Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

RSS 2018 INTERNATIONAL CONFERENCE

Tackling Selection Bias in Sentence Data Analysis Using a Severity Scale

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

Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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)
  - Assumes that sentencing is undertaken in two steps
  - Requires variables that meet the exclusion criteria

Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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)
  - Assumes that sentencing is undertaken in two steps
  - Requires variables that meet the exclusion criteria
- Models for censored data (Tobit model)
  - Assumes that sentencing is a one-step decision process
  - Assumes that non-custodial sentences are part of the same distribution (normal) as custodial durations

Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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

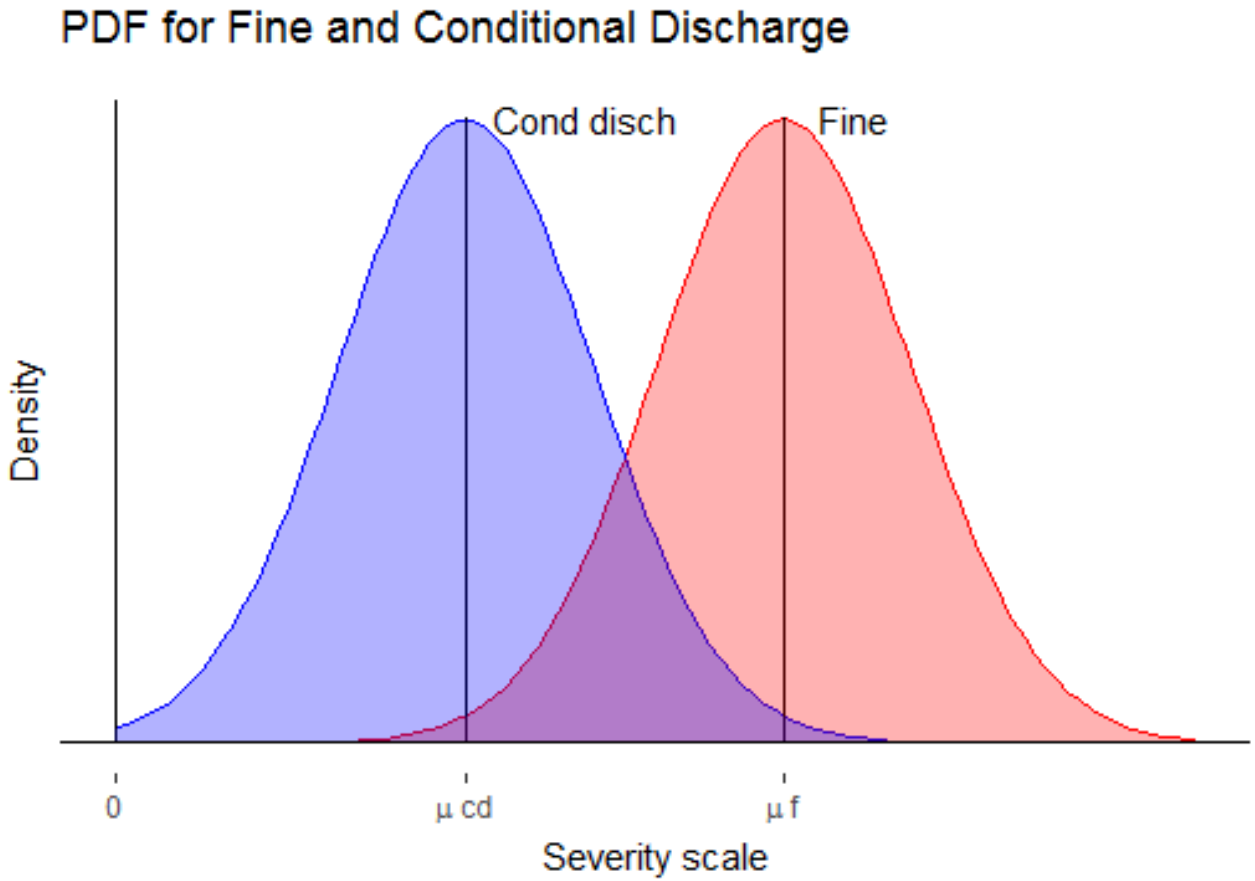
Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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)

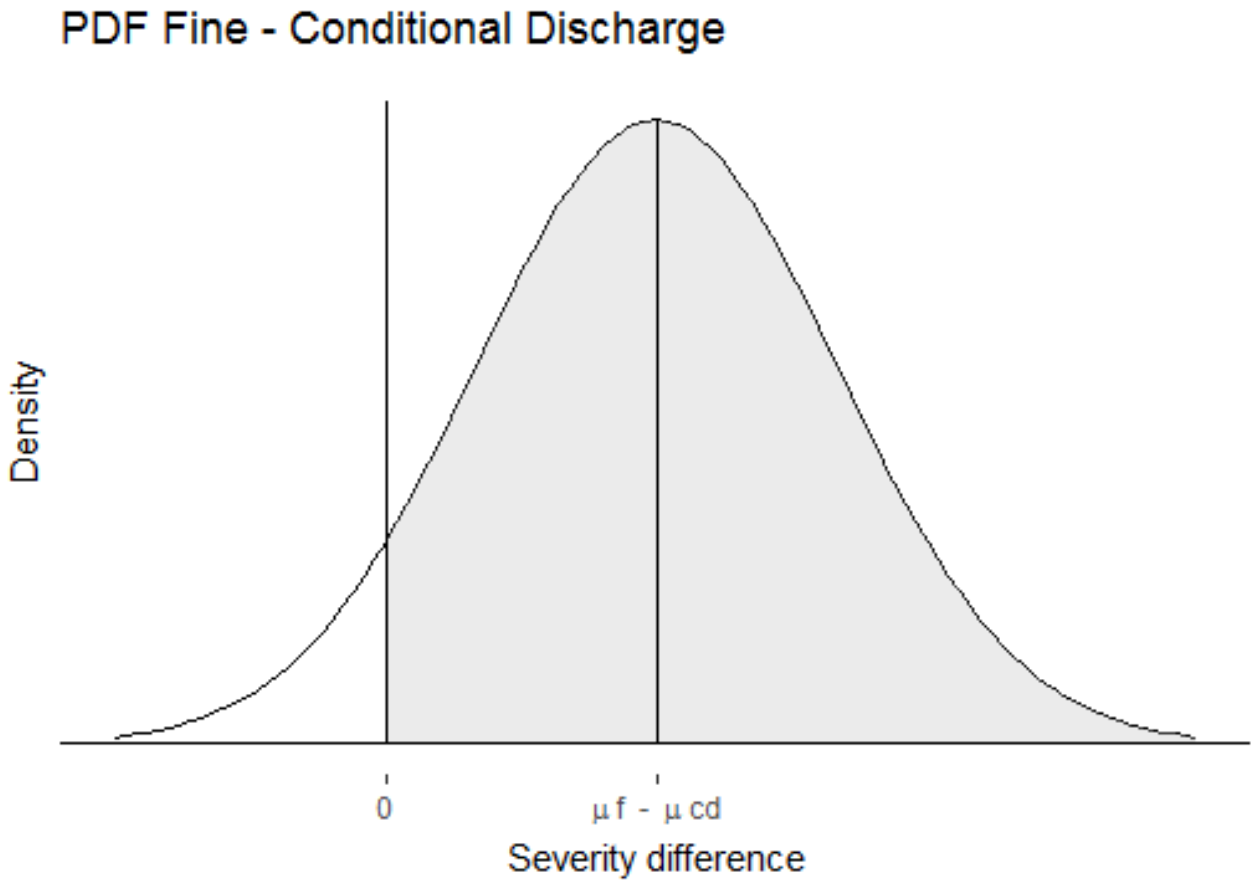
Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

Thurstone Model: Visually



Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

Thurstone Model: Visually



Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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	

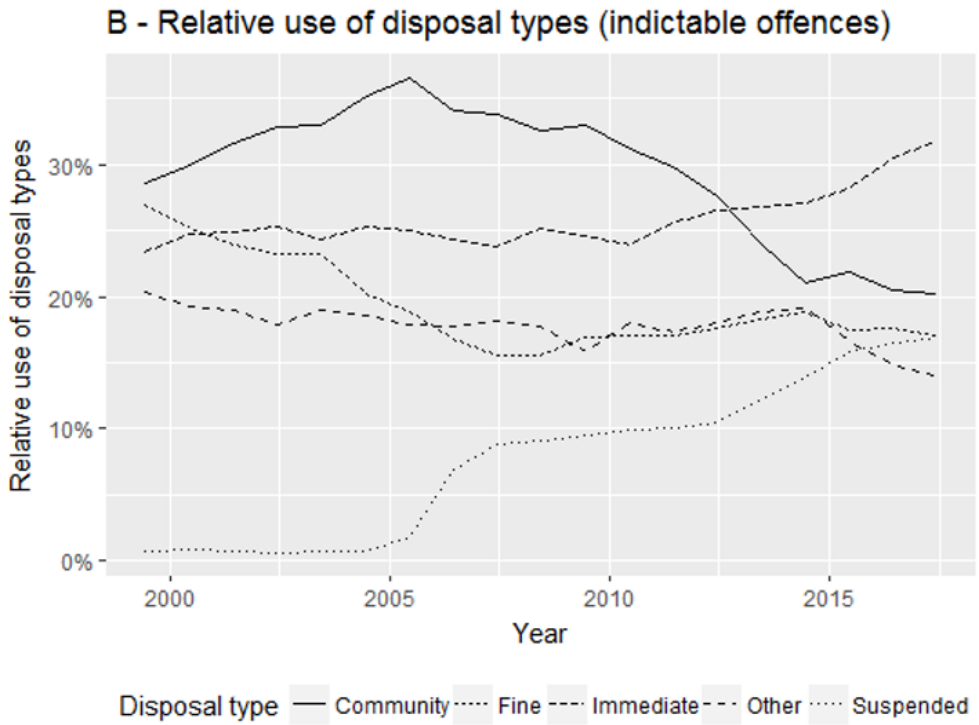
Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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

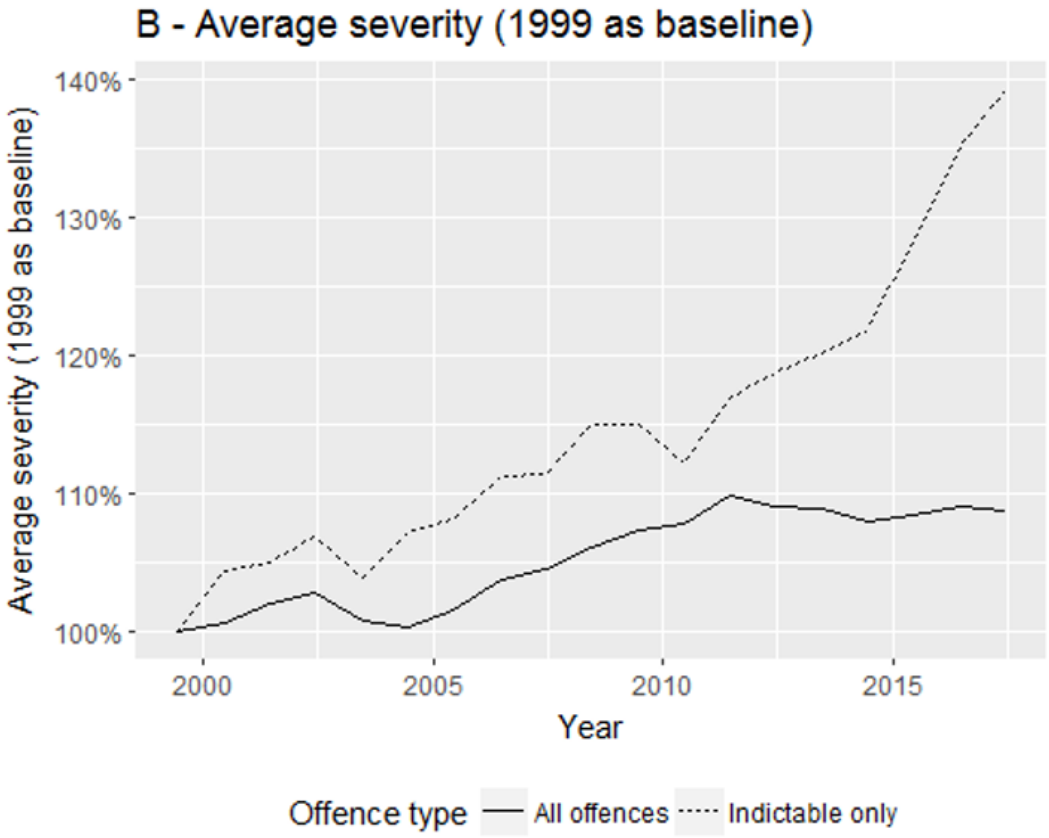
Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

Monitoring Severity

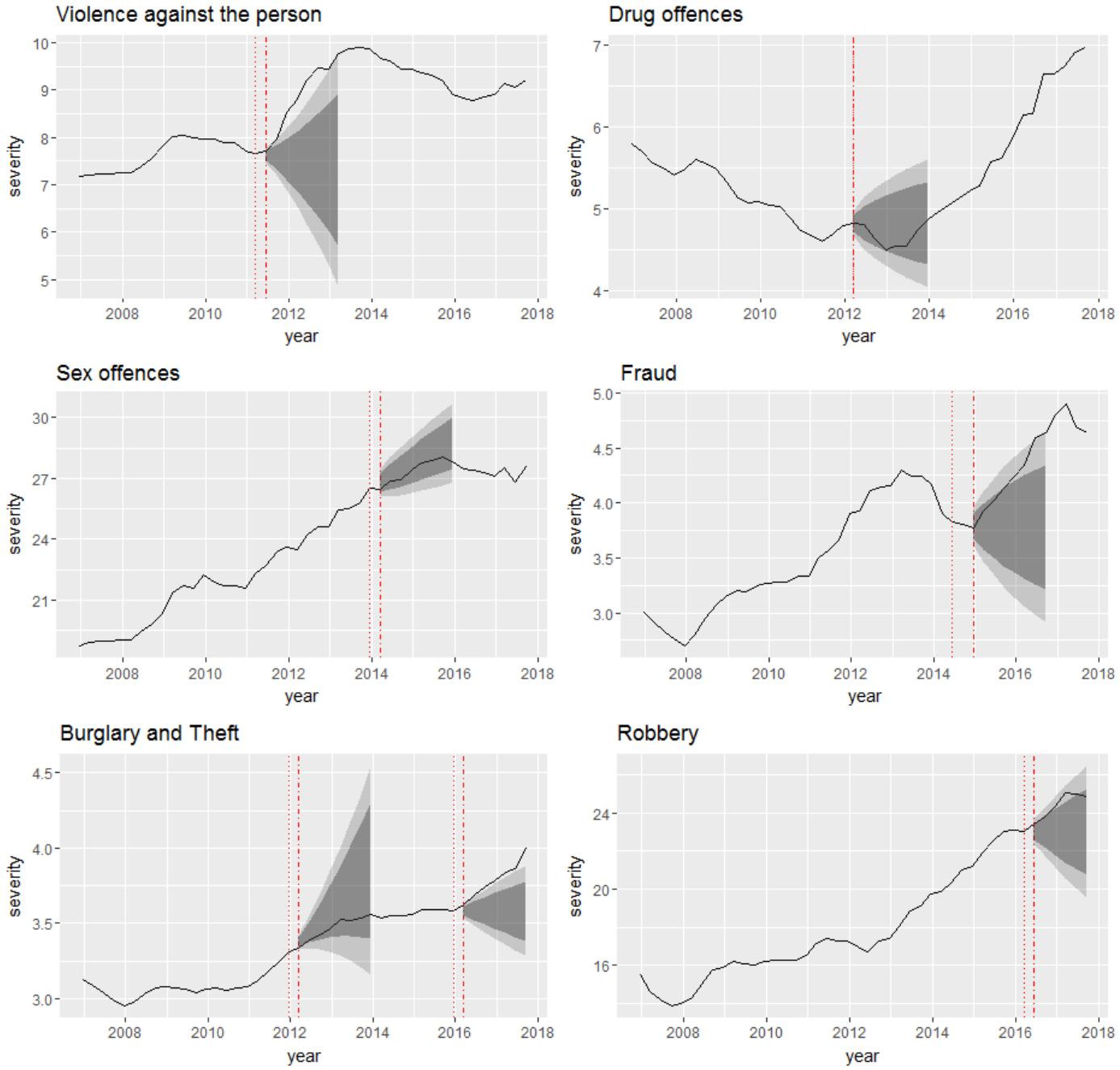


Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

Monitoring Severity



Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion



Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

Modelling Severity

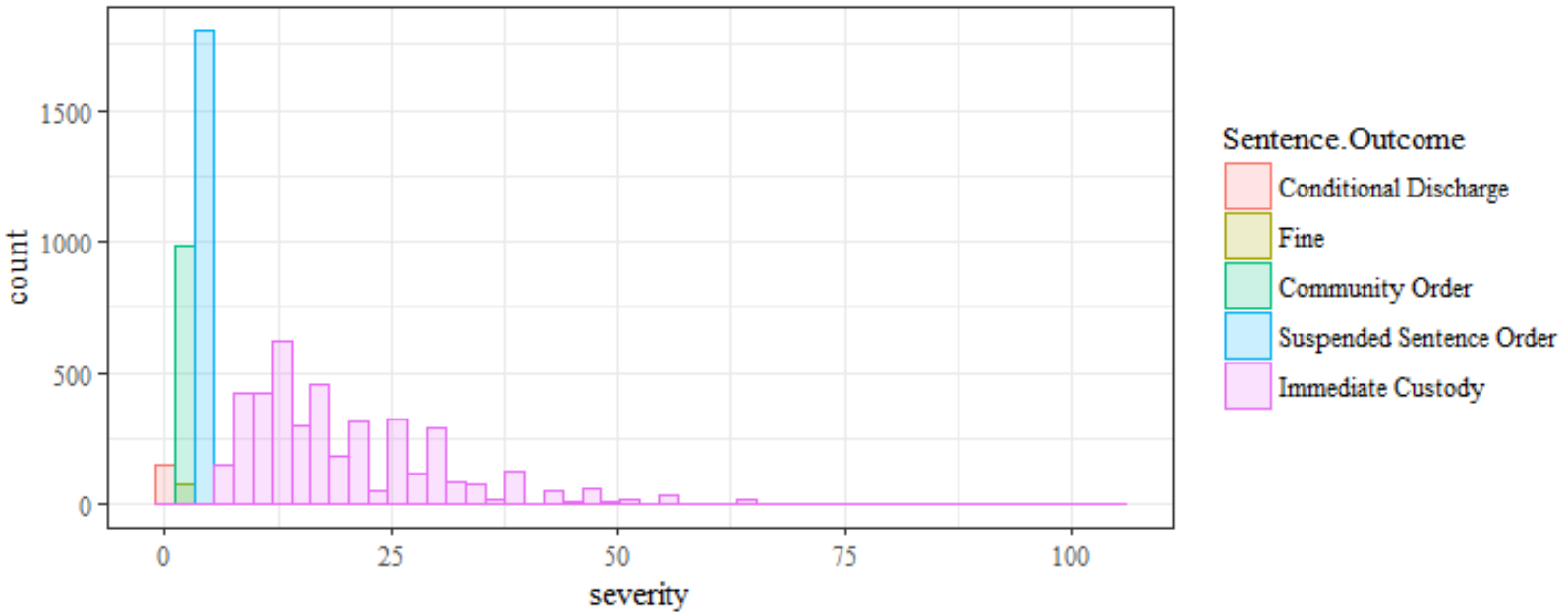
- A sample of 7,242 offences of theft
- Sentenced at the Crown Court in 2011
- 63.8% received a custodial sentence
  - 151 conditional discharges
  - 74 fines
  - 989 community orders
  - 1806 suspended sentences
  - 4220 custodial sentences

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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

Modelling Severity



Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

	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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

	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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

	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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

	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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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 use diffuse beta priors to estimate their posterior distributions
  - We take a sample from our posterior distributions over p_{rc} and use them to estimate a set of severity scores and the final model
  - and repeat this process 2,000 times to propagate that uncertainty through to our severity scale and the final model

Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

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 use diffuse beta priors to estimate their posterior distributions
  - We take a sample from our posterior distributions over p_{rc} and use them to estimate a set of severity scores and the final model
  - and repeat this process 2,000 times to propagate that uncertainty through to our severity scale and the final model
- 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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

	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	

Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

	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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

Discussion

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

Introduction
Measuring Severity
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

Discussion

- 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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

Discussion

- 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 scale of severity allows us to work on 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
Monitoring Severity
Modelling Severity
Propagating Uncertainty
Discussion

Discussion

- 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 scale of severity allows us to work on both problems
  - e.g.1 most guidelines have not increased severity
  - e.g.2 male defendants are more harshly treated than we knew
- Further work is necessary in order to:
  - refine the scale of severity
  - figure out the best way to propagate the uncertainty accurately and simply
  - export this approach to other jurisdictions