Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

MoJ Analytical Seminar Series

Tackling Selection Bias in Sentence Data Analysis Using a Scale of Severity and Bayesian Statistics

Jose Pina-Sánchez (University of Leeds)
John Paul Gosling (University of Leeds)
Sara Geneletti (London School of Economics)

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Introduction

- Lots of important research questions can be explored through the analysis of sentence data

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Introduction

- Lots of important research questions can be explored through the analysis of sentence data
- We can look into the decision-making process of judges
  - Investigate the presence of heuristics in sentencing
  - How different aggravating and mitigating factors are used
  - Compliance with the sentencing guidelines

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Introduction

- Lots of important research questions can be explored through the analysis of sentence data
- We can look into the decision-making process of judges
  - Investigate the presence of heuristics in sentencing
  - How different aggravating and mitigating factors are used
  - Compliance with the sentencing guidelines
- Measure key principles guiding the sentencing process
  - Whether subgroups of the population are discriminated
  - The extent to which sentencing is consistent
  - Individualised
  - Proportional
  - Or affected by penal populism

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Introduction

- Lots of important research questions can be explored through the analysis of sentence data
- We can look into the decision-making process of judges
  - Investigate the presence of heuristics in sentencing
  - How different aggravating and mitigating factors are used
  - Compliance with the sentencing guidelines
- Measure key principles guiding the sentencing process
  - Whether subgroups of the population are discriminated
  - The extent to which sentencing is consistent
  - Individualised
  - Proportional
  - Or affected by penal populism
- Assess the effectiveness of different punishments in
  - Deterring crime
  - Fostering compliance
  - Promoting trust in the Criminal Justice System

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

The Problem

- Five main sentence outcomes (aka disposal types)
  - *discharge* < *fine* < *community order* < *suspended sentence* < *custodial sentence*
- Most of those disposal types use different units of measurement
  - e.g. pounds for fines, days for custodial sentences, conditions for community orders

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

The Problem

- Five main sentence outcomes (aka disposal types)
  - *discharge* < *fine* < *community order* < *suspended sentence* < *custodial sentence*
- Most of those disposal types use different units of measurement
  - e.g. pounds for fines, days for custodial sentences, conditions for community orders
- For reasons of convenience we tend to focus on custodial sentences
  - However these represent only 7% of the sentences imposed in England and Wales
  - Creating a massive problem of selection bias
- Alternatively some studies focus on the probability of custody
  - This involves reducing the sentence outcome to a (0,1) variable
  - A remarkable loss of information

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Current Strategies

- Various statistical adjustments have been applied to tackle the problem of selection bias
  - But the assumptions upon which they are built are questionable (at least in England & Wales)
  - And keep treating non-custodial cases as a homogeneous group

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Current Strategies

- Various statistical adjustments have been applied to tackle the problem of selection bias
  - But the assumptions upon which they are built are questionable (at least in England & Wales)
  - And keep treating non-custodial cases as a homogeneous group
- Two stage processes (Heckman selection model)
  - Assume that sentencing is undertaken in two steps
  - Require variables that meet the exclusion criteria

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Current Strategies

- Various statistical adjustments have been applied to tackle the problem of selection bias
  - But the assumptions upon which they are built are questionable (at least in England & Wales)
  - And keep treating non-custodial cases as a homogeneous group
- Two stage processes (Heckman selection model)
  - Assume that sentencing is undertaken in two steps
  - Require variables that meet the exclusion criteria
- Model for censored data (Tobit model)
  - Assume that sentencing is a one-step decision process
  - Assume that non-custodial sentences are part of the same distribution (normal) as custodial durations

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

A Scale of Severity

- We suggest alternative approaches 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)
  - Strangely abandoned since then (a few exceptions; Tremblay, 2016)
  - Recently picked up by the Sentencing Council for England and Wales

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

A Scale of Severity

- We suggest alternative approaches 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)
  - Strangely abandoned since then (a few exceptions; Tremblay, 2016)
  - Recently picked up by the Sentencing Council for England and Wales
- key benefit: the analysis of 100% of the offences, while making the most of the information available
  - MoJ data captures disposal types, and durations of suspended and custodial sentences

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

A Scale of Severity

- We suggest alternative approaches 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)
  - Strangely abandoned since then (a few exceptions; Tremblay, 2016)
  - Recently picked up by the Sentencing Council for England and Wales
- key benefit: the analysis of 100% of the offences, while making the most of the information available
  - MoJ data captures disposal types, and durations of suspended and custodial sentences
- key challenge: to estimate the relative severity of different sentence outcomes

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Methods Used in the Literature

- 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)

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Methods Used in the Literature

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

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Methods Used in the Literature

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

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Methods Used in the Literature

- Four main methods have been used:
  - Direct ratings (Hindelang, et al., 1975)
    arbitrary; unreplicable; uninformative
  - Magnitude escalation (Leclerc and Tremblay, 2016)
    unreliable; vast variability in responses; assumes numeracy
  - Thurstone pair-comparisons (Buchner, 1979)
    assumes severity around specific sentences is normally distributed with known variance
  - Canonical correlation / correspondence analysis (Francis et al., 2005)

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Methods Used in the Literature

- Four main methods have been used:
  - Direct ratings (Hindelang, et al., 1975)
    arbitrary; unreplicable; uninformative
  - Magnitude escalation (Leclerc and Tremblay, 2016)
    unreliable; vast variability in responses; assumes numeracy
  - Thurstone pair-comparisons (Buchner, 1979)
    assumes severity around specific sentences is normally distributed with known variance
  - Canonical correlation / correspondence analysis (Francis et al., 2005)
    assumes perfect linear correlation between crime seriousness and sentence severity; can generate nonsensical severity scores

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Our Approach: Thurstone's Method

- Thurstone method 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
  - Not all combinations of pairs were included
  - Only those where an overlap in the level of severity is expected
  - e.g. high community orders attaching multiple and long requirements can be harsher than suspended sentences with no onerous conditions attached

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Question Format

4. How often can a **12-month custodial sentence suspended for 24 months** be more punitive than a **1-month immediate custodial sentence**? (Please consider all possible combinations of conditions that can be attached to a suspended sentence and circumstances of the offender)

- ○ Never (0% of times)
- ○ Almost never (10% of times)
- ○ Very infrequently (20% of times)
- ○ Infrequently (30% of times)
- ○ Almost as often (40% of times)
- ○ As often (50% of times)
- ○ More often than not (60% of times)
- ○ Often (70% of times)
- ○ Very often (80% of times)
- ○ Almost always (90% of times)
- ○ Always (100% of times)
- ○ Other

a. If you selected Other, please specify:

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Matrix of Severity

	absolute discharge	conditional discharge	fine	community order	6month suspended 1month custody	12month suspended 1month custody	6month suspended 6month custody	24month suspended 12month custody	1month inmediate custody	2month inmediate custody	3month inmediate 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 inmediate custody	0	0	0	0	0	0	0	0.59	0.5	1	1
2month inmediate custody	0	0	0	0	0	0	0	0.48	0	0.5	1
3month inmediate custody	0	0	0	0	0	0	0	0.38	0	0	0.5

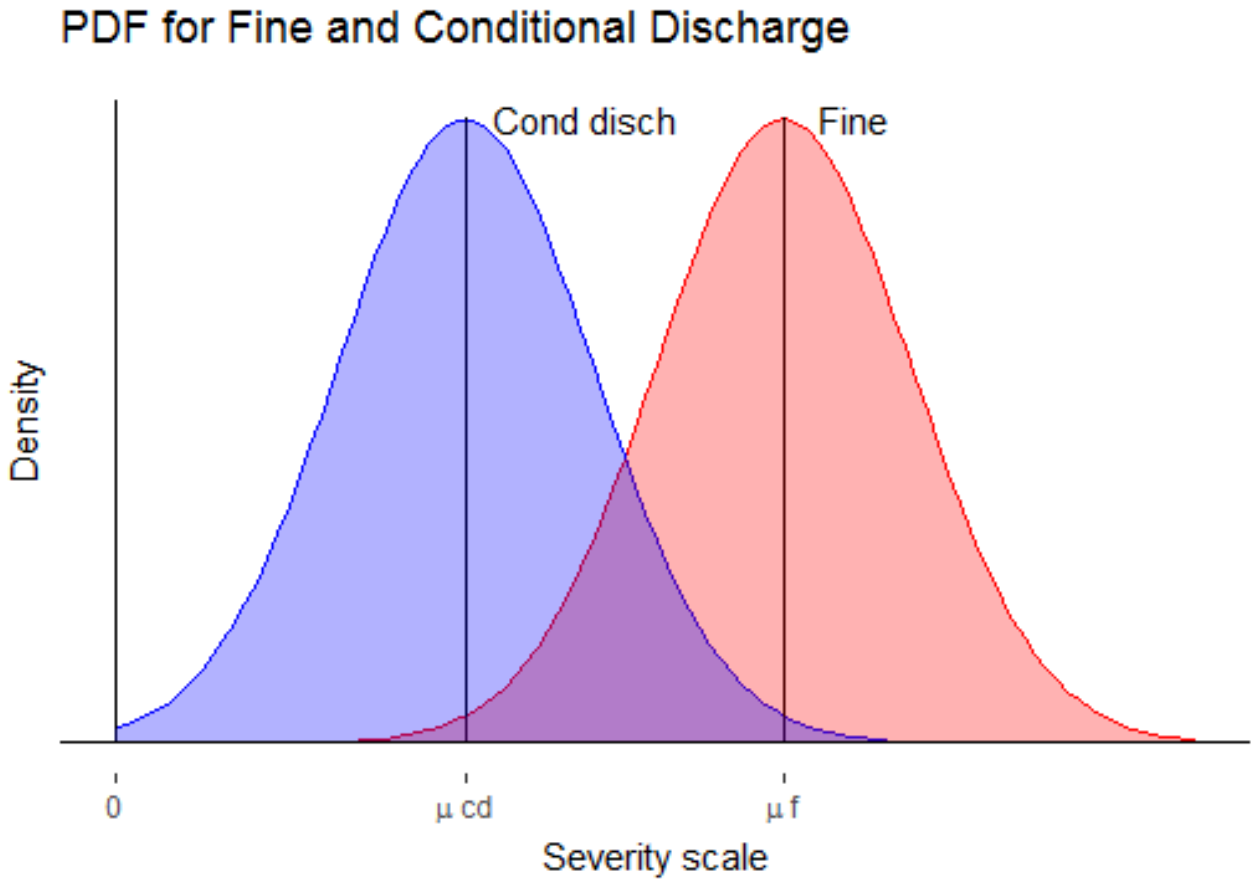
Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Thurstone Model: Intuitively

- We use the Thurstone-Mosteller model (Type V) to convert the proportions from pairwise comparisons into a severity scale
- Based on latent normal distributions for each sentence outcome included
- Each of those normal distributions will have its own mean, μ_s, and identical variance
- The amount of overlap between the distributions determines their closeness on the severity scale, i.e. their severity score, μ_s

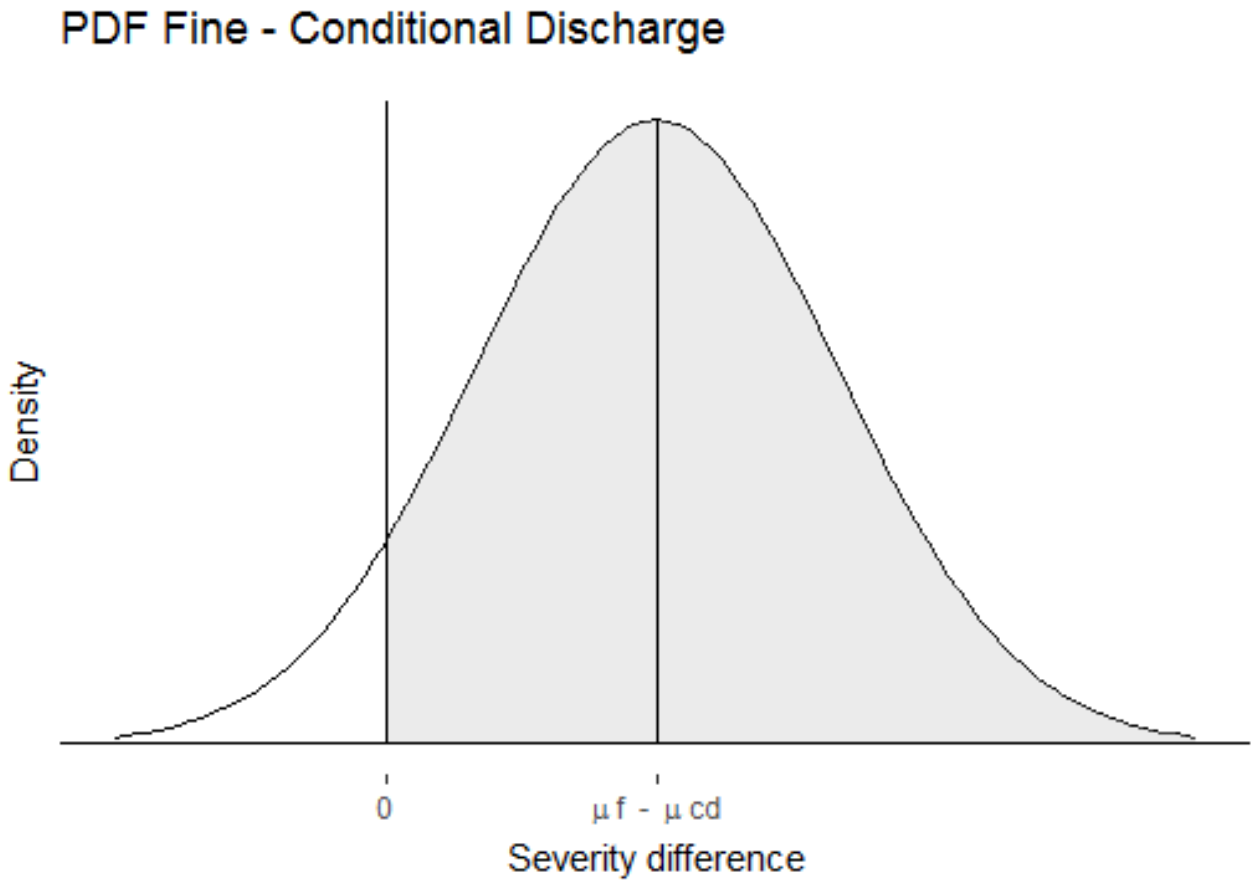
Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Thurstone Model: Visually



Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Thurstone Model: Visually



Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Severity Scores

Sentence outcome	Severity score
absolute discharge	0
conditional discharge	0.97
fine	1.33
community order	2.13
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	

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Severity Scores

Sentence outcome	Severity score
absolute discharge	0
conditional discharge	0.97
fine	1.33
community order	2.13
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

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Sensitivity Analyses

- The following sensitivity analyses were explored
  - Bradley-Terry model assuming logistic distributions of severity
  - 50-50 sample split
  - Change of population, a sample of 17 sentencing academics

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Sensitivity Analyses

- The following sensitivity analyses were explored
  - Bradley-Terry model assuming logistic distributions of severity
  - 50-50 sample split
  - Change of population, a sample of 17 sentencing academics
- No substantive differences were found
  - Values in the matrix of severity were relatively similar
  - Correlation coefficients between the different scales above .95

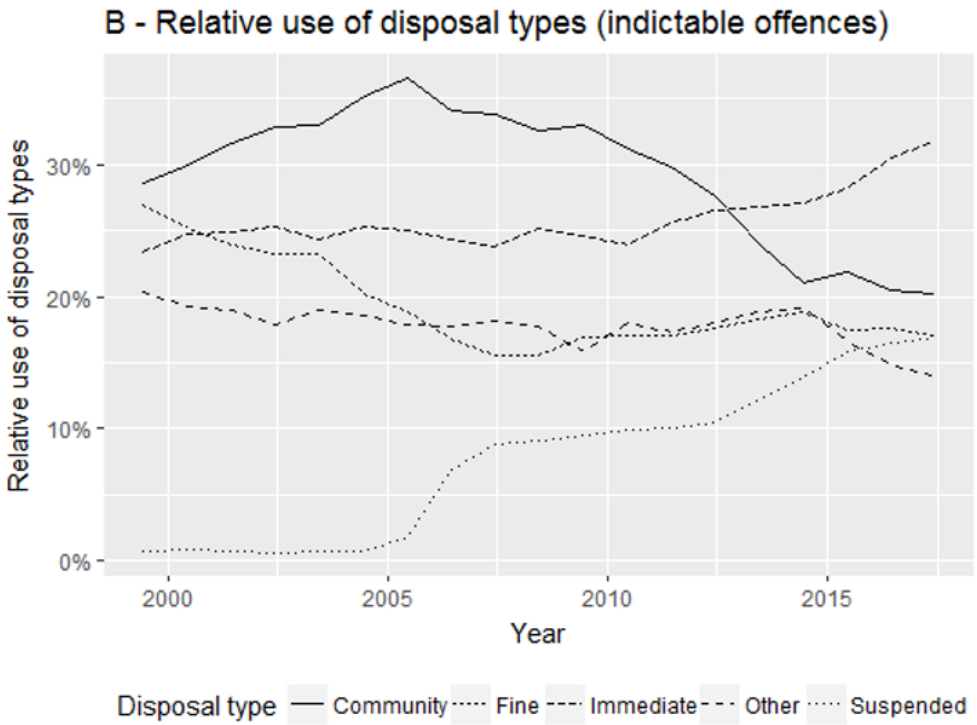
Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Sensitivity Analyses

- The following sensitivity analyses were explored
  - Bradley-Terry model assuming logistic distributions of severity
  - 50-50 sample split
  - Change of population, a sample of 17 sentencing academics
- No substantive differences were found
  - Values in the matrix of severity were relatively similar
  - Correlation coefficients between the different scales above .95
- Other important features and assumptions influence severity scores substantially
  - Most cells were locked
  - Distance between severity scores substantially affected by the assumption of equal variance

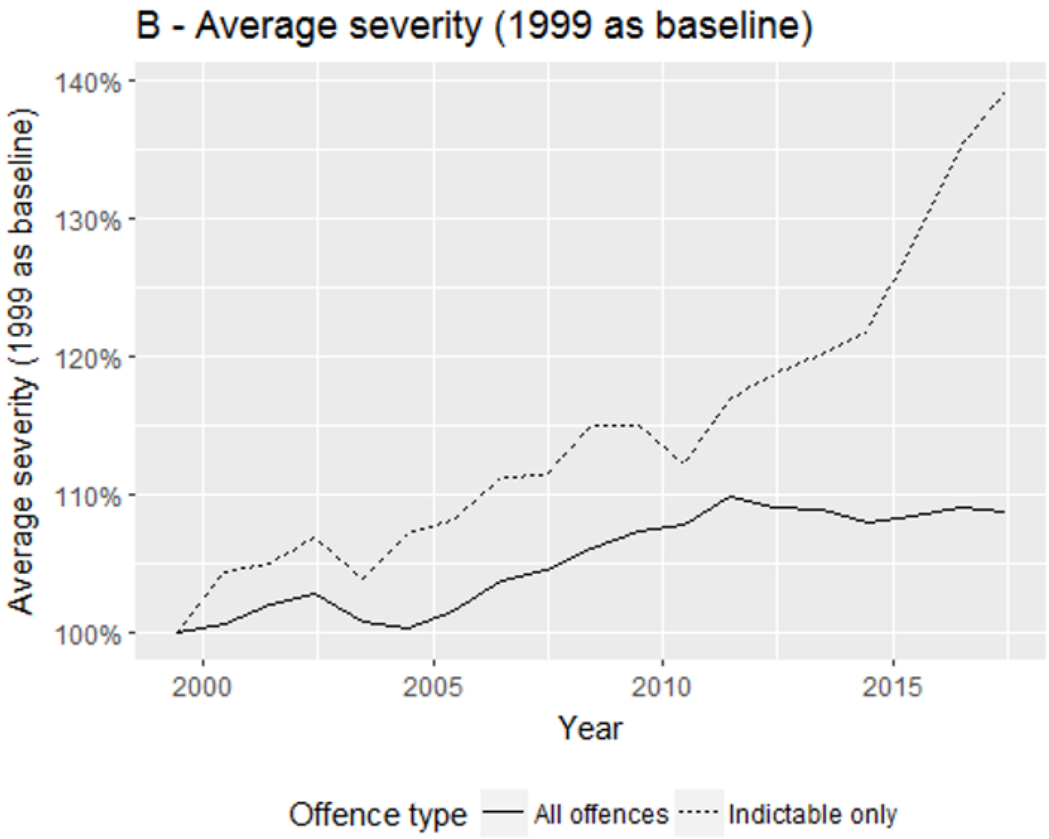
Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Monitoring Severity

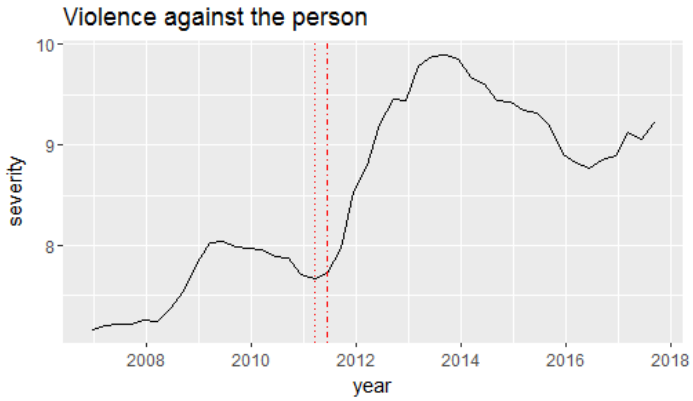


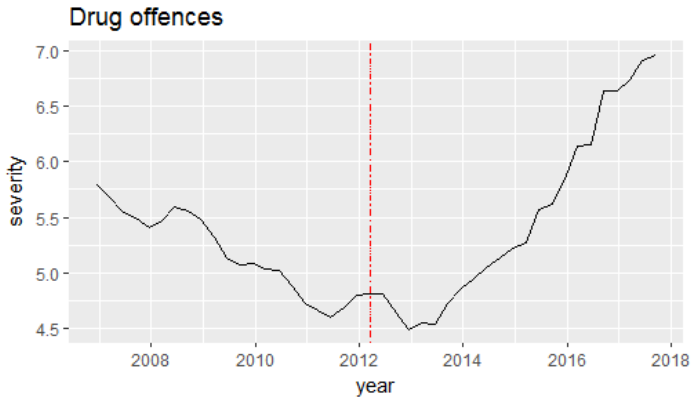
Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

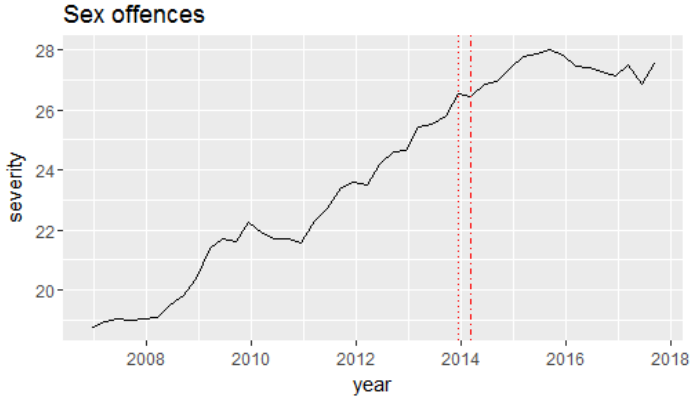
Monitoring Severity

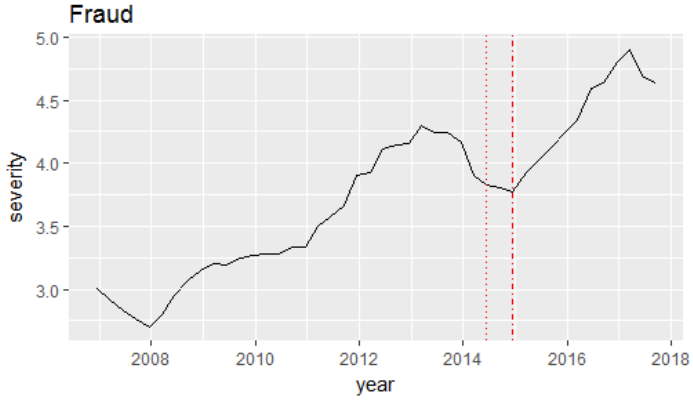


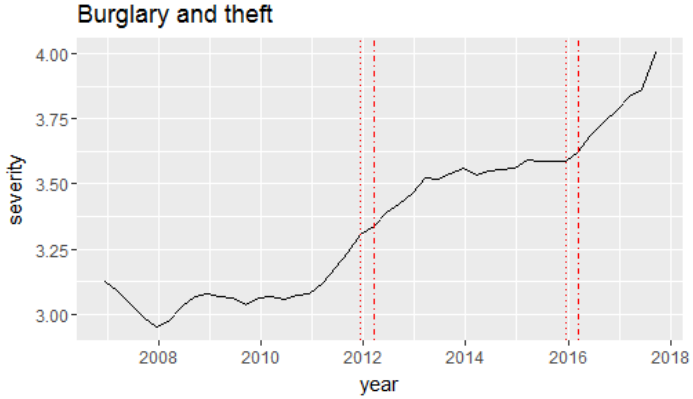
Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

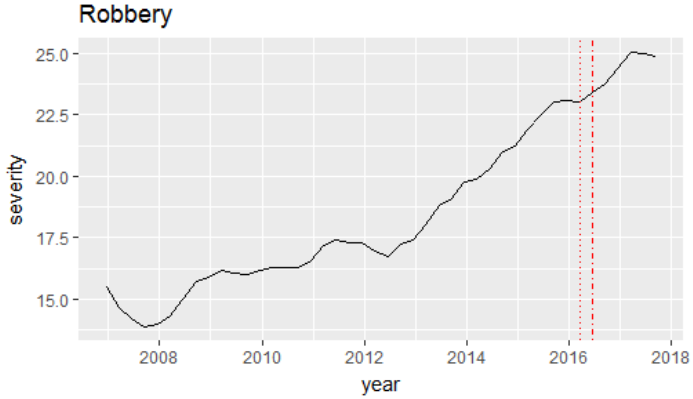




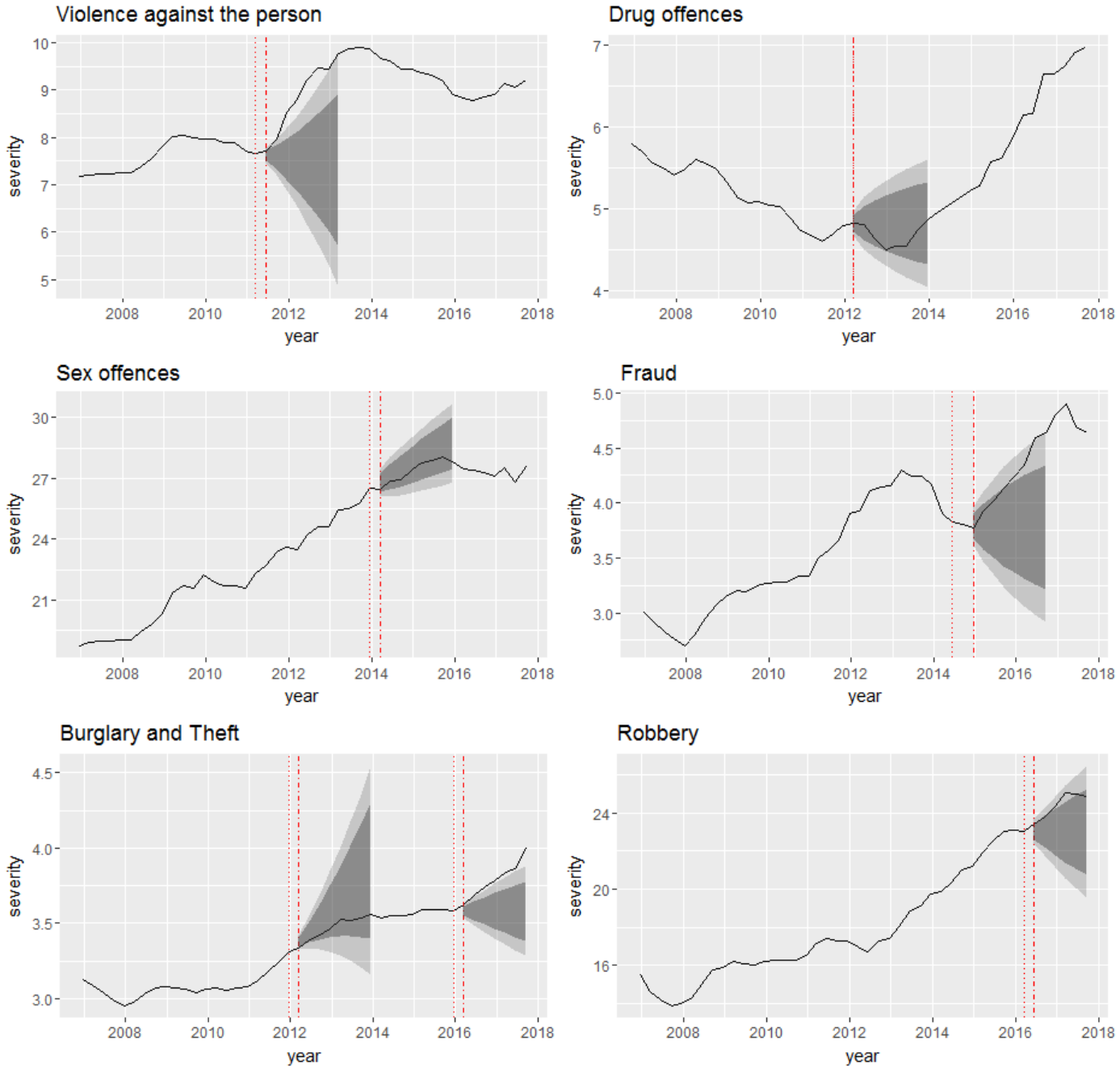








Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion



Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Modelling Severity

- A sample of 7240 offences of theft
- Sentenced at the Crown Court in 2011
- 63.8% received a custodial sentence
  - 151 conditional discharges
  - 74 fines
  - 989 community orders
  - 1806 suspended sentences
  - 4220 custodial sentences

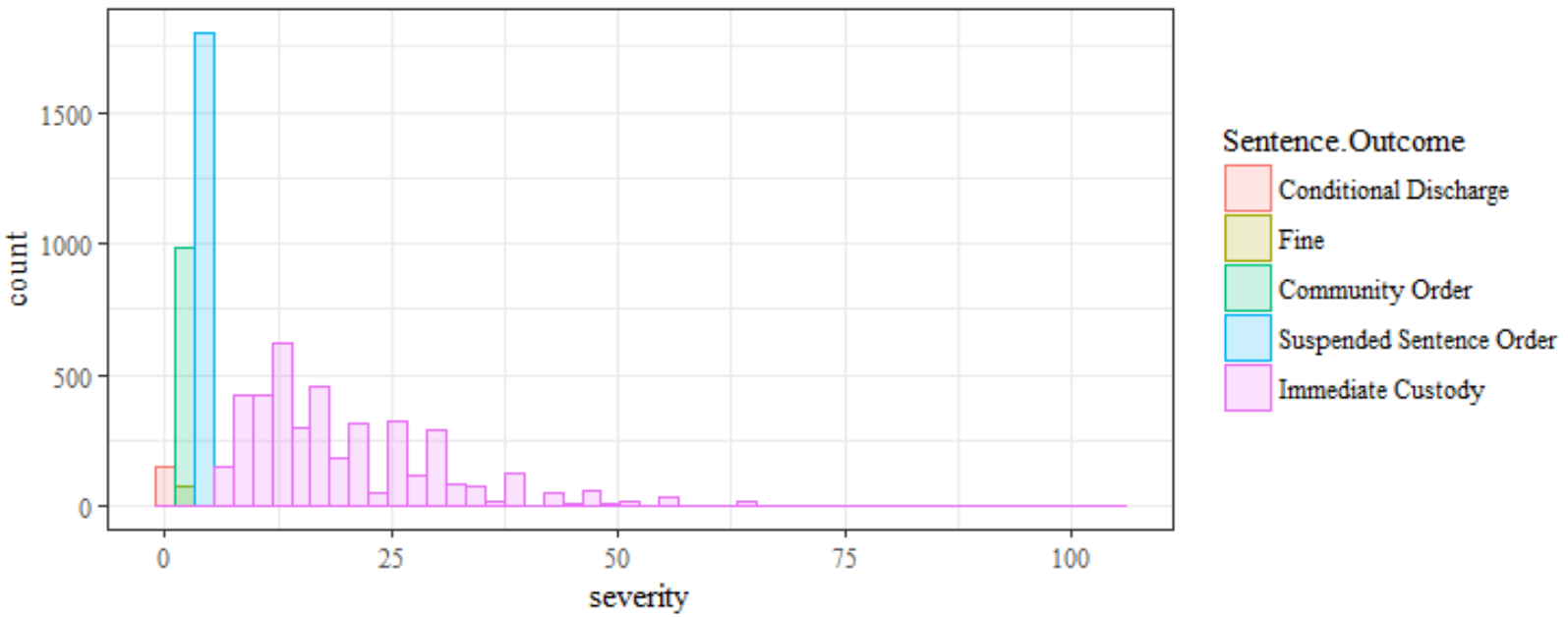
Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Table: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Max
severity	7,242	13.116	12.363	0.95	105.84
age	7,242	32.423	11.024	18	83
male	7,242	0.852	0.355	0	1
pc1_3	7,242	0.252	0.434	0	1
pc4_9	7,242	0.164	0.370	0	1
pc10plus	7,242	0.170	0.375	0	1
plea	7,242	0.847	0.360	0	1
PO_aggburgdwell	7,242	0.004	0.063	0	1
PO_aggburgunspec	7,242	0.006	0.076	0	1
PO_atttheft	7,242	0.005	0.072	0	1
PO_commercialburg	7,242	0.079	0.269	0	1
PO_conspburg	7,242	0.003	0.057	0	1
PO_conspfraud	7,242	0.007	0.084	0	1
PO_conspother	7,242	0.002	0.048	0	1
PO_conspsteal	7,242	0.008	0.088	0	1
PO_dishonestrep	7,242	0.066	0.248	0	1
PO_equipped	7,242	0.007	0.085	0	1
PO_handling	7,242	0.011	0.106	0	1
PO_immigration	7,242	0.004	0.066	0	1
PO_laundering	7,242	0.016	0.124	0	1
PO_otherfraud	7,242	0.140	0.347	0	1
PO_othertheft	7,242	0.040	0.196	0	1
PO_receivinggoods	7,242	0.066	0.248	0	1
PO_theftperson	7,242	0.048	0.215	0	1
PO_theftshop	7,242	0.061	0.239	0	1
PO_thefttrust	7,242	0.062	0.242	0	1
PO_theftvehicle	7,242	0.005	0.071	0	1
PO_falsepassport	7,242	0.035	0.184	0	1

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Modelling Severity



Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

	Dependent variable:	
	log(severity)	
	Model 1 - custody	Model 2 - all sentences
age of the defendant	0.006***	
	(0.001)	
guilty plea entered	−0.130***	
	(0.019)	
male defendant	0.052*	
	(0.024)	
1 to 3 prev convictions	0.093***	
	(0.020)	
4 to 9 prev convictions	0.184***	
	(0.022)	
10+ prev convictions	0.194***	
	(0.022)	
constant	2.836***	
	(0.040)	
Observations	4,220	
R^2	0.331	
Note:		*p<0.1; **p<0.05; ***p<0.01

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

	Dependent variable:	
	log(severity)	
	Model 1 - custody	Model 2 - all sentences
age of the defendant	0.006***	0.005***
	(0.001)	(0.001)
guilty plea entered	−0.130***	−0.103***
	(0.019)	(0.028)
male defendant	0.052*	0.181***
	(0.024)	(0.030)
1 to 3 prev convictions	0.093***	0.464***
	(0.020)	(0.027)
4 to 9 prev convictions	0.184***	0.714***
	(0.022)	(0.032)
10+ prev convictions	0.194***	0.814***
	(0.022)	(0.032)
constant	2.836***	1.905***
	(0.040)	(0.053)
Observations	4,220	7,242
R^2	0.331	0.318
Note:		*p<0.1; **p<0.05; ***p<0.01

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

	Dependent variable:	
	log(severity)	
	Model 1 - custody	Model 2 - all sentences
age of the defendant	0.006***	0.005***
	(0.001)	(0.001)
guilty plea entered	−0.130***	−0.103***
	(0.019)	(0.028)
male defendant	0.052*	0.181***
	(0.024)	(0.030)
1 to 3 prev convictions	0.093***	0.464***
	(0.020)	(0.027)
4 to 9 prev convictions	0.184***	0.714***
	(0.022)	(0.032)
10+ prev convictions	0.194***	0.814***
	(0.022)	(0.032)
constant	2.836***	1.905***
	(0.040)	(0.053)
Observations	4,220	7,242
R^2	0.331	0.318
Note:		*p<0.1; **p<0.05; ***p<0.01

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

	Dependent variable:	
	log(severity)	
	Model 1 - custody	Model 2 - all sentences
age of the defendant	0.006***	0.005***
	(0.001)	(0.001)
guilty plea entered	−0.130***	−0.103***
	(0.019)	(0.028)
male defendant	0.052*	0.181***
	(0.024)	(0.030)
1 to 3 prev convictions	0.093***	0.464***
	(0.020)	(0.027)
4 to 9 prev convictions	0.184***	0.714***
	(0.022)	(0.032)
10+ prev convictions	0.194***	0.814***
	(0.022)	(0.032)
constant	2.836***	1.905***
	(0.040)	(0.053)
Observations	4,220	7,242
R^2	0.331	0.318
Note:		*p<0.1; **p<0.05; ***p<0.01

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Propagating Uncertainty

- We are using severity scores as data but they are estimates
  - There is sampling error from having only 21 magistrates
  - We are uncertain about each paired comparison, p_{rc}
  - We estimate the Thurstone model simultaneously with the sentencing model of interest
  - Using Bayesian statistics to propagate that uncertainty through to our severity scale and the final model
  - We take each p_{rc} as a parameter to be estimated using data (likelihood function from the magistrates' responses) and prior distributions (uninformative beta distributions)

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Propagating Uncertainty

- We are using severity scores as data but they are estimates
  - There is sampling error from having only 21 magistrates
  - We are uncertain about each paired comparison, p_{rc}
  - We estimate the Thurstone model simultaneously with the sentencing model of interest
  - Using Bayesian statistics to propagate that uncertainty through to our severity scale and the final model
  - We take each p_{rc} as a parameter to be estimated using data (likelihood function from the magistrates' responses) and prior distributions (uninformative beta distributions)
- Conditional discharges, fines, and community orders are heterogeneous disposal types
  - This is akin to a problem of Berkson measurement error
  - $Y = Y^* + V$
  - i.e. the true severity scores are more variable than our estimated severity scores
  - Rather than using the severity scores as point estimates we take the entire latent severity variable, $N \sim (Y^*, \sqrt{0.5})$

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

	Dependent variable:	
	log(severity)	
	Model 2 - all sentences	Model 3 - plus uncertainty
age of the defendant	0.005	
	(0.001)	
guilty plea entered	−0.103	
	(0.028)	
male defendant	0.181	
	(0.030)	
1 to 3 prev convictions	0.464	
	(0.027)	
4 to 9 prev convictions	0.714	
	(0.032)	
10+ prev convictions	0.814	
	(0.032)	
constant	1.905	
	(0.053)	
Observations	7,242	

	Dependent variable:	
	log(severity)	
	Model 2 - all sentences	Model 3 - plus uncertainty
age of the defendant	0.005	0.005
	(0.001)	(0.001)
guilty plea entered	−0.103	−0.101
	(0.028)	(0.031)
male defendant	0.181	0.188
	(0.030)	(0.034)
1 to 3 prev convictions	0.464	0.483
	(0.027)	(0.033)
4 to 9 prev convictions	0.714	0.740
	(0.032)	(0.040)
10+ prev convictions	0.814	0.845
	(0.032)	(0.042)
constant	1.905	1.881
	(0.053)	(0.066)
Observations	7,242	7,242

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Latest Developments

- Diminishing returns of severity for every additional day in prison
  - When sentence length increases by $X4$ severity increases by $X4X.9$
  - Identify milestones in severity serving a prison?

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Latest Developments

- Diminishing returns of severity for every additional day in prison
  - When sentence length increases by $X4$ severity increases by $X4X.9$
  - Identify milestones in severity serving a prison?
- Replicated the matrix of severity under an expert elicitation format
  - This time including new categories for fines and community orders
  - Became obvious that self-completed questionnaires were not the best format

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Expert Knowledge Elicitation

- Workshop conducted last week at the Maths Society
  - 1 defence lawyer, 1 magistrate, 2 criminal law experts from the Sentencing Council, 1 criminal lawyer, and 1 penal theorist
  - A similar questionnaire under a more discursive format
  - Aimed to discuss openly the questions and reach consensus
  - It took us almost the whole day, from 11:00 to 16:00

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Expert Knowledge Elicitation

- Workshop conducted last week at the Maths Society
  - 1 defence lawyer, 1 magistrate, 2 criminal law experts from the Sentencing Council, 1 criminal lawyer, and 1 penal theorist
  - A similar questionnaire under a more discursive format
  - Aimed to discuss openly the questions and reach consensus
  - It took us almost the whole day, from 11:00 to 16:00
- Clear improvements in the validity of the responses we obtained
- Problems associated with questionnaires:
  - Too shallow, prevents dialogue, different interpretations of severity

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Expert Knowledge Elicitation

- Workshop conducted last week at the Maths Society
  - 1 defence lawyer, 1 magistrate, 2 criminal law experts from the Sentencing Council, 1 criminal lawyer, and 1 penal theorist
  - A similar questionnaire under a more discursive format
  - Aimed to discuss openly the questions and reach consensus
  - It took us almost the whole day, from 11:00 to 16:00
- Clear improvements in the validity of the responses we obtained
- Problems associated with questionnaires:
  - Too shallow, prevents dialogue, different interpretations of severity
- Allowed us to explore our underlying assumptions
  - It turns out the assumption of equal variance is not valid
  - A finding with methodological implications beyond the measurement of sentence severity

	absolute discharge	conditional discharge	fine (band A)	fine (band B)	fine (band C)	fine (band D)	fine (band E)	fine (band F)	community order (low)	community order (medium)	community order (high)	6month suspended 1month custody	12month suspended 1month custody	6month suspended 6month custody	24month suspended 12month custody	1month custody	2month custody	3month custody
absolute discharge																		
conditional discharge			0.65	0.8														
fine (band A)																		
fine (band B)																		
fine (band C)									0.65									
fine (band D)									0.55	0.75		0.75						
fine (band E)									0.4	0.6	0.95	0.7	0.75					
fine (band F)									0.2	0.5	0.9	0.6	0.65	0.8		1		
community order (low)																		
community order (medium)												0.55						
community order (high)												0.4	0.45	0.6		0.6	0.7	0.75
6month susp 1month cust																		
12month susp 1month cust														0.95				
6month susp 6month cust																0.85		
24month susp 12month cust																0.3	0.35	0.4
1month custody																		
2month custody																		
3month custody																		

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Future Applications

- Estimate a similar index in different jurisdictions
  - Scotland?

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Future Applications

- Estimate a similar index in different jurisdictions
  - Scotland?
- Monitoring proportionality
  - Ordinal proportionality: Plotting the average sentence severity imposed to different crimes
  - Cardinal proportionality: Plotting the average severity for all sentences imposed in E&W and the average seriousness of all crimes processed

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Conclusion

- Selection bias is an extremely pervasive problem in sentence data analyses

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Conclusion

- Selection bias is an extremely pervasive problem in sentence data analyses
- Solutions suggested in the literature are based on questionable assumptions and waste information

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Conclusion

- Selection bias is an extremely pervasive problem in sentence data analyses
- Solutions suggested in the literature are based on questionable assumptions and waste information
- The estimation of a scale of severity allows us to overcome both problems
  - e.g.1 most guidelines have not increased severity
  - e.g.2 male defendants are more harshly treated than we knew

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Conclusion

- Selection bias is an extremely pervasive problem in sentence data analyses
- Solutions suggested in the literature are based on questionable assumptions and waste information
- The estimation of a scale of severity allows us to overcome both problems
  - e.g.1 most guidelines have not increased severity
  - e.g.2 male defendants are more harshly treated than we knew
- Under a Bayesian framework we can treat severity scores as estimates
  - In so doing account for sampling error and measurement error

Introduction
Measuring Severity
Sensitivity Analyses
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Next Steps
Conclusion

Conclusion

- Selection bias is an extremely pervasive problem in sentence data analyses
- Solutions suggested in the literature are based on questionable assumptions and waste information
- The estimation of a scale of severity allows us to overcome both problems
  - e.g.1 most guidelines have not increased severity
  - e.g.2 male defendants are more harshly treated than we knew
- Under a Bayesian framework we can treat severity scores as estimates
  - In so doing account for sampling error and measurement error
- Measuring sentence severity remains a methodological challenge
  - A latent/unobservable and highly subjective concept
  - We need new scales based on different methodologies
  - We need to be more open about the assumptions underlying different scales
  - Ultimately deciding which scale to use will be a subjective choice