NCRM Research Showcase

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

Jose Pina-Sánchez
John Paul Gosling
Sara Geneletti

Sentencing Studies

- The concept of *punishment* is central to Criminology and Criminal Justice
- Most clearly manifested in the *sentencing* practice
- The analysis of sentence data allows exploring lots of important questions
  - E.g. identify the case characteristics considered by judges
  - Detect unwarranted disparities (discrimination)
  - Study the concepts of *deterrence*, *recidivism*, *penal populism*, etc.

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

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

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 monumental loss of information

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

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
- Two stage processes (Heckman selection model)
  - Assumes that sentencing is undertaken in two steps
  - Requires variables that meet the exclusion criteria

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
- 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

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
- 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
- Keep treating non-custodial cases as a homogeneous group

Our Scale of Severity

- We suggest an alternative approach based on the estimation of a scale of severity
  - So we can analyse 100% of the offences
  - while making the most of the information available

Our Scale of Severity

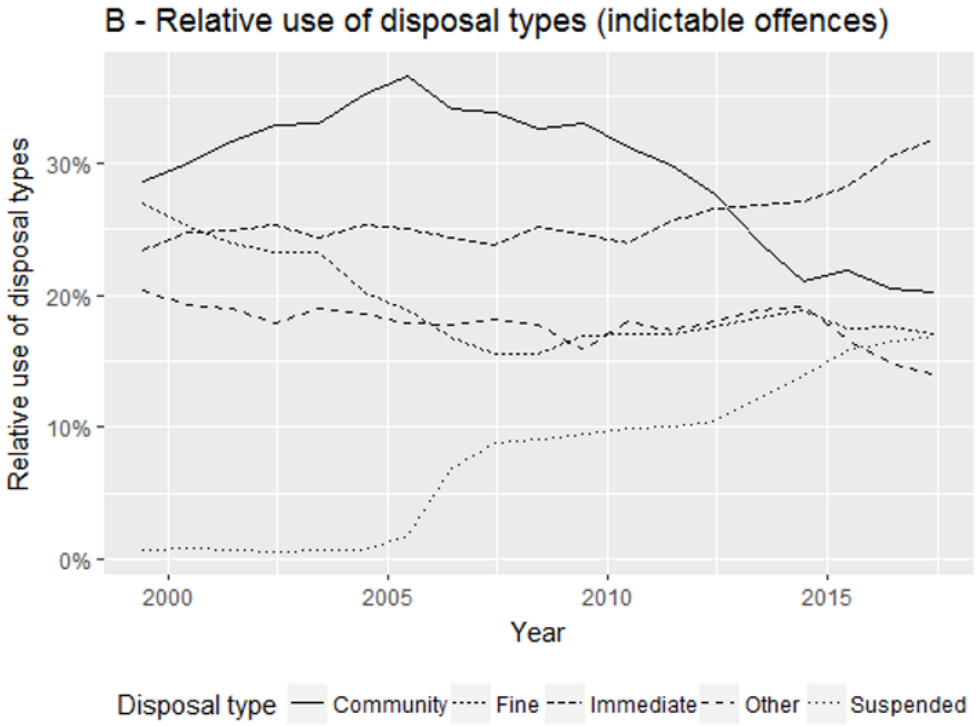
- We suggest an alternative approach based on the estimation of a scale of severity
  - So we can analyse 100% of the offences
  - while making the most of the information available
- We used...
  - The 'sentencing ladder'
  - A sample of 21 magistrates
  - Pairwise comparisons (Thurstone method)

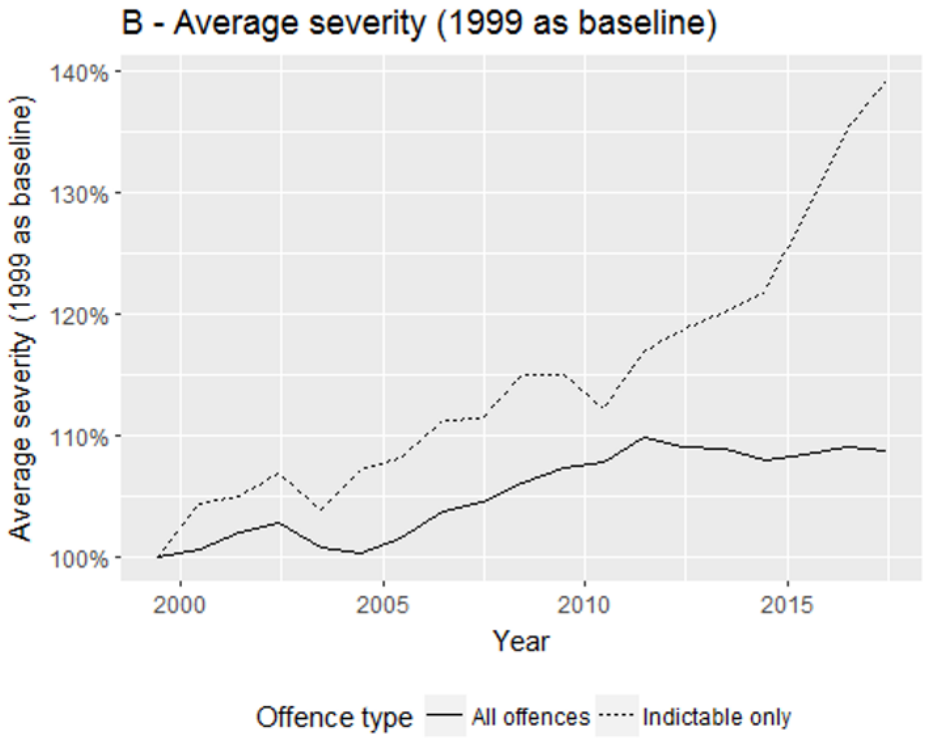
Severity Scores

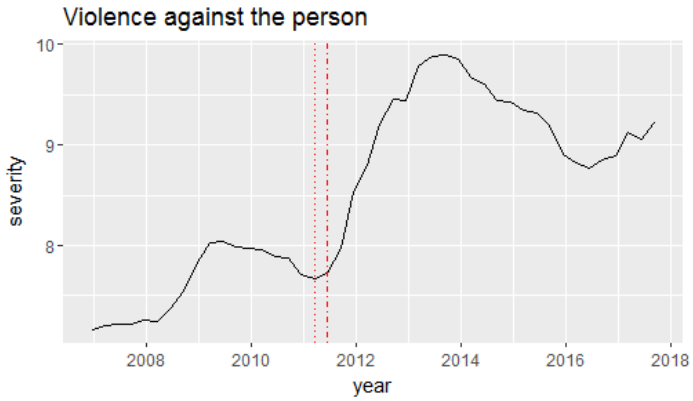
Sentence outcome	Severity score
absolute discharge	0
conditional discharge	0.97
fine	1.33
community order	2.13
1-month custody 6-months suspended	2.34
1-month custody 12-months suspended	3.66
6-months custody 6-months suspended	3.78
12-months custody 24-months suspended	5.74
1-month custody	5.05
2-months custody	5.75
3-months custody	6.45
12-months custody	13.45
5-years custody	47.05
20-years custody	173.05

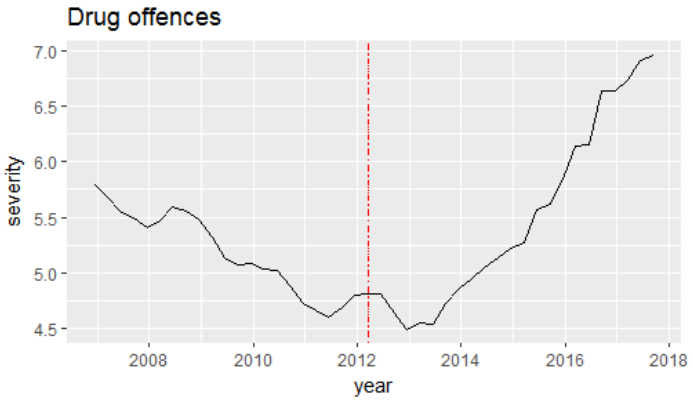
Have the Guidelines Increased Severity?

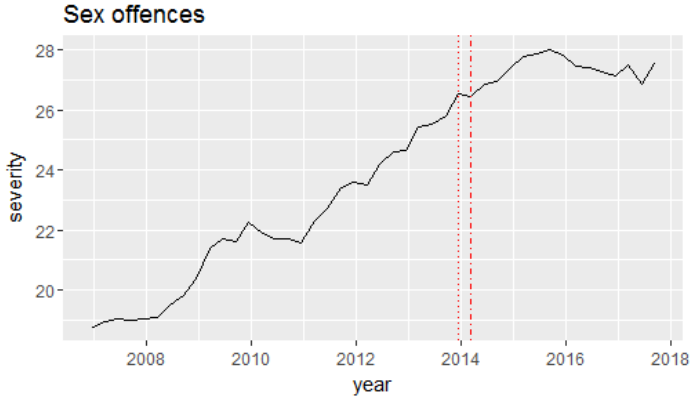
- We explore the increase in sentence severity in E&W
- Test whether the new sentencing guidelines are to be blamed

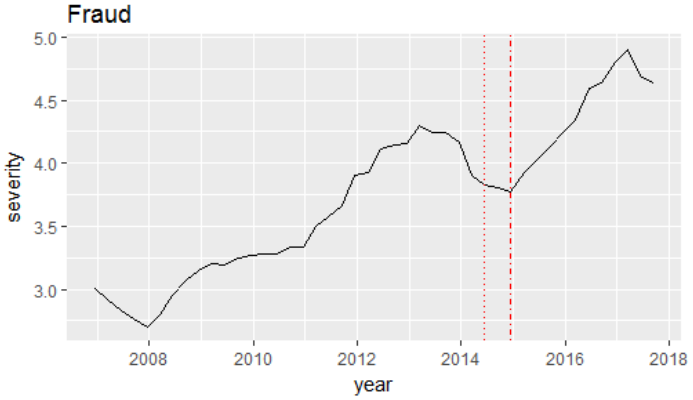


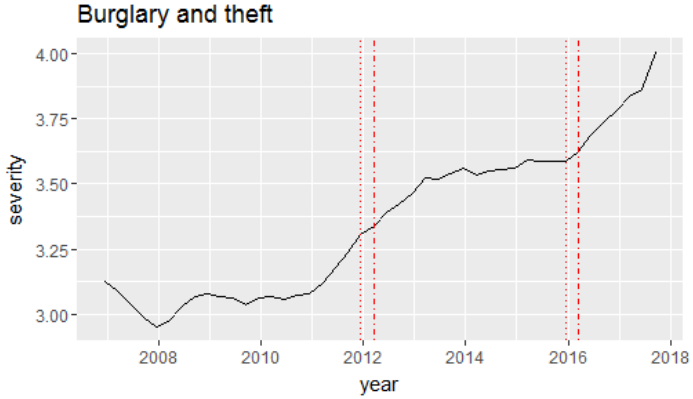


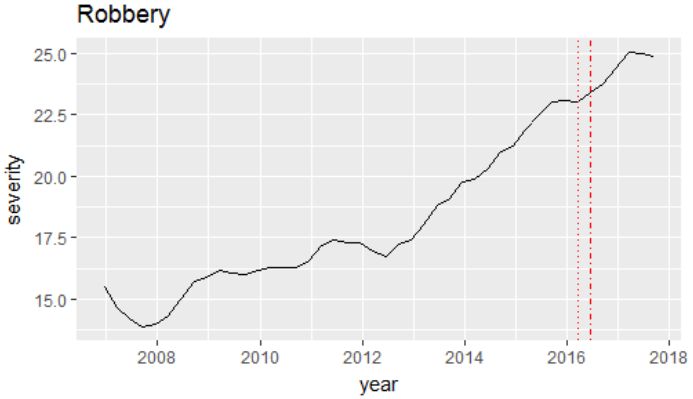


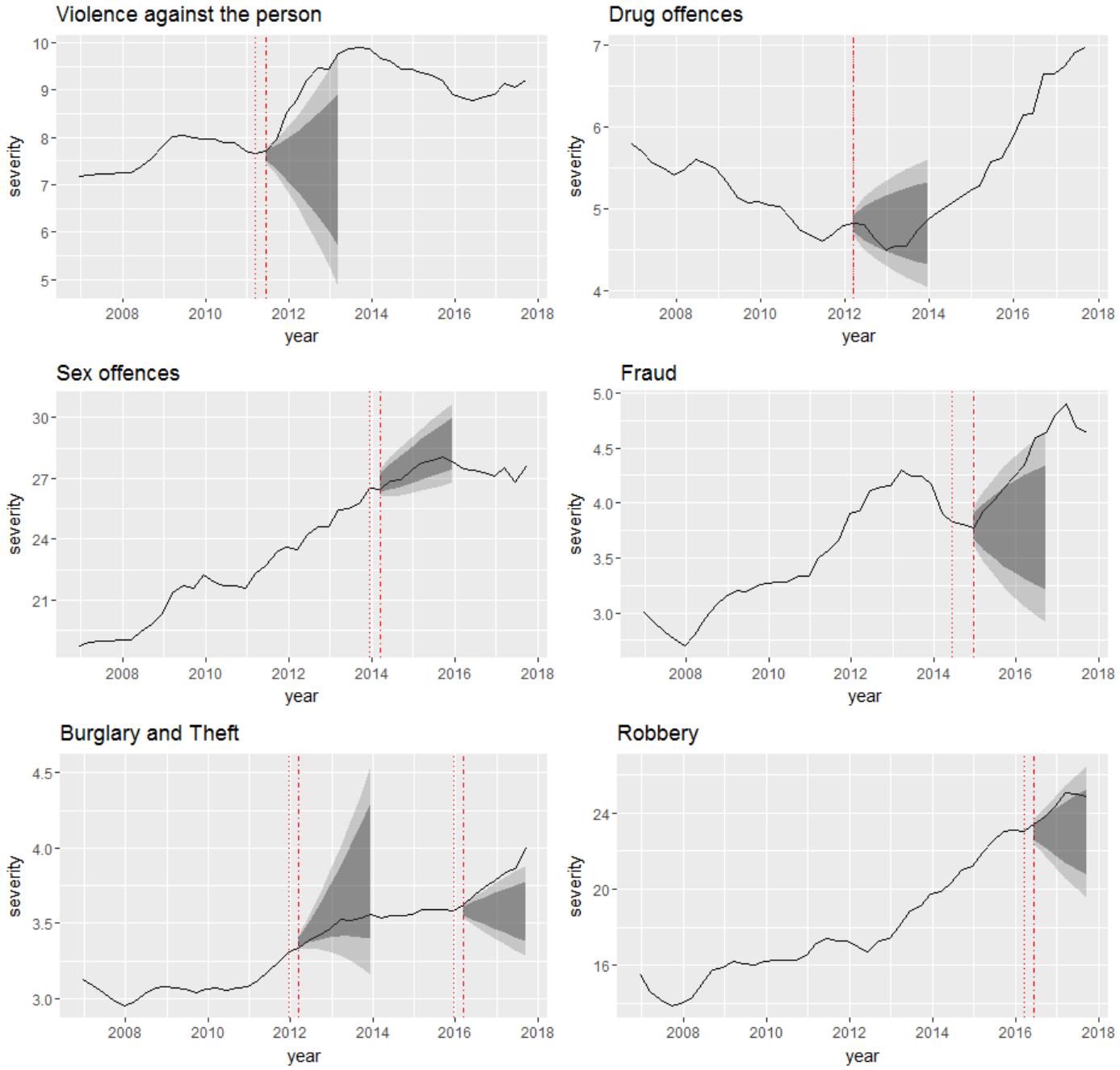












Exploring Selection Bias

- To explore the impact of selection bias we compare different models
  - we use 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)
  - model the logs of their severity scores on various offence and offender characteristics
- We use Bayesian statistics to account for...
  - the sampling error associated with the estimation of sentence severity
  - the measurement error stemming from the unobserved heterogeneity within fines and community orders

	Dependent variable: log(severity)		
	Model 1 - custody	Model 2 - all sentences	Model 3 - Tobit
age of defendant	0.006		
	(0.001)		
male defendant	0.054		
	(0.026)		
guilty plea entered	−0.132		
	(0.020)		
1 to 3 prev conv.	0.095		
	(0.021)		
4 to 9 prev conv.	0.187		
	(0.024)		
10+ prev conv.	0.199		
	0.024		
constant	2.786		
	(0.066)		
N	4,220	7,242	7,242

Note: standard deviations (within parentheses)

	Dependent variable: log(severity)		
	Model 1 - custody	Model 2 - all sentences	Model 3 - Tobit
age of defendant	0.006	0.005	
	(0.001)	(0.001)	
male defendant	0.054	0.185	
	(0.026)	(0.036)	
guilty plea entered	−0.132	−0.104	
	(0.020)	(0.032)	
1 to 3 prev conv.	0.095	0.477	
	(0.021)	(0.034)	
4 to 9 prev conv.	0.187	0.732	
	(0.024)	(0.041)	
10+ prev conv.	0.199	0.835	
	0.024	0.042	
constant	2.786	1.838	
	(0.066)	(0.068)	
N	4,220	7,242	7,242

Note: standard deviations (within parentheses)

	Dependent variable: log(severity)		
	Model 1 - custody	Model 2 - all sentences	Model 3 - Tobit
age of defendant	0.006	0.005	
	(0.001)	(0.001)	
male defendant	0.054	0.185	
	(0.026)	(0.036)	
guilty plea entered	−0.132	−0.104	
	(0.020)	(0.032)	
1 to 3 prev conv.	0.095	0.477	
	(0.021)	(0.034)	
4 to 9 prev conv.	0.187	0.732	
	(0.024)	(0.041)	
10+ prev conv.	0.199	0.835	
	0.024	0.042	
constant	2.786	1.838	
	(0.066)	(0.068)	
N	4,220	7,242	7,242

Note: standard deviations (within parentheses)

	Dependent variable: log(severity)		
	Model 1 - custody	Model 2 - all sentences	Model 3 - Tobit
age of defendant	0.006	0.005	0.006
	(0.001)	(0.001)	(0.001)
male defendant	0.054	0.185	0.275
	(0.026)	(0.036)	(0.041)
guilty plea entered	−0.132	−0.104	−0.152
	(0.020)	(0.032)	(0.035)
1 to 3 prev conv.	0.095	0.477	0.600
	(0.021)	(0.034)	(0.035)
4 to 9 prev conv.	0.187	0.732	0.880
	(0.024)	(0.041)	(0.041)
10+ prev conv.	0.199	0.835	0.990
	0.024	0.042	0.041
constant	2.786	1.838	1.615
	(0.066)	(0.068)	(0.070)
N	4,220	7,242	7,242

Note: standard deviations (within parentheses)

Key Findings

- We provide estimates of severity for non-custodial sentences
  - Which allow us to detect a strong increase in severity in E&W

Key Findings

- We provide estimates of severity for non-custodial sentences
  - Which allow us to detect a strong increase in severity in E&W
- Selection bias in sentence data analysis is truly pervasive
  - Regression coefficients biased in unpredictable magnitude and direction
  - 'Adjustments' used in the literature can make things worse

Key Findings

- We provide estimates of severity for non-custodial sentences
  - Which allow us to detect a strong increase in severity in E&W
- Selection bias in sentence data analysis is truly pervasive
  - Regression coefficients biased in unpredictable magnitude and direction
  - 'Adjustments' used in the literature can make things worse
- We suggest using a scale of severity and Bayesian statistics to...
  - Eliminate the problem of selection bias, make the most of the information available
  - Propagate the uncertainty associated with the estimation of our scale of severity
  - Adjust for other problems such as measurement error in our data

Impact

- Academic impact
  - Pina-Sánchez et al. (2019) *British Journal of Criminology*
  - Pina-Sánchez et al. (under review) *Social Science Research*
  - Pina-Sánchez and Gosling (under review) *Royal Statistical Society: Series A*
  - Disseminated through multiple events (organised, invited, and uninvited)

Impact

- Academic impact
  - Pina-Sánchez et al. (2019) *British Journal of Criminology*
  - Pina-Sánchez et al. (under review) *Social Science Research*
  - Pina-Sánchez and Gosling (under review) *Royal Statistical Society: Series A*
  - Disseminated through multiple events (organised, invited, and uninvited)
- Non-academic impact
  - Pina-Sánchez et al. (2019) *Current Sentencing Practice News*
  - Workshops organised with Sentencing Council and MoJ analysts
  - The Council has adopted our scale of severity to assess the impact of its guidelines

Next Steps

- Improve the estimation of sentence severity
  - Exploring different types of fines and community orders
  - Relaxing some of the assumptions invoked by the Thurstone model
- Bayesian statistics to detect more accurately sentencing discrimination
  - Lots of unobserved legal factors leading to confounding effects
  - Use the literature to inform strong priors on how these factors are distributed