



Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

Exploring Disparities in Sentencing Using Multilevel Modelling: Opportunities and Pitfalls

Dr Jose Pina-Sánchez
Prof Ian Brunton-Smith
Dr Guangquan Li



Introduction

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- The principle of equality under the law requires consistent sentencing:

'Like cases are treated alike'



Introduction

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- The principle of equality under the law requires consistent sentencing:

‘Like cases are treated alike’

- When we do not have that we can expect:
 - Higher proportion of sentences appealed
 - Loss of trust in the criminal justice system



Introduction

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- The principle of equality under the law requires consistent sentencing:
‘Like cases are treated alike’
- When we do not have that we can expect:
 - Higher proportion of sentences appealed
 - Loss of trust in the criminal justice system
- In 2011 England and Wales created new sentencing guidelines
 - First jurisdiction to follow the example of the US



Introduction

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- The principle of equality under the law requires consistent sentencing:
‘Like cases are treated alike’
- When we do not have that we can expect:
 - Higher proportion of sentences appealed
 - Loss of trust in the criminal justice system
- In 2011 England and Wales created new sentencing guidelines
 - First jurisdiction to follow the example of the US
- Lots of important questions to be explored further:
 - What is the true extent of the problem?
 - What are the causes of unwarranted disparities? Courts, judges, or offenders characteristics?
 - Are sentencing guidelines having the desired effect?



Introduction

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- Multilevel modelling (MLM) is a really useful technique to explore these kinds of questions:



Introduction

Compliance with
the GuidelinesWithin Court
Disparities

Conclusion

- Multilevel modelling (MLM) is a really useful technique to explore these kinds of questions:
 - We can control for relevant factors and in so doing distinguish between warranted and unwarranted disparities

$$Y_{ij} = \beta_0 + \beta_k X_{ijk} + \zeta_j + \xi_{ij}$$



Introduction

Compliance with
the GuidelinesWithin Court
Disparities

Conclusion

- Multilevel modelling (MLM) is a really useful technique to explore these kinds of questions:
 - It allows us to control for relevant factors and in so doing distinguish between **warranted** and **unwarranted** disparities

$$Y_{ij} = \underbrace{\beta_0 + \beta_k X_{ijk}} + \underbrace{\zeta_j + \xi_{ij}}$$

Introduction

Compliance with
the GuidelinesWithin Court
Disparities

Conclusion

- Multilevel modelling (MLM) is a really useful technique to explore these kinds of questions:
 - It allows us to control for relevant factors and in so doing distinguish between **warranted** and **unwarranted** disparities

$$Y_{ij} = \underbrace{\beta_0 + \beta_k X_{ijk}} + \underbrace{\zeta_j + \xi_{ij}}$$

- We can distinguish between the level at which unwarranted disparities are generated

$$\zeta_j \sim N(0, \sigma_\zeta^2)$$

$$\xi_{ij} \sim N(0, \sigma_\xi^2)$$



Introduction

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- However, in some instances the standard MLM can be inadequately restrictive
 - I am guilty of having used it indiscriminately



Introduction

Compliance with
the GuidelinesWithin Court
Disparities

Conclusion

- However, in some instances the standard MLM can be inadequately restrictive
 - I am guilty of having used it indiscriminately
- I'll present two interesting extensions of the standard MLM
 - Multivariate-multilevel model
 - Scale-location model
- Both analyses are based on similar samples
 - Offences of assault sentenced in the England and Wales Crown Court
 - Controlling for more than 30 case characteristics
 - Only court IDs (no judge IDs) were used

Mind the Intermediate Steps

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- Sentencing is a highly complex process
 - multiple preliminary decisions are considered before the final sentence is passed
 - e.g. assessments of the seriousness of an offence → prescribed starting points
 - the England and Wales sentencing guidelines consider nine steps

Mind the Intermediate Steps

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- Sentencing is a highly complex process
 - multiple preliminary decisions are considered before the final sentence is passed
 - e.g. assessments of the seriousness of an offence → prescribed starting points
 - the England and Wales sentencing guidelines consider nine steps
- The standard MLM assumes a unique data-generating mechanism
 - this is often not realistic
 - we miss how intermediate decisions are made and how they affect each other
 - comparisons of effect sizes can be misleading and measures of uncertainty will not be as precise

Multivariate Multilevel Modelling

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- We used a multivariate-multilevel model to account for three key outcomes and the relationships between them
 - assessment of the offence seriousness (Step One)
 - guilty plea reduction (Step Four)
 - final sentence outcome

Multivariate Multilevel Modelling

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- We used a multivariate-multilevel model to account for three key outcomes and the relationships between them
 - assessment of the offence seriousness (Step One)
 - guilty plea reduction (Step Four)
 - final sentence outcome

$$Y_{1ij}^* = \beta_{1k} X_{ijk} + \zeta_{1j} \text{ (offence seriousness)}$$

$$Y_{2ij}^* = \beta_{2k} X_{ijk} + \zeta_{2j} \text{ (guilty plea reduction)}$$

$$Y_{3ij}^* = \beta_{3k} X_{ijk} + \alpha_1 Y_{1ij} + \alpha_2 Y_{2ij} + \zeta_{3ij} \text{ (final sentence outcome)}$$

Multivariate Multilevel Modelling

Introduction

Compliance with
the GuidelinesWithin Court
Disparities

Conclusion

- We used a multivariate-multilevel model to account for three key outcomes and the relationships between them
 - assessment of the offence seriousness (Step One)
 - guilty plea reduction (Step Four)
 - final sentence outcome

$$Y_{1ij}^* = \beta_{1k} X_{ijk} + \zeta_{1j} \text{ (offence seriousness)}$$

$$Y_{2ij}^* = \beta_{2k} X_{ijk} + \zeta_{2j} \text{ (guilty plea reduction)}$$

$$Y_{3ij}^* = \beta_{3k} X_{ijk} + \alpha_1 Y_{1ij} + \alpha_2 Y_{2ij} + \zeta_{3j} \text{ (final sentence outcome)}$$

$$\begin{pmatrix} \zeta_{1j} \\ \zeta_{2j} \\ \zeta_{3j} \end{pmatrix} \sim MVN \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\zeta_1}^2 & & \\ \sigma_{\zeta_{12}} & \sigma_{\zeta_2}^2 & \\ \sigma_{\zeta_{13}} & \sigma_{\zeta_{23}}^2 & \sigma_{\zeta_3}^2 \end{pmatrix} \right)$$



New Insights Using Multivariate MLM

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- Between court variances at each step were of a similar magnitude
 - No specific step is more inconsistently applied than the others



New Insights Using Multivariate MLM

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- Between court variances at each step were of a similar magnitude
 - No specific step is more inconsistently applied than the others
- None of the covariances between steps were significant
 - The same courts were not systematically harsher or more lenient across each step



New Insights Using Multivariate MLM

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- Between court variances at each step were of a similar magnitude
 - No specific step is more inconsistently applied than the others
- None of the covariances between steps were significant
 - The same courts were not systematically harsher or more lenient across each step
- Several step-specific factors were found to be double counted
 - e.g. *premeditation* was used to define the offence seriousness, but it is also affecting the final sentence directly

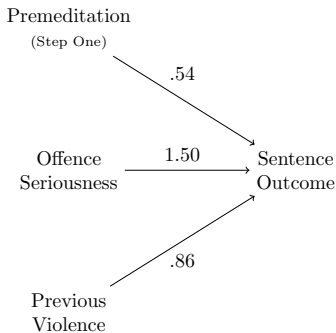
Considering Direct and Indirect Effects

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion



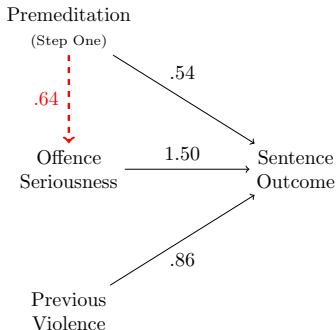
Considering Direct and Indirect Effects

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion





Compliance with the Guidelines

Introduction

**Compliance with
the Guidelines**

Within Court
Disparities

Conclusion

- Varying levels of compliance with the guidelines could be causing some of the unwarranted disparities detected



Compliance with the Guidelines

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- Varying levels of compliance with the guidelines could be causing some of the unwarranted disparities detected
- This last hypothesis is hard to test in England because the judges passing the sentence are not recorded



Compliance with the Guidelines

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- Varying levels of compliance with the guidelines could be causing some of the unwarranted disparities detected
- This last hypothesis is hard to test in England because the judges passing the sentence are not recorded
- We have relied on measures of between-court disparities to assess the level of consistency in sentencing
 - We found that roughly only 5% of unexplained variability was due to differences between courts
 - This fits nicely with the narrative of the Sentencing Council



Judge Identifiers Are Essential

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- However, this approach hides important disparities taking place within courts
 - Extreme disparities taking place within the same court could be cancelled out when looking at the court average
 - e.g. a court composed of two fair judges might seem as consistent as a court composed of an extremely harsh and an extremely lenient judge
 - By focusing on between-court disparities exclusively we have underestimated the magnitude of the problem



Within Court Disparities

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- One of the key assumptions of the standard MLM is constant residual variances

$$- Y_{ij} = \beta_0 + \beta_k X_{ijk} + \zeta_j + \xi_{ij}$$

$$\zeta_j \sim N(0, \sigma_\zeta^2); \quad \xi_{ij} \sim N(0, \sigma_\xi^2)$$

Within Court Disparities

Introduction

Compliance with
the GuidelinesWithin Court
Disparities

Conclusion

- One of the key assumptions of the standard MLM is constant residual variances

$$- Y_{ij} = \beta_0 + \beta_k X_{ijk} + \zeta_{1j} + \xi_{ij}$$
$$\zeta_{1j} \sim N(0, \sigma_{\zeta_1}^2); \quad \xi_{ij} \sim N(0, \underbrace{\sigma_{\xi_j}^2})$$

- We can relax that assumption using a location-scale model

$$\ln(\sigma_{\xi_{ij}}^2) = \alpha + \zeta_{2j}$$

Within Court Disparities

Introduction

Compliance with
the GuidelinesWithin Court
Disparities

Conclusion

- One of the key assumptions of the standard MLM is constant residual variances

$$- Y_{ij} = \beta_0 + \beta_k X_{ijk} + \zeta_{1j} + \xi_{ij}$$

$$\zeta_{1j} \sim N(0, \sigma_{\zeta_1}^2); \quad \xi_{ij} \sim N(0, \underbrace{\sigma_{\xi_j}^2})$$

- We can relax that assumption using a location-scale model

$$\ln(\sigma_{\xi_{ij}}^2) = \alpha + \zeta_{2j}$$

$$\begin{pmatrix} \zeta_{1j} \\ \zeta_{2j} \end{pmatrix} \sim BVN \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\zeta_1}^2 & \\ & \sigma_{\zeta_2}^2 \end{pmatrix} \right)$$

Between AND Within Court Disparities

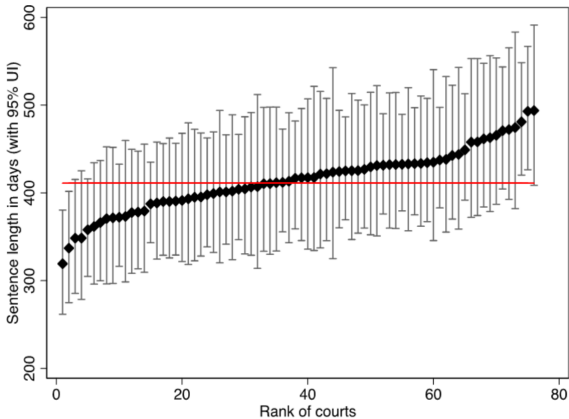
Introduction

Compliance with
the Guidelines

**Within Court
Disparities**

Conclusion

Average sentence by court



Between AND Within Court Disparities

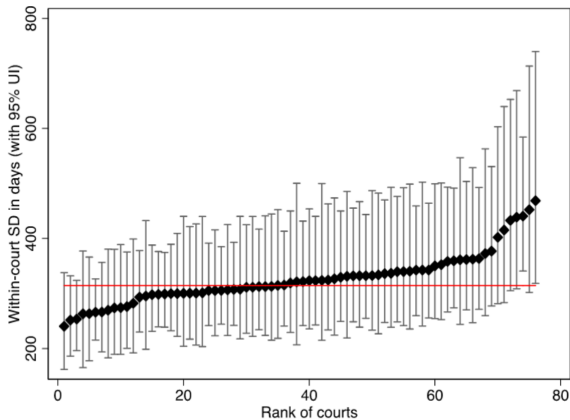
Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

Within court disparities by court





Within Court Disparities

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- We found substantial unwarranted disparities between courts
 - On average (across all courts) the typical case of assault with body harm is sentenced to 412 days
 - The same case when sentenced in the most lenient court receives just 319 days
 - While on the harshest court it will be 494
 - Perhaps not a big deal if we take into consideration that we are comparing the most extreme courts out of 76



Within Court Disparities

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- We found substantial unwarranted disparities between courts
 - On average (across all courts) the typical case of assault with body harm is sentenced to 412 days
 - The same case when sentenced in the most lenient court receives just 319 days
 - While on the harshest court it will be 494
 - Perhaps not a big deal if we take into consideration that we are comparing the most extreme courts out of 76
- Yet, we also need to consider even larger within court disparities
 - The average within court std deviation is 320 days
 - And this can also range from 240 days in the most consistent court
 - To 469 days in the most inconsistent court



Within Court Disparities

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- Disparities within courts are more substantial than disparities between courts
 - Accounting for between-judge disparities is absolutely crucial
 - Assessments based just on between-court disparities overestimate the level of consistency in sentencing



Within Court Disparities

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- Disparities within courts are more substantial than disparities between courts
 - Accounting for between-judge disparities is absolutely crucial
 - Assessments based just on between-court disparities overestimate the level of consistency in sentencing
- This is relevant policywise
 - The Sentencing Council has been targeting courts sentencing systematically harsher or more leniently than the average

Within Court Disparities

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- Disparities within courts are more substantial than disparities between courts
 - Accounting for between-judge disparities is absolutely crucial
 - Assessments based just on between-court disparities overestimate the level of consistency in sentencing
- This is relevant policywise
 - The Sentencing Council has been targeting courts sentencing systematically harsher or more leniently than the average
- To identify problematic courts a new measure of consistency should be used
 - We suggest the mean square error
 - Accounting for the court's bias ($\zeta_{1j}^2 - \beta_0$) and precision ($e^{\alpha + \zeta_{2j}^2}$)

$$mse = \sqrt{(\zeta_{1j}^2 - \beta_0)^2 + e^{(\alpha + \zeta_{2j}^2)}}$$



Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

Conclusion

- MLM is a really useful technique for the study of consistency in sentencing



Conclusion

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- MLM is a really useful technique for the study of consistency in sentencing
- In some instances, by using extensions of the standard MLM we can obtain more robust and insightful findings
 - Case characteristics are often double counted (possibly generating inconsistencies)
 - The level of internal consistency varies substantially across courts (cannot be disregarded)



Conclusion

Introduction

Compliance with
the Guidelines

Within Court
Disparities

Conclusion

- MLM is a really useful technique for the study of consistency in sentencing
- In some instances, by using extensions of the standard MLM we can obtain more robust and insightful findings
 - Case characteristics are often double counted (possibly generating inconsistencies)
 - The level of internal consistency varies substantially across courts (cannot be disregarded)
- Progress on this area relies importantly on the promotion of these types of methodological debates