# 5
# Defining and Measuring Consistency in Sentencing

*Jose Pina-Sánchez*

## Introduction

Sentencing consistency, or the extent to which like cases are treated alike, is a fundamental principle of justice. It generates transparency and predictability in sentencing practices, enhances public confidence in sentencing (Council of Europe, 1993) and helps promote the legitimacy of the criminal justice system (Roberts and Plesnicar, 2015). However, the means by which this consistency in sentencing may be achieved is far from clear-cut and highly controversial.

Over the years, critics of the sentencing process have contended that unstructured discretion results in sentencing disparity. In the influential publication 'Lawlessness in Sentencing' (Frankel, 1972), Judge Marvin Frankel claimed that unstructured discretion had led to discrimination and disparity in the United States. In response, many states introduced greater structure in the sentencing process through the implementation of sentencing guidelines. These guidelines vary in structure but all share the goal of establishing objective standards so that the degree of punishment an offender receives does not depend on the values, attitudes, and beliefs of the judge (Walker, Spohn and Delone, 2007).

This move towards structuring discretion at sentencing has been adopted by other common law jurisdictions outside the United States, including England and Wales. Sentencers in England and Wales used to enjoy a substantial degree of discretion with the only restrictions on sentencers arising from statutory limits and by the guidance offered by appellate review (Ashworth and Roberts, 2013, p. 1). The Criminal Justice Act 2003 tightened this flexible and highly discretionary framework, created the Sentencing Guidelines Council – an independent body responsible for generating definitive guidelines, and established

a duty on courts to have regard to any guidelines produced by the council.

The Sentencing Guidelines Council produced guidelines for specific offence categories (such as assault) as well as generic guidelines applicable across all offences.[1] This development was further reinforced with the Coroners and Justice Act 2009, which created the Sentencing Council for England and Wales, endowed it with new analytical responsibilities, and made the guidelines more binding on courts.[2] The guidelines process reached a state of maturity when in June 2011, a new guideline for the assault offences came into force.[3] This new format guideline replaced the assault offences guideline issued by the former Sentencing Guidelines Council in 2007.[4] Specifically, the new guideline sought to correct some problems of applicability and to introduce a clearer and more structured stepwise process to the sentencing exercise. In the words of the first chair of the sentencing council, this new structure aimed to 'increase the consistency of approach to sentencing so that offenders receive the same approach whether they're being sentenced in Bristol, Birmingham, Bolton or Basildon' (Lord Justice Leveson, 2011).[5]

In spite of the interest in the new sentencing guidelines, little is known about the degree of variability in sentencing, or the effectiveness of the sentencing guidelines in promoting greater consistency. This may be explained by the elusiveness of the concept, which makes it hard to be operationalised under a quantitative approach, and also by an absence (to date at least) of adequate data. As a result of the publication of the Crown Court sentencing survey (CCSS), a new wave of studies on the topic has emerged.[6] These studies have employed different empirical methods and have examined different aspects of sentencing. The Sentencing Council (2012) analysed departures from the guideline ranges for some offences of assault; Roberts (2013) and Roberts and Bradford (2014) explored guilty plea reductions, while Pina-Sánchez and Linacre (2013; 2014) focused on analyses of custodial sentence length.

## Overview

This chapter provides more generalizable results on the topic of consistency in sentencing by replicating most of the methods that have been implemented in earlier studies and by considering the study of custodial types. In particular, the analysis considers the use of custody versus other disposals for cases of assault. Compared to the study of consistency in sentence length, the study of disposal types is more comprehensive since it makes use of all cases, not just those that were sentenced to custody. In addition, where possible, analyses aiming to assess

changes in consistency after the new assault guideline came into force have also been carried out. The next section, 'Measuring consistency in sentencing', reviews the scope and limitations of the methods used for the measurement of consistency. 'Analysis using the Crown Court's sentencing survey (CCSS)' presents results from the analysis. Finally, I conclude with a summary of the main findings.

## Measuring consistency in sentencing

The simplest way to document consistency in sentencing is through the study of guidelines compliance statistics. This method involves calculating the percentage of offences that receive sentences within the ranges prescribed by the sentencing guidelines. In its most basic form, the study of compliance can be implemented in any jurisdiction structured by sentencing guidelines where individual-level data on the type of offence and the sentence outcome are available. Accordingly, it has been widely used to assess the effectiveness of the sentencing guidelines in the United States (Frase, 2005; Kramer and Ulmer, 2002; Minnesota Sentencing Guidelines Commission, 2010; Oregon Criminal Sentencing Commission, 2003; Scott, 2010; Tonry, 1987; Ulmer et al., 2011). The outcome of the CCSS has also permitted a comparable approach in England and Wales (Sentencing Council, 2012; Roberts, 2013).

The Sentencing Council (2012) reported the percentage of sentences imposed for actual bodily harm (ABH), grievous bodily harm (GBH), grievous bodily harm with intent (GBH with intent) and common assault that fell within the guideline ranges after the new guideline came into force. The percentages for those four offences were 97%, 97%, 92% and 99%, respectively. These results demonstrate high levels of compliance. However, since the Sentencing Council did not include analyses of trends prior to the new guideline coming into force, we cannot attribute the observed findings to the change in guidelines. These high rates of compliance may be due to the more binding nature of guidelines after the Coroners and Justice Act (2009) or simply to the very wide range of disposals available to courts within the assault guidelines. For example, for ABH (the most prevalent assault offence), the total offence guideline range runs from a fine up to three years of custody. It may be argued that almost all sentences would have fallen within this range, even in the absence of a guideline.

In addition, although guidelines-based compliance statistics can be used to shed light on consistency, they can only be approximations of the actual measures of consistency that we are seeking. By revealing

levels of departure from the guidelines, compliance statistics reflect a normative view of consistency. This is the view used in some US jurisdictions such as Minnesota or Oregon, where the guidelines aim to achieve uniformity through the implementation of a grid-based system that places offences into a limited number of categories linked to specific ranges of sentence outcomes.

Grid-based sentencing guidelines were rejected by the Sentencing Commission Working Group (2008) as being too restrictive and contrary to the traditions of English sentencing. In contrast, the system of guidelines developed in England and Wales seeks to promote consistency of approach. This is achieved by assisting sentencers in determining the most appropriate sentence outcome using a structured decision-making process that incorporates the legal factors relevant in each case. The logic of the guidelines is that if all sentencers follow the same methodical approach, sentencing outcomes will become more consistent. In this way, consistency can be promoted without undermining the judicial discretion necessary to differentiate between cases.

To obtain estimates that better differentiate consistency from uniformity, we need to use methods that overcome the simple measures of compliance rates (i.e. whether sentences fall within or without the guidelines range). One such method is 'exact matching' (Hofer et al., 1999, and Pina-Sánchez and Linacre, 2014). This method involves combining offences into groups that are as homogenous as possible and exploring the variability of sentence outcomes observed within each of these groups.[7] This procedure reflects a direct operationalization of the concept of consistency as it is commonly understood: the extent to which like cases are treated alike. By the use of matched groups, we can ensure that offences are as 'like' as possible, and by the study of the within group variability we assess how 'alike' they are being treated. Pina-Sánchez and Linacre (2014) used this approach to assess changes in the variability of sentence lengths for assault after the new guideline came into force. They found that more than half (57%) of the groups of offences showed less variability after the new guideline came into force, while the overall variability decreased by 7%.

Other methods used in the literature have complemented the study of the variability within types of offences by examining whether that variability is a result of differences in sentencing between courts. This way, different measures can be obtained that respond to a broader definition of consistency: the extent to which similar offences attract similar sentences regardless of where they are sentenced. Tarling and Mason et al. used descriptive statistics to explore the degree of variability

between courts in England and Wales. Tarling compared custody rates for burglary in 30 magistrates' courts in 1974 and 2000, while Mason et al. explored variability in custody rates, sentence lengths, and other disposal types across the 42 criminal justice areas from 2003 to 2006. Both studies found substantial differences among courts or criminal justice areas, although such variability remained relatively constant during the periods studied.

Tarling (2006) found that differences among magistrates' courts with the highest and lowest uses of custodial sentences and fines in 1975 ranged by 17.7% and 30.1%, respectively, and those differences remained relatively unchanged after replicating the analysis using data from the year 2000. Interestingly, Tarling found even greater variation in the use of various disposal types between courts after replicating the analysis for offences of burglary only. For example, the use of fines ranged by 22 percentage points, while that of custodial sentences reached 32.

Mason et al. (2007) explored disparities across magistrates' and Crown Courts in the use of custodial rates, custodial sentence length, and indeterminate sentences for public protection. Similar results demonstrating high variability were found for the three sentence outcomes and the two sentencing courts. For example, from 2003 to 2006, the range in the use of custodial rates between Crown Courts remained stable at an average of 24 percentage points. Specifically, in 2006, the use of custody rate ranged from 45% to 68%. The authors also pointed out that the majority of courts were close to the average of 56% and that just a few courts showing more extreme results created the wider range. Furthermore, disparities appeared to be systematic over time for certain courts. Of those with custody rates in the top five for 2006, two (Northamptonshire and Bedfordshire) were consistently in the top five for 2003, 2004 and 2005, whereas from those in the bottom five for 2006, one (Northumbria) was consistently in the bottom five for 2003, 2004 and 2005.

### Interpreting variability

As noted earlier, the problem with using these descriptive measures of between court variability is that one cannot tell whether the observed differences are due to a genuine lack of consistency or to differences in the composition of offences sentenced in each court. For example, in the study of variability of custodial rates using offences of assault, we would expect that courts where serious cases (e.g., GBH with intent) are more prevalent will have higher custody rates than those where common assaults are the most frequent offence.

This problem could be partially addressed by restricting the analysis to specific offences. However, we would still expect a substantial level of legitimate variability in sentencing within relatively diverse offences such as ABH or common assault. In order to overcome this problem of unexplained heterogeneity within types of offences, we can use a statistical technique known as 'multilevel modelling'. This technique allows us to control simultaneously for the multiple legal factors defining an offence while reporting different estimates of variability among courts.

The majority of examples of multilevel models applied to the study of consistency stem from the United States (Anderson, Kling and Stith, 1999, Anderson and Spohn, 2010, Ulmer, Light and Kramer, 2011). One exception is Pina-Sánchez and Linacre (2013), who applied these techniques to study consistency in sentence lengths for offences of assault, robbery and burglary in England and Wales. The authors reported substantial levels of consistency. Specifically, one of their models for assault estimated that only 1.8% of the variation in sentence length was attributable to sentencing differences across courts (Pina-Sánchez and Linacre, 2013, p. 1119). However, substantial disparities were found with regards to the way specific aggravating factors are taken into account. The most serious case related to disparities in the weight of the aggravating factor of 'sustained assault on the same victim'. For example, for an offence of ABH where this aggravating factor was present, but with no other aggravating or mitigating factors, the average custodial length was 8.1 months, ranging from 6.4 to 9.8 months depending on the court being sentenced (Pina-Sánchez and Linacre, 2013, p.1120).

## Analysis using the Crown Court's sentencing survey (CCSS)

### Compliance and exact matching

The first measure of consistency to be explored is the guideline compliance rates. The analysis compares the 'before' and 'after' scenarios for GBH with intent. Both the old and new guidelines recommended a minimum of three years custody, which makes this offence an ideal case to be studied using compliance rates since only one disposal type is covered. In 2011, 95% of GBH with intent offences were sentenced to custody, and a comparison between the same offences sentenced from January to May with those between July and December shows an increase of compliance rates from 94% to 97%.[8] This change is relatively modest, but it is statistically significant, which demonstrates a positive

effect of the new guideline.[9] Simply put, for the case of GBH with intent, sentencers were more likely to comply with the prescribed outcome after the new guideline came into force.

This preliminary analysis using compliance rates can be complemented with the implementation of exact matching. In particular, the analysis of GBH with intent offences presented above is explored through the use of matched groups using the seven most prevalent legal factors found in offences of assault:

- the number of previous convictions
- whether the offender:
  - showed remorse
  - pleaded guilty at first opportunity
  - was a member of a gang
  - acted under the influence of alcohol or drugs
  - targeted a vulnerable victim
  - perpetrated a repeated or sustained assault against the same victim

This approach permits estimation of consistency in sentencing for specific types of offences with a remarkable degree of accuracy.[10] We can explore the apparent high levels of compliance across GBH with intent offences, and determine how the proportion of cases attracting a custodial sentence varies according to the legal factors that define the offence. For example, in cases of GBH with intent with no previous convictions, guilty plea not entered at first opportunity, and no aggravating and mitigating factors, offenders are sentenced to custody 88.5% of the time. The degree of variability increases even more for offences like the one just described and where the offender also shows remorse. In such cases, the use of custody drops slightly to 85.5%. On the other hand, similar cases where no remorse was shown, and the assault was sustained on the same victim, sentences are remarkably consistent, with 96.2% of offenders being sentenced to custody.

Although very precise, these results are not perfect measures of consistency, since in addition to the type of offence and the seven most prevalent aggravating and mitigating factors, there are other relevant legal factors that were not considered in the generation of matched groups. It would be impossible to account for all the relevant legal factors using this approach, since for every additional factor, the number of groups is (at least) duplicated.[11] This fact highlights a trade-off between the precision with which matched groups are created and the number of observations

contained in each one of them. Thus, some of the variability reported here might be due to legitimate differences between the cases.

### Between court variability and multilevel modelling

If we replicate the study of the between-court variability used in Tarling (2006) and Mason et al. (2007), we can see that the average custodial rate for assault in 2011 across Crown Courts was 47.5%, with a majority of courts reporting a custodial rate close to this average. However, the level of geographical dispersion is much greater if we examine the range between courts with the highest and lowest custody rates, which amounts to 39%. This is shown in Figure 5.1, where the custody rates for each court are plotted, with Winchester (a custodial rate of 28.2%) and Dorchester (67.6%) at the lower and upper limits of that range.

Performing the same analysis for offences imposed before and after the new guideline came into force reveals a reduction in variability. The average custody rates amongst courts stayed at 47% in both periods, but their standard deviations and ranges decreased from 10.7 to 9.3 and from 53.7 to 42.7 percentage points, respectively, which suggests an improvement in consistency. The next analysis replicates those reported by Pina-Sánchez and Linacre (2013) while shifting the focus from custodial sentence length to the use of custodial sentences (as opposed to other disposals). This increases the sample from 5,527 to 12,508 cases and thus enhances the generalizability and precision of the analysis.

The models employed distinguish between the ten most common assault offences.[12] In addition, to address differences between cases, these models also control for the ten most prevalent aggravating and mitigating factors. Ideally, the analysis should incorporate all of the legal factors considered in the guidelines and recorded by the CCSS. However, the change in assault guidelines was followed by a change in the survey forms, resulting in substantial differences in structure and wording. In order to avoid inconsistencies, the analysis is restricted to factors presented similarly in both CCSS forms.

- Whether a guilty plea was entered
- Whether the plea was entered at the first reasonable opportunity
- The number of previous convictions taken into account
- Whether the offender
  - showed remorse
  - was the main carer of a dependent person, and/or perpetrated on a vulnerable person
  - was a member of a gang

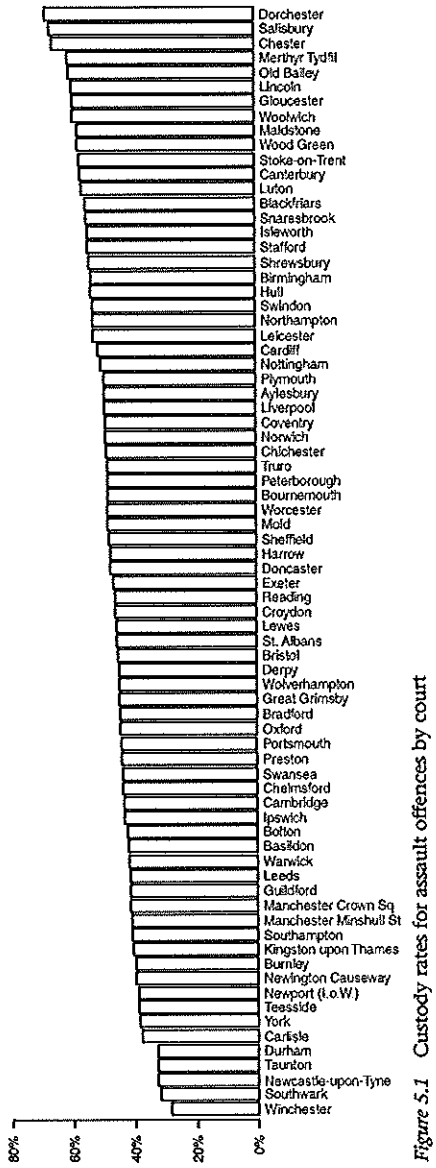


*Figure 5.1* Custody rates for assault offences by court

- committed an offence against a public worker
- was under the effect of drugs
- committed a sustained or repeated on the same victim

Results reveal levels of dispersion amongst courts that are relatively low although statistically significant.[13] For example, the custodial rate before the new guideline came into force for a reference case of an offence of common assault with no aggravating and mitigating factor ranged from 10% to 29%.[14] It is striking to compare this range of 19 percentage points with that of 39.4 percentage points found for all offences of assault shown in Figure 5.1. The two figures are not entirely comparable, as the former represents variability for a very specific type of common assault. However, this difference in ranges larger than 20 percentage points illustrates very clearly how biased measures of between court variability using descriptive statistics can be.

In addition, it is interesting to note that after the new guideline came into force the range for common assault narrowed to an interval ranging from 13% to 29%. This was due to a slight increase in sentence severity for common assault and a small reduction in between courts variability. Furthermore, although there are some differences in the use of custodial sentences between courts, those differences are not systematic. This is shown in Figure 5.2, which presents the different probabilities of being sentenced to custody for the reference case of common assault in the ten most extreme courts before and after the guideline came into force.

Specifically, none of the courts that were amongst the ten with the lowest or highest probabilities of sentencing to custody in the pre-guideline period can be found amongst the same bottom or top positions after the new guideline came into force. As can be seen in Figure 5.2,

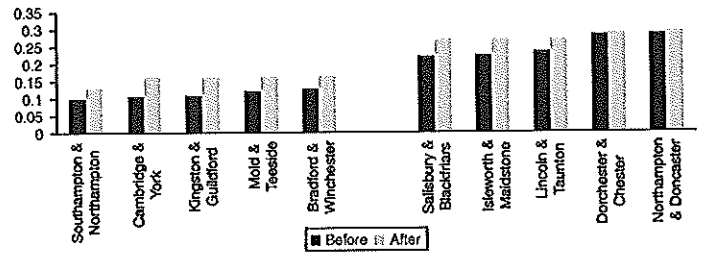


*Figure 5.2* Top ten courts with highest and lowest probabilities of custody for common assault before and after the new guideline came into force

Southampton, Cambridge, Kingston, Mold and Bradford led the ranking of 'most lenient' courts before the guideline came into force. Salisbury, Isleworth, Lincoln, Dorchester and Northampton were amongst the harshest. However, after the arrival of the new guideline, none of these courts could be found amongst the same extreme positions. This finding is interesting for two reasons. First, it challenges the results of Mason et al. that suggested systematic differences existed in the use of custody between courts over time. Second, it illustrates the greater accuracy on the measurement of consistency obtained using multilevel models, which suggest lower levels of variability than previously reported in the literature.

In order to explore the extent to which these positive findings can be generalised beyond cases of common assault, the analysis was replicated to other offences of assault such as affray, ABH, GBH, and GBH with intent, where no aggravating or mitigating factors were present in each of them. The results are presented in Figure 6.3, in which circles denote the average probability of custody for each reference case, and vertical lines around those circles reflect their between court variability.

Figure 5.3 reveals that the level of between court variability differs substantially across the cases studied. There are some cases that attract a higher consensual response than others. For offences where the probability to be sentenced to custody is either very high or very low (such as GBH with intent or common assault) the levels of consistency are

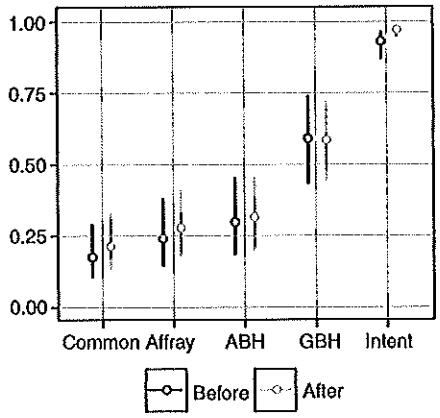


*Figure 5.3* Between courts variability in the probability of custody before and after the new guideline came into force[15]

remarkably high, whereas offences that have a probability of being sentenced to custody close to .5 (such as affray, ABH, and GBH) show lower levels of consistency. A second important finding to note is that consistency improved in all the offences studied after the new guideline came into force. That improvement was less noticeable for the reference cases of common assault and affray than for GBH with intent, which became almost perfectly consistent.

As can be seen, consistency improved in all the offences studied after the new guidelines came into effect. In addition, it is important to note that in spite of the accuracy achieved by the measures of variability plotted in Figure 5.3, they are not perfect measures since some of the between court variability will be due to legitimate legal factors considered by sentencers that could not be included in the models. Hence, the actual levels of between court variability will be lower than those reported here, which makes us more confident about the overall level of consistency amongst offences of assault sentenced in the Crown Courts.

Finally, to conclude the analysis, I present results obtained using multilevel models capable of estimating the between court variability in the application of specific aggravating and mitigating factors.[16] This approach can also explore the degree of consistency in the sentencing process, as opposed to the measures of consistency in the sentence outcome seen so far. In particular, I studied the effect of remorse and the aggravating factor 'sustained assault': two legal factors that were found to exercise a significant influence over custody rates under the exact matching approach. To observe the between court variability in the application of those factors, amongst different reference cases, before and after the new guideline came into force, results from their respective multilevel models were plotted in Figure 5.4.

Figure 5.4 reveals that the level of between court variability in the application of the aggravating factor, sustained assault, is extensive, which corroborates Pina-Sánchez and Linacre's findings (2013), where the degree of inconsistency was measured in custodial sentence length. On the other hand, for the application of remorse, we can see a more common approach across courts. These results could be due to the greater importance assigned to sustained assault than to remorse in the new assault guidelines: the former is a step one factor whilst the latter is a step two factor. However, given the high levels of disparity observed, we should consider the possibility that there are significant differences in the way this factor is interpreted by sentencers. In addition, as was observed in Figure 5.3, it is offences where the probability of being sentenced to custody is not close to 0 or 1 where variability

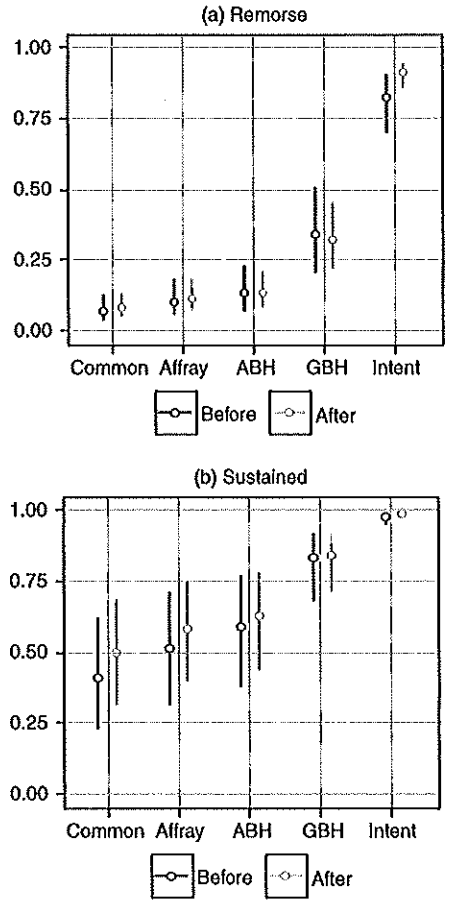


*Figure 5.4* Variability in the probability of custody, considering aggravating and mitigating factors before and after the new guideline came into force[17]

between courts is more noticeable. Finally, it is reassuring to observe that after the new guideline came into force, the level of variation between courts shrunk in all the cases studied, although that reduction was more striking in the application of the mitigating factor of remorse than the aggravating factor of a sustained assault.

## Conclusions

Until recently, studies on sentencing consistency in England and Wales were rare and methodologically limited. The few studies that attempted to study the elusive concept of consistency had to rely on methods unable to provide robust answers. The arrival of the CCSS has radically changed matters. Thanks to its level of detail with regards to the factors taken into account by courts, new and more accurate analyses have become possible (Sentencing Council, 2012; Pina-Sánchez and Linacre, 2013 and 2014; Roberts, 2013). However, our understanding of the subject remains partial since each of these studies has explored consistency using different samples of offences and different methodologies.

The analysis presented here can be summarised in two main conclusions:

1) The level of consistency in sentencing is higher than found in previous studies using simple descriptive statistics (e.g., Tarling, 2006; Mason et al., 2007). Whether the level of consistency is acceptable or not is a matter of opinion. My conclusion would be that the level of consistency for the Crown Courts of England and Wales in 2011 seems adequate, although some variation remains in the application of certain aggravating factors, such as sustained assault, that might need to be reviewed in future guidelines.
2) Consistency in sentencing improved after the new guidelines, regardless of the method used to assess that change. The majority of measures suggest modest but detectable increases in consistency after the new assault guideline came into force. More generally, results from the multilevel models showed decreases in between-court variability for all the most frequent offences of assault. In particular, considering a reference case of common assault with no aggravating or mitigating factors, the range in the use of custodial sentences between the most severe and the most lenient courts decreased from 19 to 16 percentage points. Moreover, none of the courts among the ten harshest and ten most lenient before the new guideline came into force could be found in the same bottom or top positions after the arrival of the guideline, which suggests an absence of systematic inconsistencies between courts in 2011.

Finally, more research into consistency is necessary. It could be argued that the new changes reported here take place over the long run, and the changes we have observed over a period of six months are only

temporary; or perhaps the higher consistency observed responds to the effect of a different process, such as the more binding nature of the guidelines brought about by the Criminal and Justice Act 2009. Additional releases of data from the CCSS will allow us to increase the statistical power and to increase the timespan covered by the analysis, enabling us to explore such issues in the future.

## Notes

I would like to thank Robin Linacre for his willingness to engage in methodological discussions on the measurement of consistency and for contributing to the analyses presented here.

1. For example, the definitive guideline on sentence reductions for a guilty plea (See Sentencing Guidelines Council, 2007).
2. The act provides that *'Every court must, in sentencing an offender, follow any sentencing guidelines which are relevant to the offender's case...unless the court is satisfied that it would be contrary to the interests of justice to do so'* (s.128 (1) (a) Coroners and Justice Act, 2009), whilst under previous legislation, courts were only required to 'have regard to' sentencing guidelines (s.172, Criminal Justice Act, 2003).
3. The new assault guideline can be downloaded at http://sentencingcouncil.judiciary.gov.uk/docs/Assault_definitive_guideline_-_Crown_Court.pdf.
4. The old assault guideline can be downloaded at http://webarchive.nationalarchives.gov.uk/20100305172947/http://www.sentencing-guidelines.gov.uk/guidelines/council/final.html.
5. This excerpt is part of an interview with the BBC, which can be downloaded at http://www.bbc.co.uk/news/uk-12681250.
6. See Chapter 1 of this volume for a description of the CCSS.
7. For example, the United States Sentencing Commission (1991) matched offences of cocaine/heroine distribution into groups using factors such as the amount of drugs, injury caused to any victims, the defendant's role in the offence, criminal record, and whether the defendant pleaded guilty.
8. The number of offences of GBH with intent sentenced in the before and after period as defined here are 449 and 546, respectively.
9. The difference in proportions was tested using the value of Pearson's chi-squared test statistic, which was equal to 4.24, with p-value .04.
10. In order to obtain robust estimates, I only considered matched groups that comprise at least 50 cases.
11. This problem is known as the curse of dimensionality (Bellman, 1961).
12. These are ABH, GBH, GBH with intent, common assault, affray, assault with the intent to resist an arrest, assault on a police constable, fear or provocation of violence, harassment, and violent disorder.
13. These findings are based on the statistically significant variance of the random intercepts both before and after the guideline came into force.
14. Results from the random intercepts multilevel logit models used here should be interpreted in terms of probabilities; however, for the sake of facilitating

the comparability of results with previous, they are often reported here as percentages.

15. The vertical lines around each circle represent the 95% confidence intervals calculated using the standard deviations of the random intercept term.
16. Specifically, this analysis is carried out using random slopes models, which could be considered to offer more robust measures of between court variability. See Pina-Sánchez and Linacre (2013) for a discussion on how random slopes estimates are offer more robust measures of consistency in the presence of omitted relevant variables than random intercept estimates.
17. The vertical lines around each circle represent the 95% confidence intervals calculated using the standard deviations of the random slopes term for remorse (left) and sustained assault (right).

## References


Anderson, J., Kling, J. and Stith, K. (1999) Measuring Inter-Judge Sentencing Disparity: Before and After the Federal Sentencing Guidelines. *Journal of Law & Economics*, 42: 271–307.

Anderson, A. and Spohn, C. (2010) Lawlessness in the Federal Sentencing Process: A Test for Uniformity and Consistency in Sentence Outcomes, *Justice Quarterly*, 27: 362–393.

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

Bellman, R. E. (1961) *Adaptive Control Processes: A Guided Tour*. Princeton: Princeton University Press.

Council of Europe (1993) Consistency in Sentencing. *Legal Issues*, 92. http://book.coe.int/EN/ficheouvrage.php?PAGEID=36&=EN&produit_aliasid=809.

Frankel, M. (1972) Lawlessness in Sentencing, *University of Cincinnati Law Review*, Vol. 41(1):1.

Frase, R.S. (2005) Sentencing Guidelines in Minnesota, 1978–2003. *Crime and Justice*, 32: 131–219.

Kramer, J. and Ulmer, J. (2002) Downward Departures for Serious Violent Offenders: Local Court 'Corrections' to Pennsylvania's Sentencing Guidelines. *Criminology*, 40: 807–932.

Mason, T., de Silva, N., Sharma, N., Brown, D. and Harper, G. (2007) *Local Variation in Sentencing in England and Wales*. London: Ministry of Justice.

Minnesota Sentencing Guidelines Commission (2012) Sentencing Practices: Controlled Substance Offenses Sentenced in 2010. http://www.leg.state.mn.us/docs/2013/other/130860.pdf.

Oregon Criminal Sentencing Commission (2003) Sentencing Practices: Summary Statistics for Felony Offenders Sentenced in 2001. http://www.oregon.gov/CJC/docs/SG01v2.pdf.

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

——(2014) *Impact of the 2011 England and Wales Assault Guidelines*. Available from authors.

Roberts, J. (2013) Complying with sentencing guidelines: latest findings from the Crown Court Sentencing Survey. In: Ashworth, A. and Roberts, J. (eds) *Sentencing Guidelines*. Oxford: Oxford University Press.

Roberts, J.V. and Bradford, B. (2014) Sentence Reductions for a Guilty Plea: New Empirical Evidence from England and Wales. *Journal of Empirical Legal Studies*, in press.

Roberts, J. and Plesnicar, M. (2015) Legitimacy and the Sentencing Process. In: G. Meszko and J. Tankebe (eds) *Legitimacy and Criminal Justice in Europe*. Amsterdam: Springer.

Sentencing Commission Working Group (2008) *Sentencing Guidelines in England and Wales: An Evolutionary Approach*. London: SCWG.

Sentencing Council (2012) *Crown Court Sentencing Survey: Annual Publication, 2011*. http://sentencingcouncil.judiciary.gov.uk/docs/CCSS_Annual_2011.pdf.

Sentencing Guidelines Council (2007) *Sentence Reductions for a Guilty Plea. Definitive Guideline*. London: Sentencing Guidelines Council.

Scott, R. (2010) Inter-Judge Sentencing Disparity After Booker: A First Look, *Express*. http://works.bepress.com/ryan_scott/2/.

Tonry, M. H. (1987) *Sentencing Reform Impacts*. Rockville: National Institute of Justice.

Ulmer, J., Light, M. 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.

Von Hirsch, A., Ashworth, A., Roberts, J. (2009) *Principled Sentencing: Readings on Theory and Policy*. Portland: Hart Publishing.

Walker, S., Spohn, C. and Delone, M. (2007) *The Color of Justice: Race, Ethnicity, and Crime in America*. Belmont: Wadsworth.