 1.49)			Community impact	.38 (-1.80, 2.82)
Use of weapon	**1.46** (1.27, 1.66)			Failure warnings	.52 (-.44, 1.59)
Serious injury	**1.83** (1.59, 2.07)			Failure court orders	**1.62** (1.13, 2.14)
Sustained assault	**1.50** (1.27, 1.74)			Gratuitous degradation	**.94** (.11, 1.82)

(Continued)

Appendix IV. (Continued)

Step One: Assessed seriousness		Step Four: Guilty plea reduction		Final sentence	
Vulnerable victim	**.85** (.54, 1.15)			Location	.23 (−.001, .46)
Lack premeditation	**−.73** (−.94, −.51)			While on licence	**1.85** (.88, 2.92)
Mental disorder	**−1.12** (−1.67, −.57)			Ongoing effect	**.65** (.38, .92)
Provocation	**−.61** (−.93, −.30)			Presence of others	**.49** (.23, .75)
Self-defence	**−.43** (−.80, −.06)			Previous violence	**.86** (.43, 1.29)
Subordinate role	−.20 (−.69, .29)			Timing of offence	**.38** (.03, .73)
Injury less serious	**−.87** (−1.09, −.65)			Under drugs/alcohol	**.22** (.02, .42)
				Address addiction	**−.97** (−1.27, −.68)
				Lack of maturity	−.22 (−.51, .06)
				Good character	**−.77** (−1.01, −.52)
				Isolated incident	**−.51** (−.74, −.28)
				Lapse of time	−.34 (−.83, .14)
				Medical condition	−.32 (−.81, .17)
				Mental disability	**−.62** (−1.08, −.17)
				No relevant convictions	**−.56** (−.77, −.35)
				Primary carer	**−.87** (−1.27, −.48)
				Genuine remorse	**−.29** (−.49, −.10)
				Single blow	**−.36** (−.56, −.15)

Random effects

	Step One		Step Four		Final sentence
$\sigma_{\mu 1}$	**.36** (.26, .49)	$\sigma_{\mu 2}$	**.43** (.31, .57)	$\sigma_{\mu 3}$	**.40** (.29, .54)
		$\sigma_{\mu 12}$	−.02 (−.10, .05)	$\sigma_{\mu 13}$	−.01 (−.09, .05)
				$\sigma_{\mu 23}$	.04 (−.03, .12)

Note: Estimates in bold have a 95% credible interval that do not include 0. ABH = actual bodily harm; GBH = grievous bodily harm.

Appendix V. Results from Model 2 (multilevel multivariate with mediating effects)

	Step One: Assessed seriousness	Step Four: Guilty plea reduction	Final sentence
Variables	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)
Threshold 1/2	−.76 (−1.08,−0.46)	−.32 (−.62,−.03)	−2.81 (−3.39,−2.25)
Threshold 2/3	3.42 (3.06,3.79)	1.33 (1.01,1.64)	.12 (−.43,.63)
Threshold 3/4		3.07 (2.73,3.42)	2.96 (2.40,3.50)
Threshold 4/5		9.08 (8.58,9.60)	
Seriousness (step one)			.90 (.69,1.11)
Reduction GP (step four)			−.15 (−.30,−.01)
Offence (ref: ABH)			
GBH			1.65 (1.43,1.88)
GBH with intent			5.09 (4.29,6.00)
Common assault			−1.03 (−1.35,−.72)
Step One factors			
Deliberate harm	.56 (.11,1.02)	−.11 (−.52,.30)	.43 (−.06,.93)
Intent. serious harm	.71 (.21,1.21)	−.49 (−.99,.01)	.99 (.27,1.76)
Leading role gang	.46 (.04,.88)	.26 (−.15,.67)	.52 (.03,1.01)
Hostility age/gender	.73 (−.69,2.18)	−.17 (−1.27,.95)	.11 (−1.11,1.39)
Premeditation	.54 (.19,.89)	.02 (−.30,.34)	.53 (.14,.93)
Racially motivated	.69 (−.47,1.86)	−.39 (−1.46,.73)	.20 (−.82,1.24)
Hostility orientation	1.44 (−.72,3.91)	−1.61 (−3.23,.04)	−1.54 (−3.27,.32)
Targeting vulnerable	1.00 (.61,1.40)	−.06 (−.43,.31)	1.06 (.61,1.51)
Use of weapon	1.38 (1.17,1.59)	−.07 (−.27,.13)	.70 (.49,.92)
Serious injury	1.74 (1.49,2.00)	.07 (−.16,.30)	.75 (.50,1.00)
Sustained assault	1.35 (1.10,1.61)	.13 (−.11,.38)	.69 (.43,.94)
Vulnerable victim	.71 (.39,1.04)	.01 (−.30,.31)	.32 (−.01,.65)
Lack premeditation	−.67 (−.89,−.45)	−.05 (−.28,.17)	−.57 (−.78,−.37)

(Continued)

Appendix V. (Continued)

	Step One: Assessed seriousness	Step Four: Guilty plea reduction	Final sentence
Mental disorder	**-.92** (-1.59,-.25)	.53 (-.19,1.26)	.49 (-.13,1.10)
Provocation	**-.57** (-.91,-.25)	.07 (-.27,.41)	**-.53** (-.81,-.24)
Self-defence	-.30 (-.68,.07)	.13 (-.24,.50)	-.26 (-.59,.07)
Subordinate role	-.20 (-.71,.31)	.14 (-.37,.64)	**-.49** (-.95,-.03)
Injury less serious	**-.91** (-1.14,-.68)	**-.29** (-.53,-.05)	.11 (-.10,.32)
Step Two factors			
Previous convictions (1–3)	**.34** (.03,.66)	**-.37** (-.68,-.08)	**1.21** (.83,1.59)
Previous convictions (4–9)	.23 (-.19,.65)	.08 (-.32,.49)	**1.62** (1.11,2.16)
Abuse of trust	.37 (-.32,1.06)	-.27 (-.89,.34)	-.09 (-.75,.57)
Against public	-.43 (-.88,.01)	-.21 (-.68,.26)	.18 (-.25,.62)
On bail	.22 (-.33,.77)	-.10 (-.62,.42)	**1.56** (.82,2.37)
Dispose of evidence	.03 (-1.14,1.27)	-.45 (-1.52,.61)	.91 (-1.40,4.22)
Victim forced leave	-.52 (-1.30,.25)	**-.85** (-1.66,-.03)	.47 (-.35,1.33)
Community impact	-.24 (-2.22,1.89)	.19 (-2.06,2.5)	-.84 (-2.94,1.47)
Failure warnings	-.03 (-.84,.79)	.28 (-.54,1.09)	.40 (-.56, 1.42)
Failure court orders	-.25 (-.65,.15)	-.22 (-.63,.19)	**1.60** (1.09,2.14)
Gratuitous degradation	.62 (-.08,1.34)	.06 (-.54,.68)	.60 (-.23,1.50)
Location	**.38** (.14,.62)	.01 (-.24,.24)	.21 (-.03,.45)
While on licence	-.32 (-.92,.28)	.02 (-.62,.66)	**1.76** (.80,2.82)
Ongoing effect	.24 (-.02,.50)	-.21 (-.47,.05)	**.49** (.21,.78)
Presence of others	**.28** (.01,.54)	**-.28** (-.55,-.02)	**.42** (.15,.69)
Previous violence	.26 (-.12,.65)	-.31 (-.67,.05)	**.79** (.35,1.23)
Timing of offence	-.01 (-.35,.34)	.02 (-.32,.37)	.20 (-.16,.56)
Under drugs/alcohol	**.26** (.05,.46)	**-.21** (-.42,-.01)	.18 (-.02,.38)
Address addiction	-.23 (-.57,.10)	.24 (-.11,.59)	**-.97** (-1.28,-.67)

Appendix V. (Continued)

	Step One: Assessed seriousness	Step Four: Guilty plea reduction	Final sentence
Lack of maturity	.02 (−.30,.33)	.09 (−.24,.42)	−.24 (−.54,.05)
Good character	−.02 (−.30,.26)	**.49** (.19,.79)	**−.71** (−.95,−.46)
Isolated incident	**−.37** (−.64,−.11)	.06 (−.22,.34)	**−.40** (−.64,−.17)
Lapse of time	−.38 (−.98,.21)	.36 (−.19,.91)	−.05 (−.55,.44)
Medical condition	−.04 (−.62,.54)	.03 (−.60,.67)	−.47 (−.97,.03)
Mental disability	−.21 (−.78,.37)	−.04 (−.64,.55)	**−.89** (−1.42,−.35)
No relevant convictions	−.13 (−.37,.11)	−.22 (−.47,.03)	**−.65** (−.87,−.43)
Primary carer	.42 (−.04,.88)	.05 (−.41,.52)	**−.93** (−1.33,−.52)
Genuine remorse	.14 (−.07,.36)	**.27** (.05,.48)	**−.20** (−.40,−.01)
Single blow	**−.22** (−.44,−.01)	.20 (−.04,.44)	−.15 (−.36,.07)
Step Four factors			
First opportunity	−.01 (−.22,.21)	**1.89** (1.61,2.17)	−.08 (−.29,.13)
Plea (ref: on/after trial)			
At magistrates'	.05 (−.31,.40)	**5.41** (4.95,5.87)	**.71** (.29,1.14)
Prior to PCMH	.33 (−.07,.73)	**5.28** (4.78,5.79)	**.88** (.41,1.37)
At PCMH	.06 (−.19,.30)	**4.27** (3.97,4.57)	**.56** (.21,.91)
After PCMH	−.11 (−.45,.23)	**2.06** (1.75,2.36)	.36 (−.01,.73)
Random effects			
$\sigma_{\mu1}$	**.32** (.24,.46)		
$\sigma_{\mu2}$		**.38** (.28,.54)	
$\sigma_{\mu3}$			**.36** (.27,.51)
$\sigma_{\mu12}$		.93 (−2.58,4.90)	
$\sigma_{\mu13}$			−.96 (−5.22,2.78)
$\sigma_{\mu23}$			−1.37 (−4.90,1.72)

Note: Estimates in bold have a 95% credible interval that do not include 0. GP = guilty plea; ABH = actual bodily harm; GBH = grievous bodily harm.

Appendix VI. Direct and indirect effects of Step One and Step Four factors.

Coefficient	Total effect	Indirect effect through Step One	Indirect effect through Step Four	Direct effect
Step One factors				
Deliberate harm	.96 (.32,1.61)	.50 (.10,.95)	.02 (−.05,.10)	.43 (−.06,.93)
Intent. serious harm	1.71 (.86,2.61)	.64 (.18,1.14)	.08 (−.01,.21)	.99 (.27,1.76)
Leading role gang	.89 (.28,1.52)	.41 (.04,.82)	−.04 (−.14,.02)	.52 (.03,1.01)
Hostility age/gender	.80 (−.92,2.59)	.66 (−.62,1.99)	.03 (−.17,.24)	.11 (−1.11,1.39)
Premeditation	1.01 (.52,1.53)	.48 (.16,.84)	−.01 (−.06,.05)	.53 (.14,.93)
Racially motivated	.88 (−.54,2.32)	.62 (−.41,1.72)	.06 (−.12,.29)	.20 (−.82,1.24)
Hostility orientation	−.01 (−2.5,2.75)	1.29 (−.62,3.61)	.25 (−.03,.69)	−1.54 (−3.27,.32)
Targeting vulnerable	1.97 (1.39,2.58)	.90 (.52,1.34)	.01 (−.06,.08)	1.06 (.61,1.51)
Use of weapon	1.96 (1.60,2.33)	1.24 (.91,1.60)	.01 (−.02,.05)	.70 (.49,.92)
Serious injury	2.31 (1.88,2.78)	1.57 (1.15,2.01)	−.01 (−.06,.03)	.75 (.50,1.00)
Sustained assault	1.88 (1.49,2.30)	1.21 (.86,1.60)	−.02 (−.08,.02)	.69 (.43,.94)
Vulnerable victim	.96 (.53,1.42)	.64 (.33,.99)	.01 (−.06,.06)	.32 (−.01,.65)
Lack premeditation	−1.17 (−1.49,−.87)	−.60 (−.86,−.37)	.01 (−.03,.05)	−.57 (−.78,−.37)
Mental disorder	−.43 (−1.31,.43)	−.83 (−1.51,−.21)	−.08 (−.26,.03)	.49 (−.13,1.10)
Provocation	−1.06 (−1.49,−.64)	−.52 (−.85,−.21)	−.01 (−.08,.05)	−.53 (−.81,−.24)
Self-defence	−.56 (−1.04,−.08)	−.27 (−.63,.07)	−.02 (−.13,.06)	−.26 (−.59,.07)
Subordinate role	−.69 (−1.34,−.05)	−.18 (−.65,.28)	.04 (−.01,.12)	−.49 (−.95,−.03)
Injury less serious	−.67 (−1.01,−.35)	−.82 (−1.12,−.56)	.06 (−.01,.15)	.11 (−.10,.32)
Step Two factors				
Previous convictions (1–3)	1.57 (1.11,2.06)	.31 (.02,.61)	.06 (−.01,.15)	1.21 (.83,1.59)
Previous convictions (4–9)	1.81 (1.18,2.46)	.20 (−.18,.59)	−.01 (−.09,.06)	1.62 (1.11,2.16)
Abuse of trust	.28 (−.62,1.21)	.33 (−.28,.98)	.04 (−.06,.18)	−.09 (−.75,.57)
Against public	−.17 (−.77,.42)	−.39 (−.82,.01)	.03 (−.04,.14)	.18 (−.25,.62)
On bail	1.77 (.88,2.72)	.20 (−.30,.72)	.02 (−.07,.12)	1.56 (.82,2.37)
Dispose of evidence	1.01 (−1.56,4.47)	.03 (−1.03,1.16)	.07 (−.10,.30)	.91 (−1.40,4.22)
Victim forced leave	−.13 (−.96,1.22)	−.47 (−1.20,.23)	.13 (−.01,.36)	.47 (−.35,1.33)
Community impact	−1.09 (−3.94,1.90)	−.22 (−2.04,1.73)	−.03 (−.48,.37)	−.84 (−2.94,1.47)

Appendix VI. (Continued)

Coefficient	Total effect	Indirect effect through Step One	Indirect effect through Step Four	Direct effect
Failure warnings	.32 (−.87,1.58)	−.03 (−.77,.72)	−.04 (−.02,.09)	.40 (−.56, 1.42)
Failure court orders	.21 (−.83,1.32)	−.22 (−.60,.14)	.03 (−.03,.13)	**1.60** (1.09,2.14)
Gratuitous degradation	**1.15** (.10,2.25)	.56 (−.07,1.24)	−.01 (−.13,.10)	.60 (−.23,1.50)
Location	**.55** (.23,.88)	**.34** (.12,.58)	.01 (−.07,.06)	.21 (−.03,.45)
While on licence	**1.47** (.37,2.64)	−.29 (−.85,.25)	−.01 (−.12,.11)	**1.76** (.80,2.82)
Ongoing effect	**.74** (.38,1.12)	.21 (−.02,.46)	.03 (−.01,.09)	**.49** (.21,.78)
Presence of others	**.71** (.35,1.08)	**.25** (.01,.51)	.04 (−.01,.12)	**.42** (.15,.69)
Previous violence	**1.06** (.51,1.63)	.23 (−.12,.60)	.05 (−.01,.14)	**.79** (.35,1.23)
Timing of offence	.19 (−.28,.67)	−.01 (−.32,.31)	−.01 (−.07,.06)	.20 (−.16,.56)
Under drugs/alcohol	**.44** (.17,.73)	**.23** (.04,.43)	.03 (−.01,.90)	.18 (−.02,.38)
Address addiction	**−1.22** (−1.65,−.80)	**−.21** (−.53,−.10)	−.04 (−.12,.02)	**−.97** (−1.28,−.67)
Lack of maturity	−.24 (−.65,.16)	.01 (−.27,.30)	−.01 (−.08,.04)	−.24 (−.54,.05)
Good character	**−.80** (−1.16,−.45)	−.02 (−.28,.24)	**−.08** (−.18,−.01)	**−.71** (−.95,−.46)
Isolated incident	**−.75** (−1.09,−.41)	**−.33** (−.60,−.09)	−.01 (−.06,.04)	**−.40** (−.64,−.17)
Lapse of time	−.45 (−1.18,.27)	−.33 (−.89,.19)	−.06 (−.19,.03)	−.05 (−.55,.44)
Medical condition	−.50 (−1.22,.22)	−.03 (−.56,.50)	−.01 (−.12,.11)	−.47 (−.97,.03)
Mental disability	**−1.07** (−1.81,−.33)	−.19 (−.72,.34)	.01 (−.10,.12)	**−.89** (−1.42,−.35)
No relevant conviction	**−.73** (−1.04,−.42)	−.12 (−.34,.10)	.03 (−.01,.10)	**−.65** (−.87,−.43)
Primary carer	−.56 (−1.14,.02)	.37 (−.04,.82)	−.01 (−.10,.08)	**−.93** (−1.33,−.52)
Genuine remorse	−.11 (−.39,.17)	.13 (−.06,.33)	−.04 (−.10,.01)	**−.20** (−.40,−.01)
Single blow	**−.37** (−.67,−.08)	−.19 (−.41,.21)	−.03 (−.09,.01)	−.15 (−.36,.07)
Step Four factors				
First opportunity	**−.37** (−.73,−.02)	−.01 (−.02,.02)	**−.29** (−.59,−.01)	−.08 (−.29,.13)
At magistrates'	−.07 (−.76,.61)	.04 (−.27,.37)	**−.84** (−1.65,−.05)	**.71** (.29,1.14)
Prior to PCMH	.37 (−.37,1.10)	.30 (−.07,.68)	**−.82** (−1.62,−.05)	**.88** (.41,1.37)
At PCMH	−.04 (−.54,.44)	.05 (−.17,.28)	**−.66** (−1.30,−.04)	**.56** (.21,.91)
After PCMH	−.05 (−.54,.42)	−.10 (−.41,.21)	**−.32** (−.63,−.02)	.36 (−.01,.73)

Note: Estimates in bold have a 95% credible interval that do not include 0.