

British Society of Criminology Conference 2018

Tackling Selection Bias in Sentence Data Analysis

Have the England and Wales Guidelines Increased Sentencing Severity?

Jose Pina-Sánchez (University of Leeds)

Hye-In Chung (University of Leeds)

John Paul Gosling (University of Leeds)

Sara Geneletti (London School of Economics)

Marco Dorreti (University of Perugia)

Amber Isaac (Sentencing Council)

Elizabeth Bourgeois (University of Leeds)

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Our project seeks to deal with the problem of selection bias in sentence data
 - Different disposal types use different metrics (e.g. pounds for fines, days for custodial sentences)
 - For reasons of convenience we tend to focus on custodial sentences
 - However these represent only 7% of the sentences imposed in England and Wales

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Various statistical adjustments have been applied
 - Two stage processes (Heckman model)
 - Models for censored data (Tobit model)

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Various statistical adjustments have been applied
 - Two stage processes (Heckman model)
 - Models for censored data (Tobit model)
 - But the assumptions upon which they are built are not realistic
 - And waste information available for non-custodial sentences

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Various statistical adjustments have been applied
 - Two stage processes (Heckman model)
 - Models for censored data (Tobit model)
 - But the assumptions upon which they are built are not realistic
 - And waste information available for non-custodial sentences
- We suggest an alternative approach based on the estimation of a scale of severity
 - Advocated in the 80s (Buchner, 1979; Erickson and Gibbs, 1979; Sebba, 1980; Sebba and Nathan, 1984)
 - Oddly, abandoned since then
 - Recently picked up by the Sentencing Council to carry out the evaluation of their guidelines

The Council's Scale: Rationale

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

Grievous Bodily Harm with Intent

Offence Category	Starting Point <i>(Applicable to all offenders)</i>
Category 1	12 years' custody
Category 2	6 years' custody
Category 3	4 years' custody

The Council's Scale: Rationale

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

Grievous Bodily Harm

Offence Category	Starting Point (<i>Applicable to all offenders</i>)
Category 1	3 years' custody
Category 2	1 year 6 months' custody
Category 3	High level community order

The Council's Scale: Rationale

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Based on the starting points for different levels of seriousness encoded in the guidelines
- Estimates the function mapping level of seriousness and duration of custodial sentences
 - e.g. custodial sentence length increases in 75% when going one category up in severity

The Council's Scale: Rationale

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Based on the starting points for different levels of seriousness encoded in the guidelines
- Estimates the function mapping level of seriousness and duration of custodial sentences
 - e.g. custodial sentence length increases in 75% when going one category up in severity
- Decides on one value of severity for 1-year in custody to anchor the scale
- Extrapolates to non-custodial outcomes

The Council's Scale: Problems

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- It assumes that the jumps in seriousness have an equivalent increase in severity across different offences
- The choice of a value of severity for a particular sentence outcome seems arbitrary
- Doesn't consider suspended sentences and conditional discharges

Measuring Sentence Severity: Literature

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Four main methods have been used:
 - Direct ratings (Hindelang, et al., 1975)
 - Magnitude escalation (Leclerc and Tremblay, 2016)
 - Thurstone pair-comparisons (Buchner, 1979)
 - Canonical correlation / correspondence analysis (Francis et al., 2005)

Measuring Sentence Severity: Literature

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Four main methods have been used:
 - Direct ratings (Hindelang, et al., 1975)
arbitrary; unreplicable
 - Magnitude escalation (Leclerc and Tremblay, 2016)
 - Thurstone pair-comparisons (Buchner, 1979)
 - Canonical correlation / correspondence analysis (Francis et al., 2005)

Measuring Sentence Severity: Literature

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Four main methods have been used:
 - Direct ratings (Hindelang, et al., 1975)
arbitrary; unreplicable
 - Magnitude escalation (Leclerc and Tremblay, 2016)
assumes numeracy of subjects; vast variability in responses
 - Thurstone pair-comparisons (Buchner, 1979)
 - Canonical correlation / correspondence analysis (Francis et al., 2005)

Measuring Sentence Severity: Literature

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Four main methods have been used:
 - Direct ratings (Hindelang, et al., 1975)
arbitrary; unreplicable
 - Magnitude escalation (Leclerc and Tremblay, 2016)
assumes numeracy of subjects; vast variability in responses
 - Thurstone pair-comparisons (Buchner, 1979)
based on untestable parametric assumptions
 - Canonical correlation / correspondence analysis (Francis et al., 2005)

Measuring Sentence Severity: Literature

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Four main methods have been used:
 - Direct ratings (Hindelang, et al., 1975)
arbitrary; unreplicable
 - Magnitude escalation (Leclerc and Tremblay, 2016)
assumes numeracy of subjects; vast variability in responses
 - Thurstone pair-comparisons (Buchner, 1979)
based on untestable parametric assumptions
 - Canonical correlation / correspondence analysis (Francis et al., 2005)
assumes perfect linear correlation between crime seriousness and sentence severity; generates nonsensical values

Our Approach

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Thurstone model and a sample of 21 magistrates
 - Rather than asking to compare pairs of sentences
 - We ask how often a particular disposal type can be more punitive than other
 - The questionnaire includes eleven sentence outcomes and seven comparisons
 - This gives us a matrix of severity that can be used to estimate severity scores

Matrix of Severity

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

	absolute discharge	conditional discharge	fine	community order	6month suspended 1month custody	12month suspended 1month custody	6month suspended 6month custody	24month suspended 12month custody	1month immediate custody	2month immediate custody	3month immediate custody
absolute discharge	0.5	1	1	1	1	1	1	1	1	1	1
conditional discharge	0	0.5	0.69	1	1	1	1	1	1	1	1
fine	0	0.31	0.5	0.78	1	1	1	1	1	1	1
community order	0	0	0.22	0.5	0.37	1	1	1	1	1	1
6month susp 1month cust	0	0	0	0.63	0.5	1	1	1	1	1	1
12month susp 1month cust	0	0	0	0	0	0.5	0.73	1	1	1	1
6month susp 6month cust	0	0	0	0	0	0.27	0.5	1	1	1	1
24month susp 12month cust	0	0	0	0	0	0	0	0.5	0.41	0.52	0.62
1month immediate custody	0	0	0	0	0	0	0	0.59	0.5	1	1
2month immediate custody	0	0	0	0	0	0	0	0.48	0	0.5	1
3month immediate custody	0	0	0	0	0	0	0	0.38	0	0	0.5

Comparison of Scales

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

Sentence outcome	Council's scale	Our scale
absolute discharge		0
conditional discharge		0.97
fine		1.32
community order		2.12
1-month custody 6-months suspended		2.34
1-month custody 12-months suspended		3.66
6-months custody 6-months suspended		3.78
12-months custody 24-months suspended		5.74
1-month custody		5.05
2-months custody		5.75
3-months custody		6.45
12-months custody		
5-years custody		
20-years custody		

Comparison of Scales

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

Sentence outcome	Council's scale	Our scale
absolute discharge		0
conditional discharge		0.97
fine		1.32
community order		2.12
1-month custody 6-months suspended		2.34
1-month custody 12-months suspended		3.66
6-months custody 6-months suspended		3.78
12-months custody 24-months suspended		5.74
1-month custody		5.05
2-months custody		5.75
3-months custody		6.45
12-months custody		13.45
5-years custody		47.05
20-years custody		173.05

Comparison of Scales

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

Sentence outcome	Council's scale	Our scale
absolute discharge	0	0
conditional discharge	0	0.97
fine	10.25	1.32
community order	21.65	2.12
1-month custody 6-months suspended	-	2.34
1-month custody 12-months suspended	-	3.66
6-months custody 6-months suspended	-	3.78
12-months custody 24-months suspended	-	5.74
1-month custody	29.37	5.05
2-months custody	29.93	5.75
3-months custody	30.48	6.45
12-months custody	35.26	13.45
5-years custody	55.89	47.05
20-years custody	100	173.05

Comparison of Scales

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

Sentence outcome	Council's scale	Our scale
absolute discharge	0	0
conditional discharge	0	0.97
fine	10.25	1.32
community order	21.65	2.12
1-month custody 6-months suspended	-	2.34
1-month custody 12-months suspended	-	3.66
6-months custody 6-months suspended	-	3.78
12-months custody 24-months suspended	-	5.74
1-month custody	29.37	5.05
2-months custody	29.93	5.75
3-months custody	30.48	6.45
12-months custody	35.26	13.45
5-years custody	55.89	47.05
20-years custody	100	173.05

Comparison of Scales

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

Sentence outcome	Council's scale	Our scale
absolute discharge	0	0
conditional discharge	0	0.97
fine	10.25	1.32
community order	21.65	2.12
1-month custody 6-months suspended	-	2.34
1-month custody 12-months suspended	-	3.66
6-months custody 6-months suspended	-	3.78
12-months custody 24-months suspended	-	5.74
1-month custody	29.37	5.05
2-months custody	29.93	5.75
3-months custody	30.48	6.45
12-months custody	35.26	13.45
5-years custody	55.89	47.05
20-years custody	100	173.05

Comparison of Scales

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

Sentence outcome	Council's scale	Our scale
absolute discharge	0	0
conditional discharge	0	0.97
fine	10.25	1.32
community order	21.65	2.12
1-month custody 6-months suspended	-	2.34
1-month custody 12-months suspended	-	3.66
6-months custody 6-months suspended	-	3.78
12-months custody 24-months suspended	-	5.74
1-month custody	29.37	5.05
2-months custody	29.93	5.75
3-months custody	30.48	6.45
12-months custody	35.26	13.45
5-years custody	55.89	47.05
20-years custody	100	173.05

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- We explore the increase in sentence severity in E&W (Roberts and Ashworth, 2016)
 - Which can be visualised much more clearly using our scale of severity
- Test whether the new sentencing guidelines are to be blamed (Allen, 2016)
 - Two phenomena that coincide in time
 - But correlation is not a sufficient condition for causation

Monitoring Sentence Severity

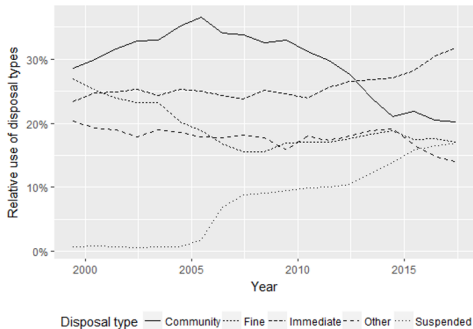
Introduction

Measuring
 Sentence
 Severity

Sentencing
 Guidelines

Conclusion

B - Relative use of disposal types (indictable offences)



Monitoring Sentence Severity

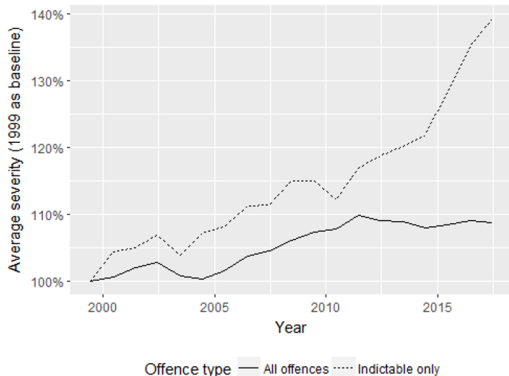
Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

B - Average severity (1999 as baseline)



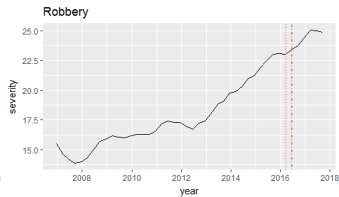
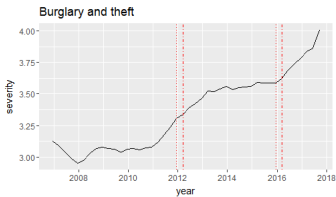
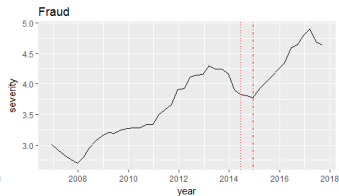
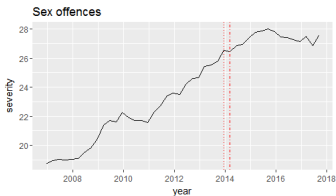
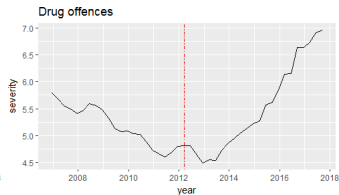
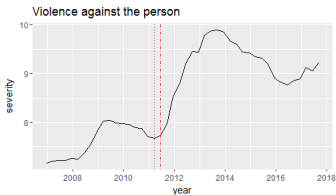


Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

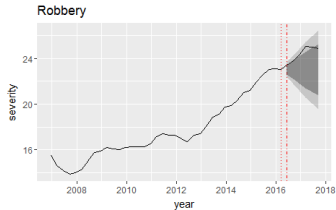
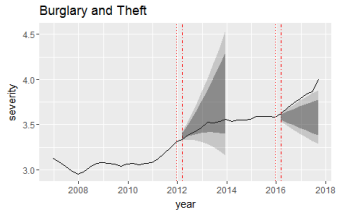
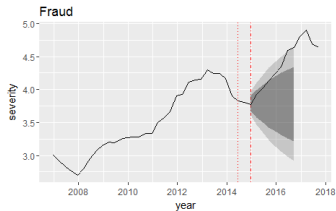
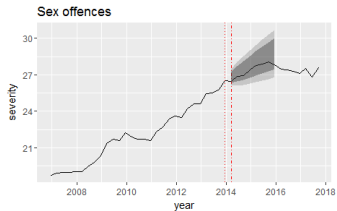
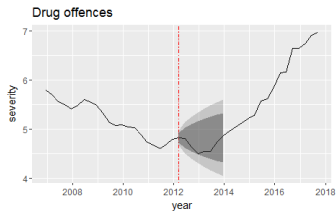
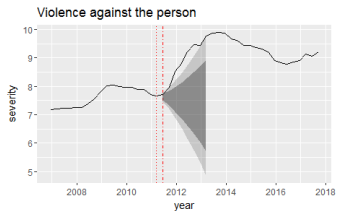


Introduction

Measuring
 Sentence
 Severity

Sentencing
 Guidelines

Conclusion



- Selection bias is an extremely pervasive problem in sentence data analyses
 - Affecting the validity of studies around key Criminology/Criminal Justice questions:
 - Effects of sentencing policy (sentencing guidelines)
 - Presence of discrimination (the Lammy review)
 - Deterrence and recidivism (Von Hirsch et al. 1999)

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Selection bias is an extremely pervasive problem in sentence data analyses
 - Affecting the validity of studies around key Criminology/Criminal Justice questions:
 - Effects of sentencing policy (sentencing guidelines)
 - Presence of discrimination (the Lammy review)
 - Deterrence and recidivism (Von Hirsch et al. 1999)
- Solutions suggested in the literature are not valid
- The estimation of scale of severity might be the way forward

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion

- Selection bias is an extremely pervasive problem in sentence data analyses
 - Affecting the validity of studies around key Criminology/Criminal Justice questions:
 - Effects of sentencing policy (sentencing guidelines)
 - Presence of discrimination (the Lammy review)
 - Deterrence and recidivism (Von Hirsch et al. 1999)
- Solutions suggested in the literature are not valid
- The estimation of scale of severity might be the way forward
 - We will never be able to obtain perfect measures of sentence severity
 - But we can improve their estimation process
 - Help us run more realistic/robust sentence data analysis
 - e.g. most guidelines cannot be blamed for the increase in severity

Introduction

Measuring
Sentence
Severity

Sentencing
Guidelines

Conclusion