Using Text Mining Techniques to Explore the Role of Judge-Court Interactions on Sentencing Disparities

Jose Pina-Sánchez, Diana Grech, Ian Brunton-Smith, and Dimitrios Sferopoulos

Background: Quantitative studies on sentencing have been constrained by the type of data accessible to researchers. Two main options are generally available - requesting access to courts to collect primary data using observations, file reviews, interviews, and questionnaires; or relying on official secondary data supplied by organisations such as Ministries of Justice or Sentencing Commissions. Primary data collection is both expensive and time-consuming, with studies typically based on relatively small samples restricted to one or a few courts. Official records are limited to the information that the judiciary is willing to disclose, with individual level information on judge characteristics being normally censored. The difficulty to access large samples from more than a few courts capturing judge identifiers has diverted much of the efforts of research on sentencing disparities towards those originating *between* courts, regions, or criminal justice areas, neglecting the disparities taking place *within* them. Important methodological and substantive implications follow. Measures of uncertainty from statistical models missing the clustering of sentences within judges are biased. Substantively, not being able to compare the disparities stemming from the court or the judge level, nor the interaction between them, prevents criminal justice researchers from pinpointing the true extent and origin of these disparities.
Here we use text mining techniques to access an online archive of sentences records processed in the Crown Court, which include the name of the judge who imposed the sentence and the court where the trial took place. Using this new dataset we compare the magnitude of the unwarranted disparities originated at the court and the judge level, and explore the interactions between them taking place when judges rotate across courts.

Methods: We accessed the archive of sentence records uploaded on www.thelawpages.com. Each sentence record was loaded in a browser, a parser written in Perl then looked for specific keywords within the HTML source code, from which relevant variables could be systematically recorded. A number of relevant case characteristics were extracted directly from the source code following this approach (e.g. sentence outcome, type of offence, and age of the defendant). Variables relating to judge characteristics were generated on a second stage. Clauses including Miss, Ms, and Her were extracted from the judges' title in order to approximate their gender. The position of the judge was also approximated to differentiate between Recorders, Queens Counsel judges (identified with the prefix QC) Circuit judges (identified using HHJ, Her Honour and His Honour) and High Court judges (using the clauses Justice and Honourable).
A subsample of 7,221 violent and sexual interpersonal offences sentenced to prison in the Crown Court from 2007 to 2017 was analysed using cross-classified Weibull models, which can account for the right censoring of indeterminate custodial sentences (with the 'minimum term' duration being the last observed point) and the practice of judicial rotation. Certain judges in the Crown Court operate from different courts, which results in a complex hierarchical structure (Figure 1). To account for this phenomenon of cross classification we use three different random effects capturing the unobserved variability stemming from: (i) the court level, (ii) judges who are observed to work from the same court, and (iii) those who rotate across courts. To explore differences between gross and unwarranted disparities we specify two nested models. Model 1 is an empty model, with the random effects capturing disparities that could be due to either legal differences between cases or to inconsistencies in sentencing. Model 2 includes the full list of predictors, with the random part of the model now capturing the unwarranted disparities obtained after controlling for relevant case characteristics.

Results: We find that both gross and unwarranted differences between judges are substantially larger than differences between courts (Table 1), highlighting the importance of correctly accounting for between-judge disparities in sentencing data analyses. Comparing the random effects for the two types of judge a more complex picture is evident. Gross between-judge disparities are larger amongst 'rotating' judges, but that situation is reversed when considering unwarranted disparities. That is, judges who rotate across courts tend to deal with a more heterogeneous workload but in spite of that, tend to sentence more consistently than judges who stay in the same court when features of the case have been taken into consideration.
We do not find evidence of significant differences in sentencing based on the judges gender (Table 2). By contrast, judicial career stage appears to have an important effect on the relative severity of the punishment, with High Court judges sentencing significantly more severely than any other type of judges. This might be expected, given the more serious cases that High Court judges are required to handle.

Conclusion: In the Crown Court it is not so much the location where the case will be processed, but the judge presiding over the case within that location that will have the potential for providing an unduly lenient or harsher sentence. This stands at odds with results from Johnson (2006) who found that judge-level disparities tended to be smaller than those originating between courts at the Pennsylvania County Court. Such discrepancies regarding the main source of unwarranted disparities highlights the importance of correctly recognising the different contexts in which different jurisdictions operate. In particular, it seems that most of the theories that have been put forward by to explain differences in disparities across courts may be more contextually specific than previously suggested, and consequently they do not have ground in the Crown Court of England and Wales.
Judges who rotate across courts seem to sentence more consistently than those who do not. This finding corroborates Hester (2016) hypotheses regarding the beneficial effects associate with judicial rotation in the form of the spread of general sentencing principles and challenging of local procedures. In a time when the jurisdiction of England and Wales has embarked in a process of sentencing reform to promote consistency through the design of guidelines, it might be worth considering complimentary strategies such as fostering judicial rotation amongst non-circuit judges. This strategy is particularly interesting for the promotion of consistency in sentencing since it is less intrusive on judicial autonomy, avoiding some of the negative side-effects that have been associated with sentencing guidelines.
The growing sophistication of text mining techniques has made it possible for us to reconstruct key features of a sample of cases processed through the Crown Court. In particular, we were able to make use of information that is not normally present in official data (such as the name of the judge, or the number of offences that featured in the case). The records available at 'The Law Pages', also capture information on additional features often ignored in official data, such as the name of the barrister that defended the case, as well as some of the remarks made by the judge in the sentence transcript. We encourage other researchers to replicate our text mining strategy to explore other important under-researched questions about the operation of the court system, including: i) the effect of legal representation; ii) the application of the totality principle; or iii) the ways that judges make use of controversial case features such as the offence being committed while the offender was intoxicated.

References

Hester, R. (2016). Judicial rotation as centripetal force: Sentencing in the court communities of South Carolina. *Criminology 55*(1), 205–235.

Johnson, B. D. (2006). The multilevel context of criminal sentencing: Integrating judgeand Countylevel influences. criminology. *Criminology 44*(2), 259–298.

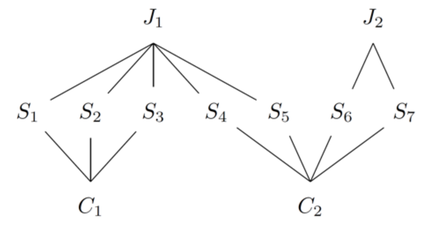

Figure 1: Clustering of Sentences at the Crown Court (an example for seven sentences, S, imposed by two judges, J, operating from two courts, C)

	Model 1		Model 2	
	Estimate	95% CI	Estimate	95% CI
Fixed Effects				
Rotate	5.94	(5.73, 6.15)	12.06	(11.7, 12.43)
Stay	5.35	(5.17, 5.53)	12.05	(11.69, 12.42)
Random				
σ_{court}	0.43	(0.35, 0.53)	0.18	(0.13, 0.24)
$\sigma_{judge.rotate}$	1.10	(1, 1.21)	0.36	(0.30, 0.44)
$\sigma_{judge.stay}$	0.69	(0.61, 0.78)	0.47	(0.39, 0.56)

Response variable: duration of custody in months

Table 1: Random Effects for Model 1 (the empty model) and Model 2 (the model controlling for case characteristics)

	Estimate	95% CI
Fixed Effects		
Judge Characteristics		
Rotating judge	12.06	(11.7, 12.43)
Staying judge	12.05	(11.69, 12.42)
Female judge	0.03	(-0.12, 0.19)
Circuit judge	-0.38	(-0.77, 0.04)
High Court judge	0.62	(0.41, 0.82)
Queen's Counsel judge	0.01	(-0.08, 0.10)
Recorder	-0.30	(-0.41, -0.18)
Defendant Characteristics		
Defendant male	0.41	(0.28, 0.54)
Defendant age	0.01	(-0.01, 0.01)
Case Characteristics		
Year	0.04	(0.02, 0.05)
Co-defendants	0.18	(0.11, 0.26)
Public protection sentence	0.92	(0.81, 1.03)
Plead guilty before trial	-0.63	(-0.70, -0.56)
Mitigating factors	-0.31	(-0.41, -0.21)
Sentenced to remand	0.52	(0.44, 0.60)
Victim impact statement	0.27	(0.19, 0.35)
Injuries caused	0.12	(0.02, 0.22)
Second offence	0.69	(0.61, 0.77)
Third offence	0.44	(0.32, 0.55)
Fourth offence	0.56	(0.41, 0.71)
Offence type (ref: murder)		
Attempted murder	-2.28	(-2.54, -2.02)
GBH with intent	-5.68	(-5.94, -5.43)
GBH	-3.64	(-3.86, -3.43)
Conspiracy to commit GBH	-3.78	(-4.22, -3.32)
ABH	-5.33	(-5.56, -5.10)
Unlawful wounding	-5.89	(-6.33, -5.45)
Common assault	-6.38	(-6.68, -6.09)
Affray	-5.72	(-6.01, -5.43)
Violent disorder	-6.06	(-6.34, -5.78)
Arson	-4.32	(-4.63, -4.01)
Arson with intent to endanger life	-4.97	(-5.41, -4.52)
Conspiracy to commit arson	-3.78	(-4.50, -2.98)
Robbery	-3.93	(-4.16, -3.71)
Attempted robbery	-4.40	(-4.72, -4.08)
Conspiracy to commit robbery	-3.55	(-3.82, -3.27)
Kidnap	-4.07	(-4.40, -3.73)
Conspiracy to kidnap	-4.01	(-4.65, -3.34)
Rape	-3.38	(-3.60, -3.16)
Indecent assault	-4.32	(-4.59, -4.06)
Rape of a child under 13	-4.24	(-4.48, -4.00)
Attempted rape	-4.92	(-5.20, -4.65)
Attempted rape of a child under 16	-3.15	(-3.45, -2.86)
Indecent assault of a female under 13	-4.19	(-4.68, -3.67)
Assault by penetration	-3.68	(-4.27, -3.02)
Sex activity with a child	-3.95	(-4.35, -3.52)
Sexual assault	-4.21	(-4.71, -3.68)
Dangerous driving	-6.08	(-6.37, -5.79)
Causing death by careless driving	-6.07	(-6.41, -5.73)
Causing death by dangerous driving	-4.06	(-4.30, -3.82)
Causing death by careless driving while over the alcohol limit	-3.98	(-4.38, -3.57)
Causing serious injury by dangerous driving	-4.79	(-5.23, -4.33)
Random effects		
σ_{court}	0.18	(0.13, 0.24)
$\sigma_{judge.rotate}$	0.36	(0.30, 0.44)
$\sigma_{judge.stay}$	0.47	(0.39, 0.56)

Response variable: duration of custody in months

Table 2: Full Results for Model 2