Scottish Sentencing Council

Exploring Unwarranted Disparities in Sentencing in Scotland

Draft submitted to the Scottish Sentencing Council in March 2025
Final version received February 2026
Published February 2026

Authors: Prof Cyrus Tata (University of Strathclyde)

Dr Jay Gormley (University of Glasgow)

Prof Melissa Hamilton (University of Surrey)

Prof Jose Pina-Sánchez, (University of Leeds)

scottishsentencingcouncil.org.uk
sentencingcouncil@scotcourts.gov.uk

Table of Contents

Executive Summary

The Scottish Sentencing Council (SSC) was established in October 2015 as an independent, advisory body as a result of the Criminal Justice and Licensing (Scotland) Act 2010. It carries out a range of work concerning sentencing in Scotland and its responsibilities include:

- Preparing sentencing guidelines for the courts;
- Publishing guideline judgments issued by the courts;
- Publishing information about sentences handed down by the courts.

The SSC can also conduct research and provide general advice or guidance and, under law, the Council must:

- Promote consistency in sentencing;
- Assist the development of sentencing policy;
- Promote greater awareness and understanding of sentencing.

The SSC has a statutory duty to publish, from time to time, data on sentencing in Scotland. To fulfil that duty, the Council commissioned us to conduct exploratory research to advance knowledge of sentencing data in Scotland. The aim was to scrutinise potential indicators of unwarranted disparities in sentencing while paying close attention to complexities, including confounding features relevant to distinguishing warranted from unwarranted disparities. In accordance with the project specification, we concentrated on sociodemographic disparities (e.g. race/ethnicity, gender, etc) rather than the absence or presence of unwarranted disparities between judicial sentencers.[1]

Consistency and Unwarranted Disparities

As in other areas of official decision-making, consistency in sentencing is central to the principle of legal equality. Genuine consistency in sentencing does not mean uniformity: treating every case as the same regardless of all relevant case differences. Simply put, the principle of consistency means that like cases should be treated alike and unlike cases should be treated differently. Unwarranted disparities relate to differences in sentencing outcomes for non-

[1] The question of the presence or absence, and nature of, any unwarranted disparities between judicial sentencers, (noted briefly below), was not one we were commissioned to investigate.

legitimate reasons (e.g. on the grounds of the sentenced person's 'race' or ethnicity[2]) between otherwise similar cases.

To be able to promote the consistency of sentencing, (and thereby reduce unwarranted disparities such as on grounds of race/ethnicity or on the basis of the individual judicial sentencer), it is essential that a body like the SSC is informed by high-quality data about sentencing patterns for different kinds of cases. Without that, it would be impossible to know whether or not there is consistency or unwarranted disparity. High-quality data about patterns of sentencing according to a range of key criteria (relative seriousness of the offending, criminal history, etc) is essential to the pursuit of consistency in sentencing.

Comparing Like with Like

A survey was conducted of the contemporary Scottish and international literature from several broadly comparable jurisdictions. Chapter 2 examines the extent to which research studies have been able to conduct controlled analyses of sentencing patterns for like cases.

Analysing sentencing disparities across the entire population of cases can mask more nuanced differences that may emerge if the research were to concentrate on specific subsets of the population. These subsets could be defined *by* various offence factors, (e.g. cases with and without multiple convictions of similarly serious offences of a similar kind), the type of sentence given, or sociodemographic groupings. Moreover, research that fails to explore intersectionality may overlook crucial variations. For instance, a study might not capture the distinct experiences of young men from ethnic minority backgrounds, thereby missing critical insights into how sentencing outcomes differ across intersecting identities and circumstances.

Nonetheless, in a number of studies in other countries where such controlled and rigorous studies have succeeded in comparing otherwise similar cases, it is possible to know whether and in what ways there may or may not be areas of unwarranted disparity.

[2] In objective terms, 'ethnicity' and especially 'race' are problematic terms. They each bear a long, troubled (and troubling) history. 'Race' refers to the pseudo-scientific notion that humanity can be divided into separate groups each with supposedly distinct, innate characteristics. The early study of crime helped to propel these racial beliefs, which continue to play out today (Phillips, Bowling and Palmer 2023). While the term 'ethnicity' may not fix human difference as immutable and biological, its study (along with the closely associated concept of 'culture') has not been immune from criticism in terms of apparent stereotyping, which easily bleeds into racialised ideas. However, ignoring the idea of race and ethnicity as key categories of perceived social difference would be to ignore the reality of unwarranted discrimination (e.g. Phillips, Bowling and Palmer 2023; Rossmanith 2023). This report therefore employs the terms 'race' and 'ethnicity' in discussing previous work on sentencing because these are terms that previous work has used – frequently interchangeably, often because the data on which it relied has done so.

Attempts to Assess the Absence or Presence of Unwarranted Disparities in Scotland

The picture in Scotland is, however, less promising. Our examination of recent attempts to assess and measure the presence or absence of unwarranted disparities using official data reveals an unconvincing picture. Recent work has been unable to control for fundamentally important case criteria (e.g. the varying seriousness of offending; the presence or absence of analogous previous convictions; the presence or absence of an early guilty plea, etc). In other words, such work demonstrates the significant challenges in producing a meaningful comparison of like cases with like cases.

This project aimed to improve on this situation. It sought access to disaggregated official data so as to conduct a properly controlled analysis in order to shed new light on patterns of Scottish sentencing practices according to a range of specified and relevant criteria. In this way, the analysis would be able to compare like with like. However, due to limitations beyond our and the Council's control, this proved not to be possible.

A detailed investigation (documented in Chapter 3) was carried out to determine what information relevant to legitimate sentencing considerations is recorded and held by official data. This was focused on well-established questions of fundamental importance to sentencing, including, for example, the ability simultaneously to control for:

- The relative seriousness of offending in different cases;
- The ability to distinguish between single and multi-conviction cases, especially in cases where there is more than one conviction of similar seriousness;
- Previous criminal history, including, for example, previous analogous convictions;
- The ability to control for plea ('guilty' or 'not guilty');
- Offender characteristics (including family circumstances, primary caring responsibilities (e.g. of children), mental and physical health, income, age;
- The backdating of custodial sentences to cover time on remand;
- Different conditions of a community-based sentence imposed by the courts.

In all of these (and other) instances, information was found either not to be collected by administrative sources (and thereby missing from official data), practically inaccessible, or, where it is collected, cannot be combined with other variables to conduct a controlled analysis of sentencing patterns according to key criteria of interest (see further Chapter 3).

To put it simply, the evidence of whether there are or are not unwarranted disparities is weak because the available official data does not appear to be capable of enabling a proper analysis which controls for the most important criteria.

Applying Advanced Mathematical Techniques

Having observed the severe limitations of official data, Chapter 4 examines the possibility of using advanced and novel mathematical techniques to try to extract greater depth and meaning from official data. A number of novel techniques are scrutinised, which may ameliorate the issues already identified. As promising as these techniques are, ultimately, they are limited by the quality of official data they are working with. No matter how novel and sophisticated the advanced techniques may be, there remain fundamental issues with that official data. Because of the limitations in official data, such techniques have to resort to making certain assumptions and drawing conditional extrapolations.

Official Data about Sentencing in Scotland is Fragmented

The limitations of official data in describing sentencing patterns in any more than the most superficial ways should not be entirely surprising. Ultimately, official sentencing data is a product of its fragmented administrative sources. Traditionally, different administrative agencies (police, prosecution, courts, social work, prisons, etc.) have collected, and continue to collect, data largely for their own daily operational purposes of processing people. Their recording systems were not established as research tools to provide meaningful insights into sentencing practice.

This fragmentation is a very long-standing (perhaps intractable) problem, and despite the admirable efforts to create one single, accessible and coordinated data set, it seems likely that this fragmentation and limited data pertinent to sentencing will remain a prominent feature of the Scottish criminal justice landscape for the foreseeable future.

The Need to Develop New Sources of Data about Sentencing

Therefore, the SSC faces the unavoidable challenge of devising a means for the systematic collection and analysis of data to support its statutory duties. This will enable SSC to carry out its core functions of promoting consistency, developing policy, and promoting public knowledge and awareness of sentencing.

Promoting the Consistency of Sentencing

To be able to promote the consistency of sentencing, it is essential to be informed by high-quality data about sentencing patterns for different kinds of cases according to a range of key criteria (e.g. relative seriousness of the offending, criminal history, etc). Such data is essential to the pursuit of consistency in sentencing.

Assisting the Development of Sentencing Policy

The development of sentencing policy requires regard to systematic information about patterns of sentencing according to different criteria. One obvious area is the development of sentencing guidelines. For example, the actual effect (or non-effect) of various key changes to sentencing policy (e.g. a reduced sentence because of a guilty plea) cannot be gauged by official data derived from administrative sources.

Promoting Public Knowledge and Understanding of Sentencing

The provision of high-quality data about sentencing patterns is vital to work seeking to promote more accurate knowledge of sentencing and thus public confidence in sentencing.

Both in Scotland and comparable jurisdictions, sentencing is widely perceived by the public to be too lenient. Other jurisdictions have been able to show that this perception of lenience may not be borne out by the evidence of patterns of sentencing. By focusing on specific kinds of cases (e.g. through the use of hypothetical scenarios/vignettes), research has asked members of the public what their preferred sentence would be and what they believe to be the typical sentence for this kind of case. Often, the public tends to think of the typical sentence as considerably more lenient than the sentence they would prefer to see passed given the same information.

Crucially, however, this can then be triangulated with the actual patterns of sentencing in that kind of case. Where data on the patterns of sentencing given the specifics of that case *is* available, it is possible to assess the extent to which the preferred sentence suggested by members of the public aligns with the reality of actual sentencing patterns. A powerful ‘good news’ story can be told. Members of the public may express some surprise and relief that the sentence they think is appropriate is more or less what is being done, but that they tended to imagine that the reality was characterised by far greater lenience than it is.

The problem which Scotland faces is that, by and large, because of the lack of high-quality sentencing data, it is very difficult to tell this good news story because triangulation between public preference, public expectations, *and* the reality of actual patterns of sentencing cannot be completed because the data about actual sentencing patterns is inadequate.

The provision of that data would allow it to be clearly demonstrated to a range of audiences that the public’s feeling of excessive leniency in sentencing may not, in fact, be borne out by the reality. It could then enable bodies like the SSC to examine with greater precision where there may be significant differences between public preferences and actual sentencing patterns.

This triangulation between public preferences, expectations of sentencing patterns and the reality of actual sentencing patterns could help to promote public confidence in sentencing and trust in the judiciary more generally.

Ways of Developing New Sources of Data about Sentencing Patterns in Scotland

Official sources of data about sentencing derived from administrative agencies in Scotland suffer from some fundamental flaws.

The SSC is currently at a distinct disadvantage compared to similar bodies, such as the Sentencing Council for England and Wales (SCEW), to the extent that it is currently unable to draw on comprehensive and relatively in-depth data about sentencing patterns. Therefore, the SSC faces the unavoidable challenge of devising a means for the collection and analysis of high-quality sentencing data to support its statutory duties. While this may be a challenge, it also presents the SSC with the opportunity to develop a method of recording and representing data that meet its information requirements.

Chapter 5 of this report outlines and evaluates a range of options to develop new and improved sources of data. These approaches include large or small census studies. Chapter 5 examines the development of the examples of census studies: the Crown Court Sentencing Survey (CCSS) and the Sentencing Information System for the High Court of Justiciary in Scotland. Because they are dedicated to the collection of data that is specifically useful to understanding sentencing patterns in genuinely similar cases, census studies produce much more in-depth and meaningful data, which can be far more useful to bodies such as sentencing councils than official administrative data.

A challenge raised by census studies is who collects such information and how (especially whether it is collected contemporaneously, rather than from court archives). As much as the question of resourcing, it underlines the importance of leadership. Clerks of court and especially judicial sentencers are far more likely to be willing to record information if they can see the point of doing so and how it will add benefit to the system in which they work: including, for instance, helping to improve public trust in and knowledge and understanding of sentencing. Even where census studies have had a limited life, they can leave a positive legacy. For example, the CCSS data has proved to be an invaluable resource, even after its discontinuation, enabling the SCEW to investigate and answer contentious topics of heightened public concern.

If it was determined that a comprehensive census approach would be too sizeable and/or too difficult to ‘sell’ to those collecting the data, a more modest alternative could be the use of a series of mini-censuses. These could be targeted to address specific issues, which the SSC is planning to address in the future, and/or matters of particular concern. The collection of some baseline data offers twin benefits. First, data on public sentence preferences and on public expectations of typical sentence patterns could be triangulated with data on actual sentencing patterns for the same kinds of cases. Such triangulation can potentially tell a powerful good news story by showing clearly that judicial sentencing is not as lenient or ‘out of touch’ as is often believed. Second, a census approach, or mini-census approach, could provide pre- and post-guideline data so enabling a more or less direct evaluation of the effect of that guideline.

An additional and more agile approach would be to deploy two sets of hypothetical case-scenario exercises run in parallel: one with members of the public and the other with judicial sentencers. These exercises would ask the two groups to consider the same case scenarios. The public would be asked for their preferred sentences and their expectations of what they think the typical sentence passed by sentencers would be. By asking the judicial sentencer group what they would in fact pass for the same scenario, it will become possible to draw a direct comparison in the responses of the two groups. This would help the SSC to identify areas of convergence and divergence between the two groups so as to inform future public engagement and education work as well as potential guidance to sentencers. It could also demonstrate instances of where public expectation (e.g. of leniency) is not borne out by the reality of what sentencers actually think and do.

Census (including mini-census) studies can be fruitfully combined with qualitative approaches, which answer different kinds of in-depth questions. Census and qualitative approaches should go hand in hand and are best seen as mutually informative. The use of AI (such as Large Language Models) may also be worth consideration, albeit cautiously given that existing work is at a relatively early stage. Using a combination of approaches will tend to provide a more rounded picture.

Conclusions

Options for developing new sources of in-depth and meaningful data about sentencing patterns present both challenges and opportunities to the work of the SSC. What is abundantly apparent is that reliance on official data derived from administrative sources cannot at present fulfil the SSC's needs in discharging its statutory functions. These issues in the official data are long-standing. It appears unlikely that this situation will improve significantly in the short- or even medium-term. Until that significant improvement is achieved, research studies, based on targeted samples of cases, should be considered.

Chapter 1: Exploring warranted and unwarranted disparities in sentencing

1.1 Introduction

The Scottish Sentencing Council (SSC) was established in October 2015 as an independent, advisory body as a result of the Criminal Justice and Licensing (Scotland) Act 2010. It conducts work concerning sentencing in Scotland and its statutory duties include:

- Preparing sentencing guidelines for the courts;
- Publishing guideline judgments issued by the courts;
- Publishing information about sentences handed down by the courts.

It also conducts and commissions research and provides general advice or guidance. Under law, the Council must:

- Promote consistency in sentencing;
- Assist the development of sentencing policy;
- Promote greater awareness and understanding of sentencing.

The Council has a statutory duty to publish, from time to time, data[3] on sentencing in Scotland. To fulfil that duty, the Council commissioned research to advance knowledge of sentencing data in Scotland.[4] The aim was to scrutinise potential indicators of unwarranted disparities in sentencing (or their absence) while paying close attention to complexities, including confounding features relevant to distinguishing warranted from unwarranted disparities.

The initial aspiration for this project was to secure access to disaggregated official data (which is derived from administrative datasets) to conduct a controlled analysis. This would shed new light on patterns of Scottish sentencing practices according to a range of specified and relevant criteria. However, due to limitations beyond our and the Council's control (see Chapter 3), this was not possible. Hence, this project focused on ways to inform the Council's strategy in this area going forward. To that end, this first chapter examines the nature of unwarranted disparities in sentencing to provide context for the report. Such unwarranted disparities could pertain to differences between, for example, protected characteristics under the Equality Act 2010 or other demographic groups.

[3] The term 'data' is used by many, even most, authors and in everyday language in the singular rather than plural. This report follows this convention.
[4] The project ran from February 2024 to March 2025.

1.2 The Need to Explore Questions of Disparity

Questions of unwarranted disparities in sentencing are vital to the principles of justice. For example, the SSC guideline on the principles and purposes of sentencing states that 'sentences in Scotland must be fair and proportionate' and, *inter alia*, this necessitates that 'people should be treated equally, without discrimination' (Scottish Sentencing Council, 2018, p. 3). Similarly, the English and Welsh Sentencing Council (SCEW) has noted that:

> 'The Sentencing Council has an overarching objective to promote a fair approach to sentencing, thus placing equality, diversity and inclusion at the heart of our work. As part of our five-year plan, we have set ourselves a strategic objective:
>
> To explore the potential for the Council's work inadvertently to cause disparity in sentencing across demographic groups by commissioning independent external contractors to undertake a project to review a sample of key guidelines and processes.' (Sentencing Council of England and Wales 2023)

To ensure these objectives are met, it is necessary for all jurisdictions to have sufficient high-quality sentencing data to monitor and evaluate potential disparities.

In addition to unwarranted disparities being a vital matter that is essential to fundamental principles of justice, it is also vital to public confidence in the justice system. Perceptions of unwarranted disparity in sentencing (even if incorrect) can severely damage public confidence (including specific demographics of the public) in the justice system. As such, it is also necessary for jurisdictions to have sufficient high-quality sentencing data to monitor for unwarranted disparities and to be able to communicate sentencing practices to groups such as the public. Indeed, as Chapter 5 explains, having access to high-quality sentencing data offers opportunities for communicating with the public and promoting awareness and confidence in sentencing. This data would be especially beneficial given that public knowledge about the realities of sentencing practice can be low (Black, et al 2019; Biggs et al, 2021; Archer et al 2022; Hockaday et al, 2025).

Another reason to be concerned with questions of disparity is that, while it is hoped that unwarranted disparities do not exist, the experiences of other jurisdictions show a need for caution. Data from other jurisdictions is covered in more detail in Chapter 2 but, for now, it can

be noted that while there are various disparities to question, racial/ethnic disparities have been one notable area of concern. Experiences in several jurisdictions show that concerning disparities may emerge, which may be for a variety of reasons. Perhaps one of the most discussed examples of disparity in sentencing is illustrated by the USA's so-called 'War on Drugs' in the 1980s, 90s and 2000s. This 'War on Drugs' included the penalisation of those people possessing 'crack' much more harshly than 'powder cocaine'. Despite their pharmacological commonality, Federal law penalised crack cocaine at a 100-1 sentencing ratio compared to powder cocaine.[5] While an ostensibly neutral provision, this has been attributed to significant sentencing disparities that resulted in black and Hispanic offenders (who were far more likely to be convicted of possession of crack cocaine than powder cocaine) receiving much harsher sentences than their white counterparts, who tended to be convicted of powder cocaine offences (Davis, 2010; Sklansky, 1995). Thus, even provisions that are not directly discriminatory can have indirectly discriminatory effects. In England and Wales, Hood's seminal work, which remains one of the most careful studies in the area, highlighted disparities earlier in the criminal justice process which had a compounding and cumulative effect on sentencing, but also (and notwithstanding these earlier discriminatory impacts) some unwarranted disparities at sentencing itself (Hood 1992; Feilzer and Hood, 2004). Recent work also notes the ongoing questions about disparities in sentencing (e.g. Veiga, Pina-Sánchez and Lewis, 2023; Lymperopoulou, 2024), and this is discussed further in Chapter 2.

1.3 Identifying Unwarranted Disparity

Sentencing is one of the most complex exercises undertaken by the criminal justice system. Depending on the specific case, sentences must serve a range of competing purposes (Scottish Sentencing Council, 2018; Tata, 2020). One implication of the individualised nature of sentencing is that it is necessary to distinguish between disparities that may be warranted (e.g. those arising due to legitimate reasons such as variations in offence seriousness) and those that are unwarranted (e.g. those that arise from discriminatory practices). This is not always an easy task as sentencing data may omit key factors meaning that legitimate reasons for

[5] 'It does seem rather strange, on the face of it, that ten grams of 'crack' (a drug of choice, apparently, of poor blacks in inner cities) has the same penal value as one kilogram of pure cocaine (the drug of choice of the rich and famous). The fact that cocaine can easily be made into crack (thus increasing the penal value of the drugs in one's possession dramatically) is lost in the 'Guidelines'.' (Doob 1995, 236-7; Provine 2002).

sentence variation are not properly reflected (see Chapter 3) and/or that potentially problematic disparities are unclear. As such, poor quality data may be overinclusive or underinclusive when it comes to identifying disparity: it may suggest disparities where there are none or may not suggest unwarranted disparities where they exist.

As a result, not only is it necessary to identify where disparity may emerge, but it is also necessary to establish why it emerges in order to determine if/when remedial action is needed. For example, in the context of ethnic disparity in England and Wales, Veiga et al note that:

> 'Despite the magnitude of these reported disparities, critics have been quick to point out that the extent to which they prove the existence of discriminatory practices in sentencing is questionable. Some scholars altogether reject the evidence of discrimination, arguing that the impossibility of controlling for every relevant variable considered by a judge means that one cannot truly know whether disparities are the result of discriminatory treatment (Halevy,1995; Wilbanks,1987; Wooldredge,1998). Many others have rightly raised questions about the robustness of these findings. They are wary in claiming that disparities are definitely the result of discrimination since the evidence is inconclusive. These researchers highlight the pervasive methodological problem that has affected all sentencing research relying on observational data (Baumer, 2013; Holdaway 1997; Reiner, 1993; Ulmer, 2012). Confounding bias is inherent to regression models, which produce approximate comparisons but are unable to produce 'like with like' comparisons (Pina-Sánchez, Roberts and Sferopoulos, 2019; Roberts and Bild, 2021). Confounding bias is inherent to regression models, which produce approximate comparisons but are unable to produce 'like with like' comparisons.' (Veiga, Pina-Sánchez and Lewis, 2023, p. 168)

These methodological challenges for statistical analysis will be explored further in this report in Chapter 4. However, for now, it is worth highlighting that sufficiently high-quality data on sentencing is essential to bodies such as sentencing councils and groups such as sentencers for at least two reasons. First, having deep and meaningful data about sentencing practice can enable a proactive account of decisions and fend off unwarranted criticism that may emerge from incorrect assumptions. Indeed, in the absence of high-quality data, even a sentencing system that was not discriminatory may be perceived as discriminatory. Secondly, such data may also be useful where legitimate issues emerge and a solution is implemented (e.g. via guidance).

Indeed, it is safe to say that questions of disparities in sentencing are one of the key issues of the moment in England and Wales. This English and Welsh experience also highlights the challenges of assessing unwarranted disparities at sentencing given that, even where it is clear that disparity exists, the important questions of how and why disparity exists and how it should be addressed can remain opaque. For example, we need to know whether/what role sentencing plays in any disparity. This question of roles is complex as it is also possible, as noted above, for example by Hood (1992), that much (but not all) of the observed unwarranted disparities can emerge from decisions prior to sentencing – including elsewhere in the criminal process through

investigative and/or prosecutorial practice. Recently, this issue has become particularly animated in England and Wales. The SCEW has noted that:

> 'In relation to offenders from ethnic minorities, there is good evidence (both from the Council's own research and other independent research) that in relation to some types of offence there is a disparity in sentence outcomes as between white offenders and offenders from an ethnic minority. Offenders from some ethnic minority backgrounds are more likely to receive an immediate custodial sentence than white offenders. In some offence specific guidelines this fact is highlighted. Why this disparity exists remains unclear.'[6]

Thus, it is vital to have sufficient high-quality sentencing data to monitor and be sure that unwarranted disparities do not emerge and, if they do emerge, to identify the source and defend the solution.

1.4 Conclusion

Unwarranted sentencing disparities are a cause for grave concern in several jurisdictions. Indeed, arguably, the importance of questions of disparities requires a proactive approach to explore this complex area. However, as this report notes (see Chapters 2 and 3), limited data on sentencing in Scotland means that it is difficult to establish what disparities may or may not exist. What is needed is better quality data on sentencing practice (see Chapter 5). This data is vital to monitor and accurately describe unwarranted disparities so as to: (1) correct any unwarranted disparities where they are identified; and (2) demonstrate evidence of improvement or even the lack of unwarranted disparities in certain areas. In terms of correcting disparities, meaningful in-depth data could help justify actions taken if it can be shown why a disparity emerges and how an action would ameliorate an issue. In terms of demonstrating an appropriate lack of disparities, meaningful in-depth data may help correct false assumptions or misrepresentations from poor-quality data, which may be misleading. Thus, whatever the situation in Scotland may be, high-quality sentencing data is absolutely essential to: the promotion of consistency in sentencing; the informed development of sentencing policy; and efforts to inform public knowledge and awareness of (and so confidence in) sentencing and the independent judiciary.

[6] Letter to the Lord Chancellor. <https://www.sentencingcouncil.org.uk/wp-content/uploads/20250310-Letter-from-Lord-Justice-William-Davis-to-Lord-Chancellor-on-Imposition-FOR-PUBLICATION.pdf>

Chapter 2: Evidence of Unwarranted Disparities in Scotland and in Comparable Jurisdictions

2.1 Introduction

This chapter examines contemporary evidence of potential disparities by certain sociodemographic characteristics in Scotland and comparable jurisdictions regarding the composition of their prison populations and sentencing outcomes. It synthesises key findings from landmark research studies, emphasising the methodological approaches and the inherent challenges associated with comparative analysis. Additionally, the chapter summarises recent research on the perceptions of professionals and the public regarding these findings.

2.2 Scotland

As of 2022, the Scottish prison population presented as 95% White (including White minority ethnic), 4% female, and the largest age proportion was in the 25-34 years group (Sturge, 2023). The latest 2023-24 Prison Population Statistics note that:

> 'The majority of individuals experiencing imprisonment in 2022-23 identified as White (94%).
>
> Reflecting the overall pattern of individuals experiencing imprisonment, there were small increases in the number of individuals experiencing imprisonment across ethnic groups between 2021-22 and 2022-23: White (+41 to 13,617); Asian (+26 to 325); Other ethnic group (+35 to 231); African, Caribbean or Black (+15 to 274); Mixed or Multiple (+1 to 79).
>
> Overall there was little change to the proportion of individuals each ethnic group represented between 2021-22 and 2022-23. In 2022-23, the rate of imprisonment for the White, Asian, and Mixed or Multiple ethnic groups remained similar to the previous reporting year (3.2, 2.6 and 4.5 per 1,000 respectively). The imprisonment rates remained higher for the African, Caribbean and Black and Other Ethnic groups, and showed an increase from the rates seen in 2021-22 (from 8.3 to 8.7 and from 7.3 to 8.5 per 1,000 respectively).'[7]

Comprehensive datasets, with appropriate controls for case seriousness, typically are not publicly available about contemporary sentencing outcomes in Scotland (Gormley, et al 2022). Nonetheless, the government has recently published a few data points that might be capable of providing some hints about potential demographic disparities.

[7] To compare, in the 2022 Census of the general population of Scotland, 93% were White, 51% female, and the largest age grouping was age 50-59. Scotland's Census 2022 - National Records of Scotland (2024), https://www.scotlandscensus.gov.uk/search-the-census#/search-by.

A new statistical analysis produced by the Scottish Government (2023a) relied upon a dataset compiled from the Scottish Courts and Tribunal Service (SCTS)[8] to begin to assess proportional representations of ethnic groups at different stages of the criminal justice process. The section of the study on sentencing focused on outcomes from mid-2017 to early 2023.[9]

Initial analyses, which, very importantly, were unadjusted for any controls, appeared on their face to suggest various disparities. However, as the authors of the report themselves emphasise, the depth and sensitivity of their investigations were severely compromised by the data available to them:

> 'Controls have not been applied for any characteristics besides ethnic group (such as sex, age, average income, geography, offence mix or offender history), so it is not possible to estimate what proportion of differences found are directly attributable to ethnicity. Whilst the differences reported may suggest areas which merit further investigation, additional work controlling for more variables would be required to help determine what the sources of these differences might be.' (Scottish Government 2023a, p.4)

Given these serious limitations, the authors describe this first effort at such an analysis as 'experimental' and conclude that it is not sufficiently robust to inform the making of policy:

> 'Due to the data issues identified in earlier sections, care needs to be taken when interpreting these results. As a consequence, the analysis in this report is experimental and is not on its own sufficiently robust to be used, for example, to inform policy recommendations.' (Scottish Government 2023a, p. 25)

Of course, all statistical analyses should be interpreted with care. One should always be mindful of the assumptions on which an analysis is based and what is and is not controlled for. This first attempt to analyse ethnic differences was, however, as the authors correctly note, no more than a tentative and 'experimental' first step. The meaningfulness of the data in terms of sentencing was low. For example, to make valid comparisons one would need to control for case seriousness. As Chapter 4 sets out, this includes the seriousness of sentenced offence/s; criminal history; and the timing of a guilty plea (if there is one), to name but a few fundamentally important characteristics. The analysis sought to do its best to control for case seriousness, but this was extremely limited. It was only able to control for extremely broad offence categories, rather than any direct way of controlling for the seriousness of the conviction/s. For example,

---

[8] The publication states that the project merged two 'experimental' datasets that were 'provided to Scottish Government JAS [Justice Analytical Services] Division from SCTS systems. They were specifically designed to assist JAS with the analysis of remand and bail and to analyse the impact of presumption against short sentences.' (Scottish Government 2023a: 29)

[9] It is noted that in terms of racial categorization in this study, it used the term 'White British' to encompass those individuals identified as either White Scottish or other British. Those identified as White, but neither Scottish nor otherwise British were classified within the White Minority Ethnic group.

using the official data available, it compared the odds of the imposition of custodial sentence by ethnic group by breaking this down according to extremely broad offence types. These included: 'Non-Sexual Crimes of Violence'; 'Sexual Crimes'; 'Crimes of Dishonesty'; 'Crimes Against Society'; 'Miscellaneous Offences'; 'Road Traffic Offences'.[10] None of these categories is capable of meaningfully controlling for the variable seriousness of convictions for which offenders may be sentenced. All of these broad categories contain within them convictions which vary enormously in terms of seriousness. This is a point recognised not only by professionals involved in the sentencing process, but also by members of the public (e.g. Archer et al 2022; Black et al 2019; Biggs et al 2021; Hockaday et al 2025).

On the basis of the fundamental limitations of the official data analysed, the report then suggested that White British individuals were more likely to be sentenced to prison compared to ethnic minorities, with Asian defendants being the least likely to receive custodial sentences (Scottish Government, 2023a). Conversely, Asian individuals were most likely to receive monetary fines, while White British were least likely. The proportions of Black, Mixed, White Minority, and Other groups fell between these extremes for receiving both custodial and monetary penalties.

Proceeding to explore sentence lengths for custodial penalties, the report found that White British individuals generally received shorter prison terms compared to minorities, with Asians receiving the longest sentences. In sum, while the White British group was said to be the most likely to be incarcerated, their custodial terms were the shortest; the results for the Asian group was the opposite in being the least likely to be incarcerated, yet with the lengthiest sentences when incarcerated. However, these trends varied by type of offence. As an illustration, for violent crimes, the Asian and White Minority ethnic groups each received longer sentences than White British, whereas Black, Mixed, and Other groups received shorter sentences.

As noted above, the authors cautioned against drawing definitive conclusions due to the extremely limited controls and data constraints (Scottish Government, 2023a). In addition, another caution is warranted. The unit of analysis was the conviction rather than the individual being sentenced, thereby potentially inflating statistics as any person sentenced in respect of multiple convictions will be represented in the data multiple times. As we note in Chapter 3, this apparently technical difference in how 'a case' is counted can present serious problems to those seeking to 'marry-up' official data from different criminal justice agencies. It is a problem not unique to Scotland (Gormley et al 2023; Tata 1997). Over time, each agency has developed its own case-counting rules, which make sense from the perspective of the daily operation of that agency, but less in terms of informing policy-making. It renders the attempt to combine data-sets derived from the various recording practices of different agencies extremely challenging.

[10] Although the publication does not explain why these were selected, it seems likely that it is because together these very broad categories comprise the overwhelming majority of cases.

In another government publication examining women's involvement in the criminal justice system, unadjusted data from 2019-2020 indicated a higher likelihood for males to be sentenced to prison in Scotland, partly attributed to their higher prevalence in convictions for more serious offences (Scottish Government, 2022). When prison terms were imposed, females tended to receive shorter sentences than males. The report highlighted a specific example of gender disparity in sentencing: with homicide or culpable homicide convictions, males were disproportionately more likely to receive life sentences compared to females (Scottish Government, 2022).

That said, as the report noted, it is extremely difficult to identify the existence or absence of unwarranted disparities given the limitations of official data, not least in terms of relative seriousness of offending. As noted above, broad categories like ‘violence’ or even ‘culpable homicide’ contain within them huge variations in seriousness, including cases with single as opposed to multiple convictions of similar seriousness. The authors observe:

> ‘Please note that sentencing decisions are reflective of a number of factors such as the severity of the crime and whether the individual has offended in the past. In addition, the decision on what type of punishment is reasonable will be based on the personal circumstances of the offender. *These statistics do not take account of these factors.*’ (Scottish Government 2022, p.25, emphasis added)

Similarly, unadjusted statistics for the 2021-2022 period reiterated the trend of men being more prone to custodial sentences than women (Scottish Government, 2023b). During the same period, women were observed to be more likely to receive alternative sentences, typically admonishments, rather than penalties involving custody, community sentences, or fines. However, given the limitations of the official data relied upon, and the weak ability to control for case seriousness, such ‘likelihood’ should be treated with the utmost caution. It cannot be safely assumed that differences in terms of case seriousness (offence, offender characteristics etc) are simply evenly distributed between males and females.

2.3 England and Wales

As of June 2023, 27% of individuals within the prison population of England and Wales identified as (non-White) ethnic minorities (compared to 18% in the general population), 96% were males, and the largest age group was 30-39 years (Sturge, 2023). Current estimates are that there are 152 male prisoners per 100,000 men in the population, compared to 13 female prisoners per 100,000 females in the population (Sturge, 2023).

The Ministry of Justice (MoJ) typically compiles its research reports on sentencing by utilising data extracted from case management systems. One such publication (Hopkins, Uhrig, and Colahan, 2016a) used data from the Police National Computer and MoJ court database for the year 2015 on recordable offences sentenced in Crown Courts or by magistrates. For the results

discussed in this paragraph, the models included controls for gender, ethnicity, age, offence type and criminal history. Importantly, however, no controls were applied for other aspects relating to culpability, harm, or aggravating/mitigating factors. This study reviewed ethnic differences in a single category of Black, Asian or Minor Ethnic (BAME) against the alternative for Whites (which included White ethnic minorities). Results revealed that the BAME group was statistically more likely to be sentenced to immediate custody than Whites (Hopkins, Uhrig, and Colahan, 2016a). Men were more likely to face a custodial term than women across offence types, though with some variations in magnitude, meaning that gender has a greater effect in some crime types than others (e.g. lowest for small theft and highest for drug importation/production), though again the seriousness of offending was not controlled for. Looking at interactions, BAME males and BAME females faced higher odds of prison than White males and White females, respectively, with the magnitude of the difference being greater in the male comparison. With respect to age, statistically significant differences in the odds of prison were found when using ordinal age rankings, though not in a linear fashion.

A second study by the same researchers used the same study year of 2015 (Hopkins, Uhrig, and Colahan, 2016b). However, this study was more limited in focusing on recordable *and* indictable offences sentenced in the Crown Courts (i.e. more serious crimes) and added the presence of a guilty plea to its controls. In one analysis, the authors operationalised ethnicity into five groups, finding that compared to Whites, each of the groups of Asian, Black, and Chinese/Other were more likely to be sentenced to prison, with no statistical difference detected for the Mixed ethnicity group. However, when the researchers drew on the dichotomous BAME versus White comparison and looked at three major offence categorizations, results varied; a statistically significant increase in the likelihood of a prison sentence for BAME offenders was present for drug offences, but no difference was present for either violent or sex offences (Hopkins, Uhrig, and Colahan, 2016b). Men were more likely than women to be sentenced to prison and this result applied to numerous offence types, though the magnitude of the increase varied by offence type. Researchers concluded that their findings overall ‘showed an independent association’ between ‘sex and the odds of imprisonment’ (Hopkins, Uhrig, and Colahan, 2016a: 9) and between ‘ethnicity and the odds of imprisonment’ (Hopkins, Uhrig, and Colahan, 2016b: 10).

A more recent MoJ publication on sentencing outcomes by ethnic groupings uses data from the Court Proceedings Database for the years 2018-2022 regarding indictable offences (Ministry of Justice, 2024). In unadjusted models (without controls), all other ethnic groups were more likely to be sentenced to immediate custody than Whites in most years, though not always reaching statistical significance. All other ethnic groups were sentenced to longer average custodial sentence lengths than Whites, with Black and Asians being the two groups with the most extended durations of custody. Still, when examining major crime categories, no statistically significant differences were observed between ethnic groups regarding sexual offences.

The Ministry of Justice (2024) report then provided results from logistic regression models for each year 2018-2022 to observe ethnic differences when using controls for sex, age, offence

type, court, and plea; there were no controls for criminal history, harm, culpability, or aggravating/mitigating factors. The Black and Mixed-ethnic groups were more likely to receive custodial sentences for indictable offences than the White group (with statistical significance). The Asian group was less likely than the White group to face an immediate custodial sentence, with statistical significance in three years (in contrast, in 2018 the Asian group was more likely to face custody but without statistical significance). The Other group had an increased likelihood of immediate custody (with significance) in two years.

The MoJ (2024) publication next disaggregated ethnicity and race into over a dozen categories and with a more limited comparator group of White British (i.e. here differentiating White ethnic minority).[11] All variations of the Black groups (i.e. African, Caribbean, and other Black) were more likely to receive custodial sentences than White British, though not always achieving statistical significance. As well, two of the mixed ethnicity pairings, being from the White/Black African subgroup and White/Black Caribbean subgroup, were more likely to receive custodial sentences in most years, though not always with statistical significance. In sum, any Black ethnicity in these models with controls was rather consistently associated with a greater likelihood of immediate custody. In contrast, disaggregating the Asian group yielded disparate results. For example, the Chinese subgroup was more likely than White British to face custodial sentences in four years (two of which with statistical significance) but less likely in the final year. The Indian subgroup was consistently less likely to face custodial sentences in all years (with statistical significance in three years). The Pakistani group was more likely to face custodial terms in two years (one with statistical significance) and less likely in three years (one with statistical significance). Overall, the MoJ report concluded that 'ethnicity was independently associated with receiving a custodial sentence' in that Black and Mixed ethnic groups were typically associated with greater odds of receiving a custodial sentence than White British offenders, while the association with Asian and other ethnic groups was less consistent (Ministry of Justice, 2024: 71).

Other recent MoJ studies conveying unadjusted statistics to observe gender differences have found that males were more likely to receive sentences involving immediate custody (Ministry of Justice, 2022) and longer average custodial sentence terms (Ministry of Justice, 2022; Ministry of Justice, 2024).

The MoJ sentencing data are relatively comprehensive. Nonetheless, the MoJ data also have some significant limitations and gaps in crucial information. There is no information on guilty pleas available in the Magistrates' Court data tool. Likewise, the latest versions of the data are missing offenders' previous convictions. In addition, no information is available on deferred sentences.

Very importantly, and in common with other official data in other jurisdictions, the MoJ statistics do not distinguish single from multi-conviction cases, let alone characterise the nature of such

[11] The MoJ (2004) used a combination of both race and ethnicity groupings.

differences. To put this in simple terms, where all other information may be the same, a case involving a single conviction of robbery is re-presented in official/administrative data as no different from a case of twenty convictions for robbery – even though the sentences may be very different. So when an individual is sentenced for more than one conviction, the case appears in the database according to the sentence which is judged to be the most serious. Most jurisdictions employ a 'most serious offence' or 'most severe sentence' rule to record and re-present information about sentencing (Gormley et al 2023). Although these counting rules simplify the presentation of data, the description of patterns of sentencing becomes potentially misleading if the differential seriousness of single and multi-conviction cases fail to be properly distinguished (Tata 1997; 2020).

Gormley et al (2022) report that MoJ data are also less comprehensive about non-custodial sanctions than custodial sanctions. Although there is reasonably detailed breakdown of the principal sanctions, no information is available about the elements of those sanctions. One example is that the nature and number of conditions imposed as part of a community order or suspended sentence order are unavailable.

2.3.1 Crown Court Sentencing Survey

A more sophisticated dataset is the Crown Court Sentencing Survey (CCSS), which was deployed around the years 2010-2015 to capture a great number of data points in sentencing decisions in the Crown Courts in England and Wales. Crown Court judicial sentencers were asked to provide information about each sentenced case. The survey was, therefore, a *census* rather than a *sample* of sentencing decisions in the Crown Court. In the CCSS return, individual judicial sentencers noted important elements of the offence and required the sentencer to indicate the factors taken into account at sentencing. ‘Information derived from the sentencer permits a much more accurate calibration of the influence of various factors upon sentence outcomes’ (Gormley et al 2023, p.86). A distinct advantage of using the CCSS is to capture information to control for a variety of aspects related to criminal history, guilty plea, harm, culpability, and aggravating and mitigating factors. The CCSS was discontinued in 2015 (Gormley et al 2022; 2023).

As we will discuss further in Chapter 5, while the CCSS represented a very significant advance on the use of MoJ data, it is worth noting a number of limitations.

Firstly, the CCSS did not, by definition, include the larger proportion of cases sentenced by Magistrates’ Courts (which roughly account for over 90% of all criminal court cases in England and Wales.

Secondly, the CCSS relied on returns being completed and response rates among courts was variable: over 90% in some Crown Court locations, but lower in others. In the final year of the survey, the completion rate across all Crown Court locations was 60%. If the sentencers with

low response rates were different from those completing all returns, the CCSS would have suffered from bias.

Thirdly, the CCSS records the information taken into account by Crown Court sentencers, but not necessarily *all* the important information affecting the sentence. The form that sentencers were asked to complete lists all the *guideline* factors and provided respondents with the opportunity to note other factors.[12] Other factors which are not captured by the form may have influenced the sentence imposed and the survey cannot detect the influence of extra-legal factors on the sentencing outcome – racial or ethnic status for example. This kind of information has to be gathered in other ways (e.g. by qualitative research involving defendants and legal practitioners, or bespoke academic research.[13]

Fourthly, as noted above, sentencing data needs to accommodate cases involving multiple counts. Although the CCSS permitted the court to note the simple presence of a multiple conviction case, it lacked the depth to identify the convictions beyond the single most serious (or the principal) conviction.

A report commissioned by the SCEW used CCSS data (and its controls) for a period from 2013 to early 2015 to test the effects of ethnicity, gender, and age relative to sentence outcomes for offences sentenced under the guidelines specific to robbery, theft, and harassment (Chen, Vuk, Kuppuswamy, and Kirsch, 2023). Results showed that, compared to Whites for robbery, Black offenders were less likely to receive custodial sentences, while Asian offenders received shorter custodial sentences. No statistical differences in the likelihood of immediate custody or the length of custody existed for any other ethnic combination for robbery or regarding any of the theft and harassment offences studied. Researchers thus observed ‘no strong nor consistent evidence of sentencing disparities for different ethnic groups, either directly, or through the impact of upward or downward factors’ (Chen et al 2023: 53).

The Chen et al (2023) study also looked at gender, finding that when multiple controls were applied, men were more likely to receive an immediate custodial sentence for robbery and several types of theft offences, but not for harassment; men received longer custodial sentences only for robbery. Researchers concluded that ‘while there might be some evidence of disparate treatment of male and female offenders, gender was a more salient factor in receiving custody than in receiving a longer custodial sentence. Even so, these effects were not consistent across all offences analysed’ (Chen et al 2023: 49). Regarding age, older offenders were more likely to receive custodial sentences for robbery but less likely for one type of theft; older offenders received longer custodial terms for robbery and four types of theft. Otherwise, no other differences based on age were found. Chen et al (2023) offer insight into two other types of demographic factors. Those with a physical or mental illness or a mental disorder or learning disabilities were less likely to receive immediate custody. With controls applied, no associations

[12] To our knowledge these free response options have not yet been analysed.
[13] See for example, Hood (1992).

for any sentencing outcomes were found with respect to the defendants' difficult/deprived background.

Pina-Sánchez and Harris (2020) deployed the CCSS, finding that men were more likely to receive immediate custodial sentences than women for burglary, drug, and assault offences, in models with or without the multiple controls available in the dataset. Hence, for these offences, men are more likely to face imprisonment even when accounting for a host of legally relevant factors. Then for sentence length and focusing on assault cases in one year (2011 - because this was the only year employing a continuous scale for sentence length) and applying controls, results found males received longer prison terms than women.

Another paper to use the CCSS, and to deploy its multiple controls regarding severity and culpability, was specifically addressed to cases sentenced for the highest volume drug crimes under the drug offence guideline (Isaac, 2020). Results indicated that compared to Whites, those in the Black, Asian, and Other ethnic groups were at higher odds of immediate custodial sentences. Asian drug offenders received longer custodial sentences than Whites, but there was no statistical difference in prison length between White and either the Black or Other groups in the drug cases studied. Further results showed that males, compared to females, had a higher likelihood of an immediate custodial sentence and on average received longer custodial sentences. The study modelled age in ordinal rankings, finding that those in the age band of 26 to 50 were more likely to face prison than those younger or older.

Pina-Sánchez et al (2019) created a unique dataset from Crown Court sentencing records for individuals sentenced to prison for violent or sexual offences between 2007 and 2017. The analyses used some controls (e.g. presence of mitigating factors, guilty plea, injury caused, multiple offences, co-defendants) and found that males and older defendants received longer prison terms than females and younger individuals, respectively. This paper did not address or control for ethnicity.

2.4 Northern Ireland

As of 2023, 97% of the prison population in Northern Ireland were males, with the largest age group being 30-39 years (Northern Ireland Statistics and Research Agency, 2023). Statistics on ethnicity within the incarcerated population are not available for Northern Ireland (Shankley and Williams, 2020).

Contemporary studies on sentences in Northern Ireland focusing on demographic characteristics were not located. Still, using unadjusted data from a publicly available dataset showed that men were more likely to receive immediate custodial sentences in Northern Ireland in 2022 (data extracted from a table provided by NI Courts and Tribunals Service Statistics, 2023).

Ireland

There is a scarcity of research on the influence of sociodemographic characteristics on sentencing decisions in Ireland (Healy and Griffin, 2023: Gormley et al 2023), likely because of a lack of systematic data collection efforts (Guilfoyle and Marder, 2021). For example, the Irish Prison Service does not offer an ethnic monitoring dataset and thus statistics on the ethnic representation of the prison population is not readily available (Brandon and O'Connell, 2018).

O'Hara and Rogan (2015) commissioned a specific project to gather data on a subset of cases in Ireland during 2011-2012. Their intent was simply to compare the assignment of community service orders with short-term prison sentences as legislation required courts to consider a community sanction in cases where a short prison sentence of 12 months or less was appropriate. In unadjusted statistics and without controls for ethnicity, results found no gender differences in the use of community service orders versus short-term prison sentences. A minor age difference was detected, whereby defendants receiving community service orders were on average one year younger than those facing short custodial terms (O'Hara and Rogan, 2015).

Other researchers used data from the Irish Prison Service which tracks whether prisoners are of Irish nationality for those sentenced to imprisonment (Brandon and O'Connell, 2018). For years 2015-2017 with unadjusted statistics, Irish nationals received lesser mean sentence lengths for four offence types and greater mean lengths for two offence types than non-Irish nationals, with no statistically significant differences in about two dozen other categories (Brandon and O'Connell, 2018). When controlling for prior custodial sentences, several of the statistical differences were negated; Irish nationals received lesser mean custodial sentences for only two offences. The study reviewed one aspect of gender, finding that non-Irish males received longer custodial terms on average than Irish males for two specific offences and non-Irish females received longer custodial terms on average than Irish females for two other specific offences; otherwise, no gender/nationality differences were reported.

2.5 United States

Demographic differences exist in the proportions of individuals in American prisons serving sentences of more than one year. At the calendar year end 2022, the highest rate of imprisonment was for Black persons: the Black imprisonment rate was 13 times greater than the group of Asian[14], Native Hawaiian, or other Pacific Islander, 5 times higher than White, 2 times higher than Hispanic, and slightly higher than for Native American (Carson, 2023). These disparities, while large, have improved over time: between 2012 to 2022, the incarceration rates for Blacks and Hispanics dropped by 36%, compared to a 23% drop for Whites and a 18%

[14] The categorisation of 'Asians' in the United States' studies is most often a single grouping, which does not recognise subsets of Asian persons that are often utilised in UK datasets with demographic sorting.

decrease for Native Americans. The imprisonment rate of those sentenced to more than one year for men was 14 times higher than the rate for women (Carson, 2023). Still, these are combined statistics across American jurisdictions, with significant variations in demographic rates of imprisonment across federal and state prison systems.

A significant body of disparities research exists for sentencing outcomes in the United States (Spohn, 2017). It is noted that country-wide conclusions are problematic considering the many distinct penal jurisdictions that exist therein with widely varying laws and practices (e.g. federal, states, municipalities, Tribal courts and military). Nonetheless, reviews of the numerous studies, generally conducted on small, geographically limited samples, often find some evidence of harsher sentencing outcomes for Black and Hispanic defendants that are not fully accounted for by criminal history or crime seriousness (see a meta-analysis by Mitchell and MacKenzie, 2004) and lighter sentences for Asian defendants who are not immigration offenders (Hunt, 2023; King and Light, 2019). Literature reviews of studies that focus on intersectionalities highlight that young Black and Hispanic males are often targeted for the harshest sentencing treatment (Lehmann, 2022; Franklin, 2018).

Fewer studies have distinguished the sentencing experiences of Native Americans. An older meta-analysis reviewing race/ethnicity determined there were an insufficient number of studies on Native Americans to support a systematic review (Mitchell and MacKenzie, 2004). A more recent literature review of small samples over time found some evidence, though not consistently, of more punitive consequences for Native Americans (Ulmer and Bradley, 2019). It is noted that harsher sentencing practices for indigenous populations were observed in other countries, with reports blaming sentencing practices in part for Indigenous overrepresentation in the prison systems of Australia (Bagaric, 2016), New Zealand (Pratt, 2017), and Canada (Canadian Department of Justice, 2023; Alberton and Gorey, 2021 (providing a meta-analysis).

Still, some contrary evidence of racial/ethnic disparities in the United States is also present. A more recent meta-analysis grouping together various types of sentencing outcomes (e.g. imprisonment, sentence length, discretionary decision points impacting the sentence) found Blacks and Hispanics were subject to harsher treatment than Whites only for drug crimes (Ferguson and Smith, 2024). As another example, a study using a 2009 dataset of felony case outcomes combining a collection of some of the most populous counties in the United States with individual-level and county-level controls found that Black defendants were more likely to receive custodial prison sentences than Whites, but there was no difference for Hispanic or Other defendants compared to Whites (Nowacki, 2020).

Researchers performing a systematic review concluded that while guidelines-based systems in America have improved consistency in sentence outcomes, disparities regarding race/ethnicity have not been eliminated (Johnson and Lee, 2013).

As for gender disparities, a meta-analysis of United States sentencing research found substantial evidence of harsher sentencing practices for men than for women, though not

universally (Bontrager, Barrick, and Stupi, 2013). These include articles published between 1991-2011 tending to show more severe penalties for men for the likelihood of incarceration and sentence length, including the subset of articles controlling for criminal history, severity, and offence type. Still, favourable treatment of women appeared to decline over time (Bontrager, Barrick, and Stupi, 2013).

Another study of a collection of large counties, with a limited number of individual controls (race, gender, age, offence type, presentence detention, criminal history, and attorney) and county-level controls, found that men were more likely to receive a prison sentence (Nowacki, 2020). Still, when viewing race and gender together, Black men were more likely to receive prison sentences than White men, while there was no statistical difference as to whether Black women and White women received prison sentences (Nowacki, 2020).

Results for age disparities in American sentencing are mixed. A meta-analysis of United States studies produced between 1970-2007 concluded that 40% found a positive association between age and sentence length, 57% found a negative association, with the overall weighted mean effect size being insignificant, meaning no evidence of a relationship between age and sentence length (Wu and Spohn, 2009).

As to the socioeconomic status of defendants, the Ferguson and Smith (2024) meta-analysis of United States sentencing studies found no evidence of a class effect.

2.6 Perceptions of Sentencing Disparities in the United Kingdom

A few investigations of the views of the public and stakeholders on issues of fairness in criminal history outcomes have in recent times been conducted in the United Kingdom. One survey of public perceptions of sentencing by Scottish residents conducted in 2019 found that over half (63%) were fairly or very confident (compared to one-third (35%) that were not very or not at all confident) that Scotland's criminal justice system as a whole is fair to all (Black, et al 2019). Respondents were less likely to be confident in the system's fairness if they either claimed to know a lot or a moderate amount about sentencing or believed that local crime was increasing.

Another recent survey enquiring about sentencing of sex offenders in Scotland indicated that the public generally did not think that the personal circumstances of the defendants, including gender, should be considered in setting the punishment (Biggs et al 2021). The survey did not explicitly enquire about ethnicity as a personal circumstance.

A 2022 survey of adults in England and Wales enquired about their perceptions of criminal justice institutions generally (i.e. not just sentencing). Just over half (53%) of respondents were fairly or very confident that the criminal justice system was fair to different groups (whereas 44% were not very or not confident at all) (Archer, et al 2022). Males and those in the White or Asian

ethnic groups were more likely to be confident in the fairness of the system; no significant differences were detected by age group, socioeconomic status, or educational level.

About 57% of respondents in a 2019-2020 survey in Northern Ireland were fairly or very confident that the criminal justice system as a whole was fair (statistic extracted from data tables provided by Northern Ireland Safe Community Telephone Survey 2021-2022 (2023)).

Interviews with criminal barristers in England and Wales revealed suspicions that judges, who are mostly of White ethnicity, might interpret background information in ways that are sympathetic toward White experiences, yet suspecting representations of minority ethnic cultures indicate higher risk (Veiga, Pina-Sánchez, and Lewis 2023). Interestingly, doctoral theses discussing interviews with sentencing judges in Scotland are generally absent of any self-reflections from them of potential ethnic or other sociodemographic biases (Velasquez, 2018, 2024; Jamieson, 2013).

The SCEW recently commissioned a report that included group discussions with representatives of civil service organisations, sentencers, and defence lawyers (Chen et al 2023). Some of the sentencers (e.g. magistrates and Crown Court justices) expressed their belief that women were treated more favourably. In contrast, the civil service partners tended to believe women faced harsher penalties, but that even if women received lighter sentences it could be explained by mitigating issues such as their lower likelihood of reoffending or recognition of being caregivers (Chen et al 2023). In terms of ethnicity, a number of the discussants thought ethnic differences in sentencing might be explained by (a) the lower likelihood of minorities to plead guilty because of a distrust of the system, (b) the disproportionate impact on minorities from such factors as gang membership, addiction, perceived remorse, and learning disabilities, and (c) the higher possibility of having no legal representation (with the inference this may be more a matter of financial status than ethnicity). Discussants also mentioned that socioeconomic status could impact how individuals might otherwise benefit from mitigating factors, such as a lack of financial resources undermining the ability to do charity work or to comply with supervisory requirements.

2.7 Limitations of the Comparative Analysis

International comparisons of sentencing practices are potentially enlightening, though it is nonetheless necessary to emphasise the complexity of drawing broad conclusions or generalising findings, as highlighted by Gormley et al (2022). Variability can arise from many factors, including these examples:

- Despite sharing core sentencing philosophies among common law jurisdictions (i.e. retribution, deterrence, public safety, and rehabilitation), there can be variability in how these principles are prioritised or applied, owing to statutory or cultural differences;

- Inconsistencies in legal definitions and concepts;
- Discrepancies in how data is collected, recorded, maintained, and reported;
- Disparate perspectives of what can or should constitute aggravating or mitigating factors;
- Diverging views on which direction specific circumstances should lean towards (e.g. mental impairment, intellectual disability, youth, or substance use), such as mitigating versus aggravating culpability;
- Distinct practices on how and when young defendants are adjudicated in the adult system;
- Mixes in the range of penalties legally available (e.g. warnings, shame sentences, fines, community orders, suspended sentences, immediate imprisonment, life sentences, or the death penalty);
- Variations in incentives offered for early guilty pleas or for cooperating with law enforcement;
- Differences in the amount of discretion lawfully allowed to sentencers;
- Dissimilar impacts of available options for early release on initial sentencing outcomes;
- The presence of statutory or guideline systems that may be rigid or flexible, including the use or absence of mandatory minimum sentences;
- Requirements, or lack thereof, for sentencers to provide detailed explanations for their sentencing decisions;
- The potential inclusion of sentencers with little or no legal training;
- The legal standard and degree of flexibility in appealing sentences or having them modified or substituted by appellate authorities.
- As a methodological point, much of the research on disparities that explores custodial decisions is reported using odds ratios. Odds ratios should not be used for comparisons across studies since they are contingent on the baseline rate (i.e. the custody rate for the reference category, normally whites). Meta-analyses of sentencing that ignore or do not address this problem are, therefore, severely limited.

Each of the foregoing aspects can influence sentencing outcomes, highlighting the diversity and complexity of judicial systems across multiple jurisdictions. Furthermore, countries may categorise ethnicity and race differently, adding another layer of complexity to comparative research on demographic traits (Tonry, 1997). For example, the Hispanic ethnic category, a mainstay in disparities research in the United States, lacks a direct counterpart in almost any other jurisdiction. Or, Scotland's recognition of a White ethnic minority group distinguished from Whites generally is a classification not commonly found in other Western nations.

2.8 Conclusion

Certain of the studies mentioned in this chapter also face additional limitations that hinder the ability to draw definitive conclusions. Analysing sentencing disparities across the entire population can mask more nuanced differences that might emerge if the research were to concentrate on specific subsets of the population. These subsets could be defined by various factors, including the primary category of offence (such as drug possession, drug trafficking, violent crimes, or major fraud), the severity of the offence, the existence or absence of multiple convictions of similarly serious offences, sociodemographic groupings, or the type of sentence given. Moreover, research that fails to explore intersectionality may overlook crucial variations. For instance, a study might not capture the distinct experiences of young men from ethnic minority backgrounds, thereby missing critical insights into how sentencing outcomes differ across intersecting identities and circumstances.

As well, it may be unfair to make assumptions about placing responsibility for seeming disparities in penalties primarily on the sentencing stage when one cannot entirely isolate the effects to that particular point in time. The criminal justice process is a multi-step journey involving many decision points by various decision-makers that can impact the resulting outcomes, such as resulting in cumulative disadvantage (Mitchell, 2018). This is further complicated by difficulties in properly ‘marrying-up’ data from different agencies which may use different case-counting rules. Thus, discretionary and nondiscretionary choices at various stages might have varying degrees of influence on the ultimate sentence (Veiga, Pina-Sánchez, and Lewis, 2023). As illustrations, actions may be influenced by biases ranging from whether the victim formally reports the crime in the first instance, police make an arrest, prosecutors pursue the case, the person pleads guilty or a jury finds guilt, and probation officials make specific recommendations (Sporer and Goodman-Delahunty, 2009).

These differences between cases cannot be assumed to be evenly distributed among cases involving different sociodemographic groups and so may well skew results. The limitations of official data, collected primarily for the operational purposes of processing individuals and cases, may mean that the apparent absence or presence of unwarranted disparities cannot, at least as things currently stand, be fully described nor explained.

Chapter 3: Assessing Official Datasets Relevant to Sentencing

3.1 Introduction

In order to explore questions around sentencing disparities, it is necessary to analyse data about sentencing practices. One potential source of this information is official data such as that routinely published by the Scottish Government. Indeed, there are several organisations that collect large volumes of administrative data relevant to sentencing in Scotland. Some notable organisations include: Police Scotland, the Crown Office and Procurator Fiscal Service (COPFS), the Scottish Prison Service (SPS), local authorities, and the Scottish Courts and Tribunals Service (SCTS). Each agency produces and publishes data that reflects its own particular functions and responsibilities, as is entirely reasonable. However, this focus on operational needs means that it can be difficult to derive meaningful conclusions about sentencing patterns from an organisation's administrative data. From the perspective of sentencing, the fragmented manner in which each organisation produces data in different ways can make it challenging, if not practically impossible, to engage in 'follow-through' of the progress of samples of specific cases from the point of initial reporting or detection to final disposal and the implementation of the sentence.[15]

The net effect of this fragmentation of data from different administrative sources is that there are major barriers to being able to understand and describe patterns of sentencing practices according to key conditions. For example, COPFS publishes charge-level statistics in publications such as 'Domestic Abuse Charges reported to COPFS' and 'Hate Crime in Scotland'. Yet, since Criminal Proceedings statistics measure the main charge in a case,[16] this means that these figures are not directly comparable.

Similarly, the fragmentation of data means there is little ability to enable plea status (a key factor which has been the subject of much legal, public and policy debate)[17] to be linked to final sentence outcomes. The Management Information and Analysis Team (MIAT) within SCTS provide monthly information derived from the Criminal Operations digital case management system (COP2).[18] If we look at routine (e.g. quarterly and annual) publications, we see that publicly available information is provided in aggregate form. As such, it is of limited use in terms

[15] See further Appendix A.
[16] Additionally, the Scottish Government's statistics on criminal proceedings report on the convictions and sentencing of accused persons. This is different from the recorded crime statistics, which count crimes and offences at the time that they came to the attention of Police Scotland.
[17] Gormley, McPherson and Tata (2020); Gormley, Roberts and Tata (2025). See Section 196 of the Criminal Procedure (Scotland) Act 1995. This provision has been discussed in a series of cases such as *Du Plooy v HMA* 2005 1 JC 1, *Spence v HMA* [2007] HCJAC 64, and *Gemmell v HMA* 2012 JC 223.
[18] 'Digital Strategy – 2018-2023', 22 June 2018, 7

of facilitating detailed analyses. While there is some frequency information on pleas, this information does not link to sentences passed and all bulletins carry the important proviso that:

> 'The statistics in this bulletin do not have information relating to accused persons in terms of what they were charged with or their resulting conviction or sentence as there are already well-established National Statistics on these aspects of criminal justice.'[19]

This fragmentation means that the impact of key policies, such as the impact of pleas on sentencing, cannot be adequately measured by available official data.[20]

Despite these limitations, official data from different agencies may have the potential to be interesting and have important implications for the understanding of sentencing practices. Additionally, given that large-scale, comprehensive data collection can be labour-intensive, there is a perennial intrinsic appeal to the possibility that existing datasets could be utilised for purposes such as understanding sentencing or sentencing trends (such as predicting prison populations).

Consequently, this project sought to gain access to official data so that it could be analysed to explore questions relevant to warranted and unwarranted disparities in sentencing with the intent to shed new light on these important matters. However, the reality is that there are serious restrictions that prevent such data from being a viable option for the SSC to make use of in order to provide the data it needs on sentencing.

Therefore, this chapter details some of the key limitations of official data in Scotland and how this impacted the ability of the project to explore questions about warranted and unwarranted disparities.

3.2 Limitations of Official Publications

Publications based on official data can be relevant to questions about sentencing. For example, the Criminal Proceedings in Scotland publications[21] and reconviction publications[22] have data drawn from the Criminal Case History System for Scotland (CHS). Likewise, the Recorded Crime in Scotland publications[23] 'are based on data which Police Scotland extract from their data repository (called the Source for Evidence Based Policing (SEBP)) and submit to the

[19] 'SCTS Official Published Statistics'
[20] For an analysis of *some* of the challenges to understanding the operation of not guilty and guilty pleas in sentencing, see Gormley and Tata (2020); Gormley, McPherson and Tata (2020, 2022); Gormley, Roberts and Tata (2024).
[21] Available online: https://www.gov.scot/collections/criminal-proceedings-in-scotland/.
[22] Available online: https://www.gov.scot/publications/reconviction-rates-scotland-2018-19-offender-cohort/pages/35/.
[23] Available online: https://www.gov.scot/collections/recorded-crime-in-scotland/.

Scottish Government.'[24] Similarly, some basic data exists on matters such as 'index sentence length' (Scottish Government, 2024).[25] Such data can be informative about matters relevant to sentencing policy and is easy to access.

However, while such publications can be relevant to understanding sentencing, alone, they are insufficient. Much more information is needed for guideline development and monitoring – including exploring questions around warranted and unwarranted disparity. Additionally, the utmost caution is required when relying on bald statistics without relevant controls that enable like-for-like comparisons for the purposes of exploring questions around potential disparities. Without the required details and controls, more meaningful insights into sentencing are simply not possible.

In part, these limitations are unsurprising. While matters linked to sentencing are of interest to certain audiences (such as the prison population for the purposes of predicting future capacity), understanding sentencing is not typically the focus of official data. Thus, routine publications do not offer, (and are not intended to offer), the in-depth and nuanced insights into sentencing that would be needed to explore questions around potential sentencing disparities. For instance, the Scottish Government use crime groups in publications. The Recorded Crime in Scotland publications include:

- Total Crime and Offences;
- Total Crime;
- Non-sexual crimes of violence;
- Sexual crimes;
- Crimes of dishonesty;
- Damage and reckless behaviour;
- Crimes against society;
- Total Offences;
- Antisocial offences;
- Miscellaneous offences;
- Road traffic offences.

These categories are not legal categories as such. Classifying offences without considering features of the offending nor the offender (e.g. harms caused to victims, mitigating and aggravating factors, etc) makes it extremely difficult, from the perspective of sentencing, to compare similar real-world cases.

[24] (*Recorded Crime in Scotland, 2020-2021*, 2021, para. 7.1)
[25] Note that Page 54 states that for 'a comparison of index sentence length and overall sentence length is provided in the Technical Manual for 2013-14, the latest year when the previous Official Statistics publication and cellWise data were both available.' This data is over a decade old at the time of writing.

Furthermore, the Scottish Government publications group criminal behaviours according to whether it is deemed to be a 'crime' or an 'offence'. There is no distinction in Scots criminal law between 'offences' and 'crimes'. Neither is such a distinction commonly made in policy or academic literature. Instead, this is a distinction, which dates back to the 1920s, for the purposes of publications by the Scottish Government and its forebears:

> 'All criminal conduct, whether classed as a crime or an offence, is serious and the separation of crime and offences statistics is a long-running practice, to facilitate understanding of the statistics. However, this distinction is solely for statistical purposes and ensures that the figures are consistent and keep track of trends over time.
>
> Information on the classification of crimes and offences and their split for statistical purposes can be found alongside the Recorded Crime in Scotland statistical bulletin. The separation of crime and offences statistics has been in place since the 1920's and the statistics have been presented in the bulletin in the same format since 1983.'[26]

It is a distinction, which because of its longevity, might be said to offer a constancy. On the other hand, it is an abstraction with little relevance to real-world criminal proceedings. Further, the long-standing distinction drawn by governmental analysis between what it deems to be the more serious group of 'crimes' and the less serious group of 'offences' might seem rather peculiar. For example, alongside murder, rape and robbery in the more serious grouping of 'crimes' appear 'shop-lifting' and 'possession of drugs'. While the deemed less serious category of 'offences' includes 'urinating etc', 'drunkenness and other disorderly conduct', there is also 'dangerous and careless driving' and 'driving under the influence' (Scottish Government 2003: Appendix A).

Therefore, this project could not progress simply on the basis of routinely published data derived from administrative sources. We also considered one-off publications for this report, and the 'experimental' (i.e. first of its kind) statistical analysis produced by the Scottish Government (2023a) – which relied upon a dataset compiled from the Scottish Courts and Tribunal Service – analysed in Chapter 2. However, this was unable to provide meaningful insight into sentencing practice capable of answering questions related to warranted/unwarranted disparity in sentencing.

3.3 Limitations of Official Data

For present purposes, official publications have limits due to their main focus not typically being on sentencing. However, as the statistical analysis produced by the Scottish Government (2023a) shows, official publications also have limits due to the data that underlies them. Thus, while there might be better ways (theoretically) to use official data, there will be inherent limits to

[26] Email correspondence with Justice Analytical Services.

the insights official data can provide Scotland (see Chapter 4). It is important to set out these limitations as they will impact the ability of future research to produce new insights based on official data.

The general limitation of official data in Scotland is that it is often held in databases that are not designed with a focus on collecting and analysing sentencing data. That these systems are not focused on sentencing has at least two (potentially related) consequences for this project:

(1) Key sentencing factors are not sufficiently covered by official data;

(2) Extracting official data to make it usable is difficult.

3.3.1 The Omission of Key Sentencing Factors

The first consequence of official datasets not focusing on sentencing is that they lack key variables of interest. Data is collected by organisations to meet their own operational needs and obligations. For example, as our discussions with SCTS show, in summary cases, SCTS only audio records sentencing (a potential data source) in summary sexual offence cases (to pass onto Justice Social Work) and in cases where the offender pleads guilty. Likewise, while data on total effective sentences may be something SCTS could, in principle, provide the SSC in specific individual cases, since the details may only be contained within case notes it does not appear to be practically amenable to data analysis. Similarly, data on sentence reductions under section 196 of the Criminal Procedure (Scotland) Act 1995 may be recorded in individual cases, but it does not appear to be practically accessible for analytical purposes.[27] This approach to collecting and storing data may fulfil operational needs for SCTS, (e.g. recording some summary cases where there was a guilty plea), but not in cases where there was a conviction following a trial. This is significant given that plea status may be an especially important issue in terms of potential sentencing disparities (Gormley, et al 2020; Gormley, Roberts and Tata 2025; Gormley, McPherson and Tata 2020).

More broadly, the lack of accessible data on matters such as single-conviction as opposed to multiple-conviction cases and criminal offending history is hugely problematic given that these can be extremely influential at sentencing (Scottish Sentencing Council, 2021). Indeed, it almost seems self-evident to say that, other features being equal, multiple convictions may justify a more serious sentence than a single conviction.[28] Yet, one of the most severe limitations of official data, (also an issue in other countries), is that it tends to deal poorly with providing a meaningful distinction between single-conviction cases and those involving more than one conviction. For example, all else being equal, a case involving ten convictions of robbery is likely to be seen as more serious than a case involving just one robbery conviction. Yet, official data in

[27] See further, Gormley and Tata (2025); Gormley 2025, 2024.
[28] Although this is not to say that, all else being equal, two convictions will result in double the sentence of a single conviction.

Scotland, (and indeed in other jurisdictions), is remarkably poor at dealing with such a basic distinction (Gormley et al 2023; Tata 1997; 2020). Instead, analysis of available data tends to seek to focus upon a 'principal offence' (i.e. a single offence) and detail on other offences is limited (either absent entirely or lacking in specificity).[29] This is an issue which can thwart attempts to provide a meaningful analysis of sentencing.[30] Similarly, other key variables are often missing, such as the relative seriousness of criminal history (e.g. previous convictions analogous to the case).

In terms of exactly what is missing from Scottish data, this remains a troublingly difficult question to answer. Because data is not collected or stored with sentencing analyses in mind, organisations may struggle to provide clear answers about the data they hold and, indeed, may be unclear about what data they record and what is practicably accessible for investigation. Thus, even establishing what is and is not missing from official data is an arduous task. For example, it appears that SCTS does not have a list or document setting out the variables recorded in their systems. In the absence of a definitive document which can be shared, the main way to know what factors are recorded in the data and what *might*[31] be accessible is to ask SCTS 'shot in the dark' questions about what sentencing variables may be available and for SCTS to then investigate and report back. This process tends to be slow, cumbersome, and easily leads to potential mutual misunderstandings.

Accordingly, a detailed investigation had to be carried out to determine what information relevant to legitimate sentencing considerations is recorded and held by official data, whether or not it was viable to access, and in what form (if at all) it could be made available so as to enable analysis. This investigation was focused on well-established questions of fundamental importance to sentencing, including for example, the ability simultaneously to control for:

- The relative seriousness of offending in different cases;
- The ability to distinguish between single and multi-conviction cases (especially in cases where there is more than one conviction of similar seriousness);
- Previous criminal history (including, for example, previous analogous convictions);
- The ability to control for plea ('guilty' or 'not guilty');

[29] The 'principal offence' may often (but not always) be the most serious offence.
[30] If possible, for an analysis, it would be ideal to have the sentence imposed for each offence and whether it is consecutive, concurrent, or *in cumulo*. However, if that is not available and only the final sentence for a notional principal offences is available, at the very least an analysis would need to know whether multiple offences were considered by the judge. It would also be helpful to have a measure on whether the principle of totality was applied.
[31] There are issues around what data, even if held, is within the ability of SCTS to access and provide to the SSC in a manner that facilitates analysis. For example, as above, some data may exist in case notes and be prohibitively time-consuming to extract.

- Offender characteristics (including family circumstances such as primary caring responsibilities of children, mental and physical health, income, age);
- The backdating of custodial sentences to cover time on remand; and
- Different conditions of a community-based sentence imposed by the courts.

In all of these (and other) instances, information was found either not to be collected (and thereby missing from official data), or practically inaccessible for analysis purposes. This limits the ability to conduct an in-depth, controlled analysis of sentencing patterns according to key criteria of interest.

3.3.2 Access Limitations

A second consequence of organisations not focusing on collecting and analysing sentencing data is that accessing data held by organisations can be a tremendous challenge. For instance, COPFS notes:

> 'The COPFS case management database is a live, operating database. It is designed to meet our business needs in relation to the processing of criminal cases, and the information within it is structured accordingly. We do have a separate statistical database and hold only operational data needed for business purposes.' (McManus 2024: 28-29)

As such, COPFS and other organisations may not readily be able to provide data that will assist in the development and evaluation of sentencing guidelines, nor the description of sentencing patterns.

To put it simply, the evidence of whether there are or are not unwarranted disparities is extremely limited because the available official data does not adequately enable a proper analysis which controls for the most important criteria.

Of course, data issues such as these do not mean that no insights are possible from official data (see Chapter 4). However, these issues do mean that the insights available are limited (sometimes to the point where few meaningful conclusions can be drawn). In theory, a fuller picture may be obtained by drawing on data from multiple organisations. Yet, for a variety of reasons, it appears clear that it is (especially for a project of this size) impractical to link different datasets from the various organisations in Scotland at this time, or even in the medium-term future to yield in-depth analyses of sentencing patterns according to a range of key criteria.

In part, the lack of a cohesive picture is a consequence of information coming from different databases that cannot be linked due to different recording practices (including case counting rules and different ways of classifying offences):

> ‘Each of the main criminal justice bodies measures activity differently. Police Scotland count standard prosecution reports (SPRs), COPFS count the number of cases and accused people, and SCS [now SCTS] count numbers of cases and accused appearances. There is no consistent way to identify distinct individuals (be it as witnesses, victims or accused) across the whole sheriff court system.’ (Audit Scotland 2015: 12).

Therefore, in this research we selected the organisation most likely to hold relevant data and focused our efforts on that: SCTS.

3.3.3 SCTS Data

Sentencing debates in England and Wales are able to draw on official datasets such as those held by the MoJ. This data, while not without limitations (especially compared to the Crown Court Sentencing Survey), can then be used to inform the development and monitoring of SCEW guidelines. In Scotland the closest equivalent might be considered to be SCTS. Section 10 of the Criminal Justice and Licensing (Scotland) Act 2010, states:

> ‘(1) The Scottish Court Service must provide the Council with such information relating to the sentences imposed by courts as the Council may reasonably require for the purposes of its functions.
>
> (2) The information must be provided in such form and by such means as the Council may require.
>
> (3) The Council must from time to time publish information about the sentences imposed by courts.’

Thus, the research team requested that the SSC secretariat explore what data SCTS might be able to provide to the SSC in order to facilitate an analysis of warranted and unwarranted disparities in sentencing.[32] Regrettably, despite persistent efforts, it proved impossible, in practical terms, to gain access to SCTS data for analysis.

Nevertheless, it was possible to explore the nature of what data relevant to describing patterns of sentencing is and is not collected, and what is and is not electronically available, in principle, from SCTS.[33] In itself, knowing this is valuable to the future direction of sentencing research. Future research projects, including those conducted or commissioned by SSC, need to know what data is held and could, in principle, be made available. This can then allow exploration of

[32] We thank the SSC secretariat for their significant efforts in this area in light of the significant challenges that emerged over the course of several months.
[33] This is aside from various hurdles, such as the need for a data sharing agreement. By this stage in the project, it was made clear that since there was no longer time to conduct an analysis of SCTS data, a data-sharing agreement would not be required.

the of any meaningful knowledge about patterns of sentencing given different criteria. Thus, the goal was to explore with the SSC and SCTS:

(1) What data is held?

(2) What data could reasonably be provided for a potential future analysis of warranted and unwarranted disparity?

In written correspondence, we asked the Management Information and Analysis Team (MIAT) of SCTS a series of specific questions about the collection and availability of key information relevant to the attempt to describe patterns of sentencing. We then arranged an online meeting between the research team, SSC, and SCTS. Unfortunately, no member of MIAT was able to attend the meeting, and so our questions (listed below) were discussed with another SCTS official who in turn consulted MIAT.[34] Written responses were provided subsequently. Here we outline answers to the questions raised.[35]

1. Timescale.

What are the limits to numbers and time periods SCTS can provide data on? Are there technical or practical/resource limits?

Following consultation with the Management Information and Analysis Team (MIAT) of SCTS, we were informed via the SCTS official that COP2 was established in 2006 so records could be obtained from then to date. However, we were informed that this would have significant practical and resource constraints, which would require to be considered in full prior to any data provision being undertaken. In relation to physical case papers, it was reported that these are retained locally at courts for a period of 25 years and thereafter forwarded to the National Records of Scotland for future preservation. We were informed that summary cases are retained locally at courts for a period of 10 years and then subsequently destroyed.

2. Offences.

How are offences recorded in SCTS in data? For example, is it based on offence codes or something else? Does SCTS have a list of charge codes if relevant?

Following consultation with MIAT, it was reported by the SCTS official that offences are recorded by charge code. The charge codes are allocated to a case by the Crown Office and Procurator Fiscals Service (COPFS) - SCTS has no say in what someone is being charged under. As charge codes are allocated by COPFS, it was recommend that COPFS be contacted to see if a document exists and whether it can be shared.

---

[34] The note of that meeting was submitted to and agreed by the SCTS representative in consultation with the Management Information and Analysis Team of SCTS.
[35] This was described as its 'final' response.

3. Multi-Conviction Cases (i.e. individual accused convicted of more than one offence).

How, if at all, are multi-conviction cases reflected in SCTS data? Is there a principal offence approach or is each offence noted? If the latter, then how are sentences recorded (e.g. is there detail on matters such as whether sentences are consecutive, concurrent, cumulo, etc)?

Following consultation with MIAT, it was reported via the SCTS official that each charge code is noted, each charge can receive one or more disposals, and that there is no principal offence approach. It was reported that details on consecutive, concurrent, cumulo, etc sentences are contained within case notes and that these are not disclosable due to sensitive/personal information being contained within these.

4. Aggravating and Mitigating Factors.

How are these reflected in SCTS data? Is information about significant health issues (e.g. physical or mental illness) recorded, or would this not be typical?

Following consultation with MIAT, we were informed by the SCTS official that as it is recorded in the original minute, the following information is only available via case papers or by audio recordings (where available), and not by electronic records:

- Mental/physical health
- Whether accused is main carer for a child
- Employment status
- Income, decile or SMID
- Previous convictions

We were informed that age is recorded and can be provided by point of case registration, point of sentence or point of disposal, and that this information has already been provided to the SSC on occasion when requested by range.

We were further informed that audio recordings are available in solemn cases and in summary cases where the offence is a sexual offence.[36]

Following consultation with MIAT, we were informed by the SCTS official that data on previous convictions is not something that is able to be provided from electronic records.

5. Plea Status.

How are not guilty/guilty pleas reflected in the data? E.g. What data does SCTS have about plea (guilty vs not guilty) to charges where there was a conviction (whether by guilty plea or found guilty by the court)? Is the timing of the plea held in SCTS data? Is there any data on a discount/reduction under section 196?

[36] We were informed that is to allow provision of the COPFS narrative to Justice Social Work.

Following consultation with MIAT, the SCTS official informed us that pleas are recorded at charge level by verdict date. These can be co-related to hearing type.

We were further informed that it might be possible for data to be provided on discount/reduction under section 196, but that this would require investigation.

6. Race/Ethnicity of the Sentenced Person.

Is data on the race/ethnicity of the person sentenced available and, if so, how is this data generated?

Following consultation with MIAT, we were informed by the SCTS official that this is sometimes recorded by the police at the outset, but is variable.

7. Sex.

Is data on the sex (e.g. male or female) of the person sentenced available and, if so, how is this data generated?

Following consultation with MIAT, we were informed by the SCTS official that this is sometimes recorded by the police at the outset, but is variable.

8. Sentence Disposal Type.

How does SCTS data reflect cases involving more than one type of sentence? (e.g. a Community Payback Order (CPO) and fine). How is this categorised in SCTS data?

Following consultation with MIAT, we were informed by the SCTS official that SCTS does not categorise the data, but gives the disposal and that this will show as multiple records: e.g. fine, victim surcharge, compensation, and CPO at charge level.

9. Sentence quantum (e.g. in terms of months, hours, pounds etc).

How is sentence quantum recorded? In multi-conviction cases, is the total effective sentence extractable?

Following consultation with MIAT, we were informed by the SCTS official that the total effective sentence is possible to know. However, the details are contained within case notes, and these are not disclosable due to sensitive/personal information being contained within this area.

10. Backdating.

How is backdating for time spent on remand recorded and presented in the data? Is it extractable electronically and in a way which can be cross-referred to other key sentencing information?

Following consultation with MIAT, we were informed by the SCTS official that this is not available in the data.

11. Deferred Sentences.

Are deferred sentenced recorded in and extractable from SCTS data in a way which can be cross-referred to other key sentencing information?

Following consultation with MIAT, we were informed by the SCTS official that there is a diet type for deferred sentence which could be extracted. However, MIAT was unsure how this could be cross-referenced or used.

12. List of variables

Is there a document (e.g. CSV, Excel, or word format) that provides a list of what variables are held in the systems (e.g. COP2) used by SCTS? If there is such a document, is there information on how these variables are held (e.g. a variable held may be 'date of birth' and this may be in a format of DD/MM/YY).

Following consultation with MIAT, we were informed by the SCTS official that this is not available.

3.4 Conclusion

Despite extensions to this project, it proved impossible to obtain SCTS data to analyse the presence of warranted or unwarranted disparities in sentencing. However, it was possible to make significant steps in establishing: what data is and is not recorded; whether it is or is not held electronically; and whether or not can and cannot be made available to SSC.

It is apparent that SSC faces profound challenges in using official data to meet its statutory objectives. What is striking from SCTS' response is just how limited the information is and how difficult it may be to make available in a form which can be subject to statistical analysis. Critical information about case seriousness is not available (except through a trawl of case papers). Information about single as opposed to multi-conviction cases, previous convictions, backdating of a sentence to cover time on remand, and vital features of the sentenced person appears to be unavailable.

That such information may not be practically available (or at all) in electronic form should not, perhaps, be surprising. The operational recording practices of official bodies like SCTS, COPFS, the Scottish Prison Service have, historically, been geared to recording information useful to the processing of individuals. A further problem is that each agency uses its own recording practices (e.g. case counting rules). As such, it can be difficult (if not practically impossible) to combine

data from different agencies, thus preventing the ability to analyse case flows through the criminal justice system. Of course, to say that official bodies' recording practices have not been primarily geared towards providing meaningful insights into sentencing practice should not be understood as a criticism of those bodies. It does, however, present a fundamental challenge for the SSC and others who need meaningful and in-depth information about sentencing patterns so as to develop informed policy.

While some jurisdictions have made some advances in the use of sentencing data from official sources, the relative lack of high-quality data is not a problem which is unique to Scotland. Although vital for operational purposes, the available official statistical data is often less than adequate for the specific purpose of providing a meaningful picture of patterns of sentencing (Gormley et al 2023).

Nevertheless, while we should not lose sight of the profound limitations of official data, if it was made available it might be that by using advanced statistical techniques a more penetrating picture of the patterns of sentencing can be developed. It is this aspiration which the next chapter explores.

Chapter 4: The Use of Statistical Analysis of Official Data

4.1 Introduction

This chapter examines existing sources of official data from administrative sources to assess whether and how they could be useful in determining the existence or absence and reasons for unwarranted disparities in sentencing. In keeping with emphasis on the question of unwarranted sociodemographic disparities, the chapter does not examine the question of inter-judge disparities, though much of what is discussed in this chapter about statistical techniques can also be applied to that question and vice versa.

One would hope that official data collected by administrative agencies for operational purposes should be able to provide a relatively accurate description of the sentence outcome. After all, this would allow for a wholly comprehensive picture of sentencing practices, and it would do so without having to expend any further effort in collecting sentencing information. Despite the many serious flaws in the ability of official data in Scotland identified in Chapter 3, it may, nonetheless, be possible with the use of advanced statistical techniques, to use the existing data to help provide more revealing information. This could help to inform further research (assessed in Chapter 5) using a range of different techniques, whether by large census exercises or by specific in-depth qualitative research. This chapter, therefore, assesses the suitability of different statistical techniques to support the analysis of official data.

Official administrative data refers to datasets which are made available to researchers derived from information collected by courts or governmental bodies for their own operational purposes. These datasets tend to provide certain basic information about the sentence outcome, but often they only record a few offender and case characteristics, such as the offence type, age, and gender. Examples of administrative sentencing data are the new Magistrates and Crown Court Datasets made available by the Data First project for the jurisdiction of England and Wales (Ministry of Justice, 2022), or the dataset compiled to undertake the Scottish Government (2023a) report on ethnic disparities in sentencing, and which is currently only available to government researchers in Scotland.[37]

4.2. Isolating Unwarranted Disparities

When we study sentence disparities, we adopt the view that defendants find some outcomes more preferable than others. However, given the wide range of disposal types used in modern criminal courts, sentence outcomes can be expressed in different units (e.g. fines measured in pounds, and custodial sentences in days), which hinders their operationalisation. A variety of methods have been used to integrate disposal types based on different measurement units into

[37] This is discussed in greater depth in Chapter 2.

a single scale of sentence severity (Leclerc and Tremblay 2016; Pina-Sánchez et al 2019). These methods, however, invoke a series of assumptions and subjective choices. Hence, for reasons of parsimony, the preferred option is to rely on specific disposal types that allow for objective comparisons of severity, the most common ones being whether an immediate custodial sentence is imposed and its length.[38]

After deciding on how the sentence outcome will be operationalised, the key methodological challenge to be addressed is finding a strategy to isolate unwarranted disparities in sentence severity (Pina-Sánchez and Linacre 2016). This is because differences in sentence outcomes across say, offenders from different demographic groups, or more broadly, changes in the overall variability in a given sentence outcome ($var(O)$), are not necessarily indicative of unwarranted disparities. Such patterns might be reflecting differences in sentence severity associated with legitimate legal factors, such as the seriousness of the offence, or the culpability of the offender ($var(L)$). Unwarranted disparities would instead represent all other forms of illegitimate variability ($var(I)$) in the sentence outcome that cannot be explained by the legal factors defining the case. Hence, we can see differences in sentencing to be composed of two key elements, differences due to legitimate and illegitimate factors, yet our interest lies on identifying the latter:

$$var(O) = var(L) + var(I)$$

Broadly, we can distinguish two main methodological approaches used to isolate $var(I)$ from $var(L)$, random allocation and statistical controls.

4.2.1. Random Allocation

Random allocation was first exploited for the study of sentencing disparities over 80 years ago (Gaudet et al 1933). This method relies on practices followed by certain criminal courts in the US, where cases are assigned randomly to judges to try to prevent so-called 'judge shopping'. This random process means that over many cases, judges operating in a given court should, through the processes of chance, in a large sample typically be assigned a similar mix of cases. Hence, any systematic differences in sentence severity across judges might be indicative of unwarranted disparities.

This is the most parsimonious and unbiased approach to explore unwarranted disparities in sentencing. However, it suffers from two key limitations. Conceptually, inter-judge disparities are only a subset of all possible unwarranted disparities. They do not tell us anything about court or regional disparities, nor are they particularly informative of inter-judge disparities (e.g. Eren and

[38] This analytical choice is not without its limitations, as it involves a loss of information, which can lead to selection bias. We will get back to this point in Section 4.3 'Other Methodological Challenges.'

Mocan 2008) or possible discriminatory practices. Practically, random allocation of cases is not a common procedure across most jurisdictions, and Scotland is no exception.

4.2.2. Statistical Controls

When cases are not randomly allocated, researchers need to seek to isolate unwarranted disparities by 'adjusting' for legal factors using statistical controls. There is no commonly agreed list of legal factors, but most researchers point to all the legitimate offence and offender characteristics that sentencers are required or permitted to contemplate in deciding the most appropriate sentence. Using these controls researchers seek to create comparable groups of cases, so any remaining differences in sentence outcome provides a better approximation of unwarranted disparities. For example, instead of comparing custodial sentence length between Black and White drug traffickers, we can use information about the offence (e.g. the type of drug, the quantity traded, etc) and the offender (e.g. their role, personal mitigating factors, etc) to more closely approximate 'like with like' comparisons.

To adjust for legal factors, researchers have relied on two main analytical approaches: matching methods (Bales and Piquero, 2012; Kemp and Varona, 2023) and regression modelling (Zatz, 1987). The former aims to explore combinations of legal factors to find groups as similar as possible so they can be compared directly. This approach provides more intuitive findings and eliminates the type of parametric assumptions used in regression modelling. Regression modelling seeks to condition for (i.e. strip away) any variability in sentence severity that can be attributed to the legal factors present in the cases explored. To differentiate between explained and unexplained variability (i.e. residuals) in the outcome, regression modelling invokes a series of parametric assumptions indicating how each of those parts is expected to be distributed and related to the outcome.

This involves a less parsimonious approach than matching methods, but in exchange it offers additional flexibility to explore different forms of unwarranted disparities. For example, beyond comparing systematic differences in sentence severity across demographic groups after conditioning for legal factors, we can explore questions of: i) intersectional disparities, which take place when various individual sociodemographic characteristics combine to form specific social strata at risk of being particularly marginalised; ii) identifying the level where disparities arise, assessing the relative contribution of case, offender, judge, court, or even regional factors; and iii) estimating the overall level of consistency, which the United States Sentencing Commission (Hofer et al 2004) and the SCEW (Isaac et al 2018), have operationalised as the ratio of unexplained over explained variability, or $var(I)/var(L)$. As could be expected, regression modelling is the most common approach used in the literature; it is the method used in the first attempt in Scotland to conduct an analysis of official administrative data to illustrate *potential* ethnic disparities in sentencing undertaken in Scotland (Scottish Government 2023a).

4.2.3. Confounding Bias

Regardless of whether matching or regression methods are used, the key limitation of studies of official data relying on statistical controls stems from their inability to adjust for relevant legal factors. Assuming that such a finite set of legal factors exist and can be agreed upon, it is often impossible for practical reasons to control for all of them. This can happen when the sample size is not big enough relative to the number of legal factors to be controlled for, or, more commonly, because, as Chapter 3 explained, many of those legal factors are simply not recorded in the dataset.

Two implications follow. First, the overall extent of unwarranted disparities will be overestimated since some of the unexplained variability in sentence severity will be due to legitimate differences across cases. Secondly, if the unobserved legal factors are distributed unevenly across demographic groups, then estimates of discrimination in sentencing will also be biased. This is known as 'confounding bias' and it represents the most commonly acknowledged methodological problem in the exploration of sentencing disparities using administrative official data (Halevy, 1995; Wilbanks, 1987). For example, Pina-Sánchez et al (2019) demonstrated how offenders with a Muslim sounding name appear to receive longer custodial sentences in the Crown Court, but that difference becomes statistically non-significant after controlling for the specific offence type.

Engen and Gainey (2000) and Mitchell (2005) have also demonstrated how confounding bias could still be present when the quality of controls is inadequate. In the context of American jurisdictions operating grid-based sentencing guidelines, whereby offence seriousness and offender criminal history lead to a recommended sentence, Engen and Gainey (2000) showed how ethnic disparities become statistically significant after controlling not just offence seriousness and criminal history, but also for the recommended sentence derived from these two legal factors. Based on a meta-analysis of studies from the US, Mitchell (2005) demonstrated how when using broad offence groups (e.g. property offences) as opposed to more specific offence types (e.g. domestic burglary, fraud, etc.), or unduly dichotomising continuous legal factors (e.g. noting whether the offender has previous convictions, rather than recording their count), underestimates ethnic disparities in sentencing.

The Scottish Government's (2023a) initial analysis of ethnic disparities does not indicate the specific level of measurement used in their set of controls, but since these only included extremely broad offence groupings and did not control for the presence or absence of multiple convictions (see Chapter 3), we should expect that its findings were severely affected by confounding bias. To assess the extent and direction of such bias, we can use sensitivity analysis (Pina-Sánchez et al 2024; Ward et al 2016). How to do so is elaborated below.

4.2.4. Post-treatment Bias

There is little doubt that confounding bias is a widespread methodological problem affecting most studies. This is an uncontroversial and widely acknowledged view amongst sentencing researchers. However, the over-zealousness with which many have aimed to control for all (available) relevant factors to minimise confounding bias, has likely given rise to a far less commonly understood methodological problem: post-treatment bias. This takes place when the set of controls used in a study includes variables which act as mediators as opposed to confounders (Cinelli et al 2022; Vanderweele and Staudt, 2011). Put simply, the mediators precede the causal variable of interest, whilst the confounders follows it. For the case of studies exploring ethnic disparities in sentencing, we could think that sentencers' perceptions of offenders' ethnicity represent the independent variable of interest (Pina-Sanchez et al 2024). If so, post-treatment bias arises when we control for judicial decisions that take place after the sentencer learns about the ethnicity, for example, of the offender, such as decisions of remand or bail, or whether the sentence imposed involved a departure from the guideline sentencing range. Specifically, if these decisions are affected by the same sentencer's perception of the offender's ethnicity, we would be controlling away the causal effect that we seek to estimate, leading to a downward bias in estimated ethnic disparities.

A similar type of bias could be expected when case characteristics being controlled for are influenced by how sentencers perceive the sociodemographic identity of offenders. For example, Guilfoyle and Pina-Sánchez (2024) suggest that some of the guideline factors used by judges in England and Wales are not objectively defined, but rather they are 'racially constructed', i.e. they are affected by judicial perceptions of offenders' race. The authors came to this conclusion after noticing how most factual legal factors listed in the sentencing guidelines, (e.g. whether a guilty plea was entered), were uniformly distributed across ethnic groups. However, factors where judges have greater discretion in determining their relevance (e.g. 'genuine remorse', 'good character', or 'the ability to rehabilitate') tended to be present more frequently amongst White offenders.[39]

In sum, sentencing researchers should be careful about what they control for, do so with restraint, and consider the causal role played by different legal factors.

[39] It should be remembered that judicial sentencers in England and Wales (as in other countries) are far from alone in determining assessments of 'genuine remorse', 'good character', 'the ability to rehabilitate'. Although judicial sentencers make the formal determination, these assessments are made as part of an inter-professional process (Tata 2020, 2023) - including the work and communication of defence and prosecution lawyers and the work of pre-sentence report writers and others (Field and Tata 2023).

4.3. Further Methodological Challenges

Besides confounding and post-treatment bias, there are other forms of bias commonly affecting research on sentencing disparities regardless of the type of data used. Most notably, we should highlight selection bias, which could arise in the presence of neglected upstream disparities in the criminal justice system (Gaebler et al 2022; Knox et al 2020), inappropriate operationalisations of the sentence outcome (Bushway et al 2007), or in the presence of missing data.

For example, Stockton et al (2024) demonstrate how even as few as 10% of missing cases could substantially bias estimates of racial disparities in sentencing when the data is not missing at random, a likely scenario when offenders' ethnicity is self-reported. A similar form of selection bias can take place when the sentence outcome considered involves discarding a share of the sentences imposed. For example, Pina-Sánchez and Gosling (2020) demonstrate how studies that focus on sentence length could lead to biased estimates of disparities when the probability of being sentenced to custody is not equal across demographic groups. A similar type of selection bias may also arise when the same judicial body both decides on convictions and imposes the sentence - as in summary (non-jury triable) cases in England and Wales and in Scotland - yet the analysis focuses only on sentencing, neglecting upstream decisions, such as the nature of any differences in propensity to convict or acquit.

One last common problem affecting studies exploring ethnic disparities stems from the misclassification of ethnic minorities within the dominant ethnic group used as the reference category. For example, Mitchell (2005) shows how studies in the US that see White Hispanics as Whites tend to report smaller ethnic disparities than studies that make such a distinction. Similarly, in their study of ethnic disparities in decisions of charge, the Crown Prosecution Service (2023) shows that White Europeans and Other Whites are more likely to be charged than British Whites. It appears that this problem of misclassified ethnic groups may be present in the Scottish Government (2023a) report on sentencing disparities, which correctly differentiates White Scottish from other Whites. However, it is unknown whether problems of selection bias affect that study. Notably, no information is provided about missing data present in the study.

4.4. Exploring Unwarranted Disparities in Scotland

In this last section of Chapter 4, we proceed to review the types of analyses based on official data that could be carried out in Scotland to explore sentencing disparities. We do that by differentiating the types of analyses that could be carried out with sentencing data already available to government researchers, from more sophisticated analyses that could be considered if new sentencing data is collected.

4.4.1. Analytical Options Using Current Data

By using ready-made official data, we refer to the dataset used in the report on ethnic disparities discussed earlier in Chapter 2 (Scottish Government 2023a). The specific details of the dataset compiled to undertake that analysis are not discussed in the Scottish Government (2023a) report. Assuming that this dataset is similar to other administrative sentencing datasets, such as the latest sentencing data made available in England and Wales (Ministry of Justice 2022), we *might* choose to infer that the Scottish report is based on a census of all sentences imposed, although an unknown share of cases might be missing, in particular if self-reported questionnaires are used to derive offenders' demographic information. We can also gather that this dataset captures offenders' age, sex, ethnicity, income, criminal history, the offence type, and a measure of the location where the sentence was imposed, although its spatial resolution is not clear, nor is it evident whether location refers to the residence of the offender or the location of the court.

Using the already available data, and while also noting the serious limitations already discussed in this and earlier chapters, we could consider a range of analytical procedures to delve into the exploration of sentencing disparities in Scotland and further assess the robustness of the analyses reported in the Scottish Government (2023a) report. Specifically, one could use hurdle models to consider custody and sentence length together. This could even be expanded to consider judicial decisions of guilt. Another straightforward addition would be to expand the focus beyond ethnicity to consider sex and income-related disparities. One could also take this forward using interaction effects, which would allow combining different individual characteristics to consider intersectional disparities. For example, we could assess whether the observed ethnic disparities are uniform across men and women, or those at the top and lower income deciles.

If the area variable available captures the region where the court is located, we could also expand some of the regression models used in the Scottish Government (2023a) report by employing multilevel modelling. This would allow us to explore two key research questions: i) 'could or could not sentencing in Scotland reasonably be described as a postcode lottery?'; and ii) 'to what extent are the ethnic disparities reported uniform across courts?'. For context, previous studies based on the England and Wales Crown Court have found remarkably low levels of inter-court variability, ranging from 2% to 7% of all unexplained variability in sentence length (Pina-Sanchez and Linacre, 2013).

Based on the latest advances in multilevel modelling we could also explore the question of intersectional disparities in a more comprehensive way. Specifically, we could employ 'multilevel analysis of individual heterogeneity and discriminant Accuracy' (MAIHDA) (Evans et al 2024; Pina-Sánchez et al forthcoming2). These models provide two key advantages over more traditional approaches. First, whereas traditional models tend to struggle to estimate more than two-way interactions, MAIHDA models have been shown to offer an adequate level of precision to estimate a combination of individual characteristics with as few as just ten cases. This would

therefore allow exploring the 36 social strata that would result from combining: i) the six ethnic groups recorded in the Scottish data; ii) two sex categories; and iii) three income groups (e.g. bottom decile, second to ninth, and top decile). Second, the MAIHDA approach allows us to estimate the share of variability in sentence severity that stems from offenders' socioeconomic characteristics. In doing so, we can move beyond the 'tyranny of the averages'[40], to reflect the excessive emphasis placed by Social Epidemiologists exploring health disparities through comparisons of systematic differences between specific sociodemographic groups, generally using White male individuals as the reference, which inevitably provides a limited understanding of intersectional disparities. Instead, MAIHDA allows us to estimate disparities in a more holistic way, contemplating disparities across all the combinations of individual characteristics that define the social identity of offenders simultaneously.

Lastly, besides exploring new research questions, we could also use the available Scottish data to enhance the robustness of the findings reported. For example, just like failing to control for relevant legal factors could lead to confounding bias, this could also be the case when the set of controls is incorrectly specified. In that respect, continuous variables like offenders' age, or number of previous convictions, are normally introduced as a single coefficient, assuming a linear effect. However, these would be more realistically specified using polynomial terms, which would account for their diminishing marginal effects, e.g. after nine convictions, every additional conviction does not seem to influence sentence severity, however going from none to one conviction has a large effect (Ri and Cheng 2024; Roberts and Pina-Sánchez 2014).

To assess the robustness of the reported findings to other forms of confounding bias, it would also be useful to undertake sensitivity analysis. Because of their simplicity, we recommend the 'E-Value' (Vander Weele and Ding 2017) when exploring disparities in the probability of receiving a custodial sentence, and the 'Robustness Value' (Cinelli and Hazlett 2020) when differences in sentence length are considered. In essence, both techniques allow us to estimate the required 'strength' of a legal factor that we fail to control in our model to render the estimated ethnic disparities non-significant. This value can then be used to assess qualitatively the likelihood of such a scenario. For example, Pina-Sánchez et al (2025) estimated 42% higher odds of incarceration for ethnic minority compared to White drug offenders sentenced in the Crown Court. Using the E-value, they were able to report that for such disparity to be nothing more than confounding bias, it would be required that the legal factors left unadjusted in their model increase the probability of receiving a custodial sentence by 45% while simultaneously these factors must be 45% more common amongst ethnic minority than White offenders. The authors concluded that the first condition is possible since they failed to account for harm and culpability. However, the second condition can be outrightly rejected since Guilfoyle and Pina-Sánchez (2024) have argued that most legal factors listed in the sentencing guidelines appear to be relatively evenly distributed across ethnic groups. As a result, due to confounding bias, the authors were able to reject the hypothesis that the ethnic disparities reported are robust.

[40] A term coined by Merlo and Wagner (2013).

Lastly, to avoid post-treatment bias, it would be recommendable to consider using offenders' income as a control. This apparently minor detail may, in fact, exert a strong bias. That would be the case if income tends to mediate the effect of ethnicity. For example, if higher income offenders tend to be treated more leniently, and, as would be expected, the average income level is lower for ethnic minority offenders than for the White Scottish group, we should expect the effect of ethnicity to be biased. This absence of control for income might help to explain some of the 'experimental' findings reported by the Scottish Government (2023a).

4.5 Conclusion

The statistical methods outlined above are capable of assisting in advancing our understanding of sentencing disparities in Scotland (although it is vital to bear in mind the important caveats). However, the official administrative data on which such advanced statistical techniques is based remains profoundly problematic. The application of advanced statistical techniques discussed here is inevitably premised on the need to make certain assumptions and draw out conditional extrapolations. These tend to be based on probabilistic and inferential reasoning. The extent to which such techniques produce accurate results is contingent on the reliability of those assumptions and inferences. Ultimately, advanced mathematical techniques are based on the quality of the underlying data available. To make more significant strides forward, high-quality data collection for the purposes of describing and understanding sentencing patterns is needed so that more valid and reliable comparisons can be made, especially so that relative case seriousness can be controlled for. In other words, it will allow more meaningful comparisons, so that like can be compared with like. It is to the discussion of these additional data collection approaches to which we now turn.

Chapter 5: Exploring Ways of Collecting Sentencing Data

5.1 Introduction

This report has found profound limitations in the ability of official administrative data to provide in-depth and meaningful insights into sentencing that the SSC requires to fulfil its statutory functions. This is not to say that official data as it currently stands has no value: it can potentially provide some descriptions of sentencing patterns at a very broad initial level, including possible unwarranted disparities. Nonetheless, there are, as we have described, fundamental weaknesses. These weaknesses include:

- The varying seriousness of multi-versus-single conviction cases is either simply not represented, or very inadequately represented;
- Due to the fragmentation of official data, a key difficulty lies in the practical inability to 'cross-tabulate' sentencing patterns with criteria of particular interest in sentencing. While it can be relatively easy to produce frequency scores (e.g. numbers of men and women sentenced to imprisonment), cross-tabulating that with other information highly relevant to sentencing proves more of a challenge. So, for example, while it is relatively straightforward to describe the proportions of men and women in prison and that women are less likely to be sentenced to imprisonment than men (Scottish Government 2022), we cannot make any definitive statements as to whether or not this shows unwarranted sentencing disparities.[41] These apparent differences may or may not mask differences in case seriousness (e.g. the seriousness of offences, previous criminal offending, propensities to plead 'guilty' or 'not guilty', etc). The categories are simply far too crude to answer key questions such as 'are women or men sentenced more or less severely for otherwise comparable cases?' In other words, at present, with the available data, we cannot begin an approximate comparison of like cases with like cases;
- The inability to evaluate the effects of policy change (e.g. guidelines) on sentencing patterns. We saw in Chapter 3 that the possible effect of changes to policy (e.g. a reduced sentence because of a guilty plea) cannot be gauged by official data derived from administrative sources.

Standing back from these fundamental problems, perhaps the limitations of official data in describing sentencing patterns in any more than the most superficial ways, is not entirely surprising. Despite the admirable efforts of government statisticians and analysts and the

[41] We saw earlier for example that the Scottish Government (2023a) report which sought 'experimentally' (i.e. for the first time) to determine the existence of sentencing disparities on grounds of ethnicity struggled using the available official data to make meaningful comparisons. Meaningful comparisons, for example, cannot be made simply by slicing data up according to extremely broad offence categories such as 'Non-Sexual Crimes of Violence'; 'Crimes Against Society'; 'Miscellaneous Offences'; 'Road Traffic Offences', etc. None of these categories are capable of meaningfully controlling for the variable seriousness of convictions for which offenders are sentenced. All of these broad categories contain within them convictions which vary enormously in terms of seriousness.

application of advanced statistical techniques,[42] ultimately official sentencing data is a product of its administrative sources.[43] Traditionally, different administrative agencies have collected, and largely continue to collect, data for their own daily operational purposes of processing people. They were not established as research tools to assist in the development of policy and understanding of the reality of patterns of sentencing practice. Furthermore, the picture is, as we suggested in Chapter 3, one marked by fragmentation. Each agency collects data in different ways, including different ways to define 'a case', making the ability to marry-up datasets very challenging.

Consequently, while official data can provide us with a relatively superficial picture of frequencies against the broadest of categories, we must ask what other approaches could be used to try to describe sentencing patterns and so begin to assess whether or not there are unwarranted disparities. In this chapter, we outline a range of approaches other than the use of official data. It should be borne in mind that none of these approaches has to be used by itself. All approaches have their strengths and weaknesses. Official data, for example, has the strength of (relative) breadth and time, but is weak in terms of depth. Bespoke research studies tend to be much stronger in terms of depth and nuance but tend to be drawn from very specific samples. Combining the complementary strengths of both official data and bespoke studies will yield a fuller picture of reality.

We begin by looking at larger datasets and move on to more qualitative approaches.

5.2 Census Approaches

Here we look at 'census' approaches in which data was collected routinely about every case sentenced by a court over a duration. Like the official recording of data by administrative agencies, a census seeks to undertake a complete recording of all cases. However, unlike data collected by agencies for operational purposes, a census prioritises obtaining information which is specifically meaningful to understanding and comparing patterns of sentencing. It may be conducted over a specific window of time or on an ongoing basis. Here, as examples relevant to sentencing, we consider the Crown Court Sentencing Survey in England and Wales, and the Sentencing Information System for the High Court of Justiciary in Scotland.

[42] See also Chapter 4.
[43] It is possible, as Chapter 4 suggested, to seek to use advanced, even novel, statistical techniques to try to improve the utility of that official data, by making a series of probabilistic assumptions and inferences about how cases may or may not be randomly allocated. Of course, the extent to which such techniques produce accurate results depends in significant part on the quality of those assumptions and the quality of the underlying data for the purposes to which it is being put to use.

5.2.1 Crown Court Sentencing Survey in England and Wales

The Crown Court Sentencing Survey in England and Wales represents a more sophisticated dataset than official administrative data and was brought together by the MoJ. The Crown Court Sentencing Survey (CCSS), which was operational from 2010 to 2015, was a tool created to enable a more meaningful and deeper understanding of sentencing patterns.

Crown Court judicial sentencers were asked to provide information about each sentenced case. The survey was, therefore, a *census* rather than a *sample* of sentencing decisions in the Crown Court. In the CCSS return, individual judicial sentencers noted important elements of the offence and required the sentencer to indicate the factors taken into account at sentencing. Information derived from the sentencer affords a much closer and deeper understanding of considerations important to sentence outcomes. A distinct advantage of using the CCSS is to capture information to control for a variety of aspects related to criminal history, guilty plea, harm, culpability, and aggravating and mitigating factors, which tend either to be completely missing or superficially described by administrative processes. Furthermore, as an approach dedicated to collecting information to understand sentencing patterns, the fragmentary nature of data from different agencies, each with different recording practices, simply does not apply. As such, it is straightforward to conduct searches of the patterns of sentencing according to specific criteria of interest.[44]

Nevertheless, the CCSS has at least four important limitations.

- Firstly, the CCSS did not, by definition, include the larger proportion of cases sentenced by Magistrates' Courts (which roughly account for over 90% of all criminal court cases in England and Wales;
- Secondly, like all census approaches, the CCSS relied on returns being completed. Response rates among courts were variable: over 90% in some Crown Court locations, but much lower in others. In the final year of the survey, the completion rate across all Crown Court locations was 60%. If it were the case that the judicial sentencers with low and high response rates were substantially different in their sentencing practices, the data in the CCSS would have suffered from bias;
- Thirdly, as discussed above, sentencing data needs to meaningfully represent the differential seriousness of single- and multi-conviction cases. Although the CCSS took a step in the right direction by permitting the court to note the presence of a multiple conviction case, it did not ask the court to identify the convictions beyond the single most serious (or the principal) conviction;

[44] Although the CCSS has been praised for their level of detail, Pina-Sanchez et al (2025) has argued that it is possible that such level of detail is not entirely necessary. For example, for the case of shoplifting offences sentenced in the Magistrates' Court, after controlling for the top-20 most consequential guideline factors, adding new factors to the model did not provide any meaningful information.

- Fourthly, although the CCSS recorded the information identified by Crown Court sentencers, this would not necessarily include *all* the important information affecting the sentence. The form that sentencers were asked to complete listed all of the sentencing *guideline* features and provided respondents with the opportunity to note other characteristics.[45] Other factors which are not captured by the form may have influenced the sentence imposed and the survey cannot detect the influence of the outcome. One obvious example would be demeanour and the attitude of the individual towards the offending (Field and Tata 2023), which might also be contextualised by the content of pre-sentence reports (Tata et al 2008). A more troubling example might be perceived racial or ethnic status. This kind of information has to be collected in complementary ways (e.g. through bespoke research such as court observations, interviews with defendants and legal practitioners, or bespoke academic research).

Despite its limitations, the CCSS represents a significant advance compared with the extreme limitations of data derived from administrative agencies. Indeed, even today, it is still commonly drawn on as a source of information.

However, the CCSS was discontinued in 2015 (Gormley et al 2022; 2023). Today, instead of an ongoing census, the SCEW conducts one-off, bespoke data collection exercises in both the Crown and Magistrates' Courts. The Council justified the decision to replace the CCSS with periodic data collections on the grounds of efficiency. The Council wished to reduce the burden on sentencers who had to complete a return for every sentencing decision. Thus, the experience with the CCSS suggests that a census approach was felt to be too burdensome, although it might be argued to be akin to a necessary ‘start-up cost’ for the first few years to provide a ‘base-line’ of data.

Nevertheless, the superior depth of data of the CCSS (compared to that brought together by the MoJ) has enabled more accurate analysis of long-standing contentious claims. An example of this is the relationship between ‘race’/ethnicity and sentencing practices, which is discussed later in this chapter.

5.2.2 Sentencing Information System for the High Court of Justiciary, Scotland

Another census approach is the idea of a Sentencing Information System (SIS) where the primary end-user *may* be conceived as the judicial sentencer.

Over the period of around a decade (1993 to the early/mid-2000s), a project was conducted to research, develop, and implement a Sentencing Information System (SIS) for the High Court of

[45] Our understanding is that these free response options have not been analysed.

Justiciary. It was initiated by members of the senior judiciary who were closely involved in its direction, together with the support of an academic research team from the University of Strathclyde (Hutton and Tata 2010; Tata 2000).

The aim of the SIS was to enable High Court judges (and the Court of Criminal Appeal) to pursue consistency in sentencing. The information system was created and developed with the goal of making it useful to its end users: first instance High Court and Appeal Court judges deciding matters of sentencing.

By allowing judicial users quick and easy access to the patterns of sentencing in similar cases, judicial sentencers could, for example, see how a sentence she or he was considering imposing might compare with other reasonably similar cases. Initiated and driven forward by the senior judiciary, the SIS was seen as an alternative and better way to pursue consistency in sentencing to head-off the imposition of legislative guidelines (or mandatory minima) - most especially of the more intrusive kind government ministers were proposing (Hutton and Tata 2010; Tata 2020).

From the outset, it was strongly agreed that consulting the SIS was to be a voluntary choice for individual judges and, unlike an artificial intelligence (e.g. Expert System) approach, there was no question that it would dictate to the judicial sentencer what 'the correct' sentence should be. Judges were free to pass the sentence they thought appropriate, but they were encouraged to consult the SIS as part of their ruminations.

Encouraging judicial sentencers to pursue consistency in sentencing by accessing systematic data about patterns of sentencing in similar cases is not a new idea. As long ago as 1953, Morris proposed that sentencers be given information about sentences imposed so that they could 'see clearly where they stand in relation to their brethren' (Frase 1997: 366). While the principle was simple enough, it was not until the 1980s when advances in information technology meant that databases about sentencing decisions were pioneered in British Columbia (Hogarth 1988); in various Canadian provinces (Doob and Park 1987: p.54); and then by the Judicial Commission in New South Wales.[46]

The SIS project began by investigating the feasibility of using existing official administrative data sources to describe patterns of sentencing in similar cases. After a lengthy investigation, it was concluded that administrative data was not capable of providing sentencing data of the kind needed to represent existing sentencing patterns sufficiently accurately or meaningfully. It also concluded that, regrettably, there was no prospect of official data from administrative sources becoming useful for this purpose in the short or medium-term. It was clear that while there was some high-level in principle interest in a shared system-wide approach to data collection (e.g.

[46] The New South Wales SIS is now part of a wider interlinked Judicial Information Research System run by the NSW Judicial Commission. See Potas (2005) and Tata (2000).

with shared case identifiers and case-counting rules), daily operational imperatives in collecting data for its own administrative purposes were the overwhelming priority (Hutton et al 1996; Tata 1997; Tata et al 2003).

Therefore, it was decided that the SIS would have to collect its own data. The key issue was how to describe case 'similarity' (Tata 1997; 2020). Acutely aware that the SIS had to provide information about patterns of sentencing in cases which were reasonably comparable, considerable work was done to develop ways of collecting case information. In regular consultation with its judicial users and through a range of sentencing exercises, the SIS project sought to reflect the ways in which judges thought about sentencing so that they, as its eventual end-users, would find it helpful. The SIS created a taxonomy and means of collecting data. Initially, data was collected from court archives, and then later it was planned to be collected contemporaneously. Because it was dedicated to sentencing, the Scottish SIS contained some of the most in-depth and detailed information from the perspective of sentencing anywhere in the world.

Rather than seeking to minimise the need to distinguish between single and multi-conviction cases to a technical footnote, the SIS project sought to tackle the problem. It used two approaches to mirror the two ways in which judges said they thought about such cases. The first was to employ a 'Principal Offence Approach' in which other convictions could be recorded as ways of adding further information to that principal offence. However, it was recognised that some cases involved two or more convictions of similar gravity and so recording only one as the 'principal offence' could be problematic. The SIS sought to overcome this problem by offering users a second, complementary way of exploring the data. A 'Whole Offence Approach' to the recording and representation of sentencing data was developed with the judiciary to reflect multiple conviction cases where there was 'a course of conduct'. This whole offence approach was complementary to 'the Principal Conviction Approach' so as to capture the inter-relationship between offences:

> 'This approach is less fragmented than the Principal Conviction Approach and reflects the holistic way in which judges, (in common with other skilled discretionary decision-makers), often conceive and make sense of the narrative or 'course of conduct' of the offending behaviour.' (Tata and Hutton 2003: 69)

The SIS contained information about all sentences passed over a 15-year period (1989-2003 - some 15,000 cases) by the High Court, including appeal decisions, as well some brief narrative remarks made by the sentencing judge in individual cases. It was planned that information should initially be collected retrospectively by the research team from court/trial papers from the High Court archives, but that once there was a large baseline of information, the recording of information should be taken over by clerks of justiciary. The plan was that they would record information contemporaneously, according to a template, with judges being able to add narrative information. In that way the database would be continuously updated and expanded (Tata et al 2003).

5.2.2.1 An SIS Today?

The story of the SIS for the High Court was a product of its time in the 1990s and early 2000s. By the early 2000s, the wider political context appeared to have changed. The threat of intrusive legislative guidelines seemed to be off the political agenda, (at least for a while), and judicial discretion and independence seemed once again to be safe. This, combined especially with changes in senior judicial personnel, meant that there was much less support for the idea of an SIS as a means to pursue (and be seen to pursue) consistency in sentencing. Predictably, a lack of an institutional home for the SIS and lack of institutional authority without a visible judicial champion left the SIS vulnerable.

As a direct consequence, the SIS was not maintained by the courts service. After the project was fully handed over from the university to the courts service for maintenance and updating, senior managers in the courts service took the deliberate decision not to implement the recommended data-entry training of justiciary clerks nor the recommended data quality control regime.[47] Over time, therefore, the justiciary clerks who were supposed to input data about new cases felt less requirement to do so. The predicted consequence was that judges began to lose interest in an information system which was which was 'allowed to atrophy' (Hutton and Tata 2010: 275).

Yet, if an SIS was developed today, it would be in a potentially stronger and more sustainable position. Since the early 2000s, the institutional situation has changed radically. In the 1990s and early 2000s, there was no obvious institutional home which would have a stake in championing, maintaining and improving an SIS in line with the needs of its users. During its lifetime the SIS for the High Court was labelled as 'pilot project', which was a collaboration between the judiciary and a university team. If it were developed today, an SIS could have an institutional home in, for example, the SSC or Judicial Institute.[48]

5.2.3 Wider uses of Census Approaches

Census approaches, like the CCSS and an SIS, enable the user to view information according to different criteria and see instantaneously how the patterns change. Although the SIS for the High Court was envisaged to be used as a point of sentence tool for individual High Court judges considering individual cases, a number of others (sheriffs, crown, defence) expressed

[47] It was asserted by the senior clerk that training and quality-assurance monitoring was not necessary because justiciary clerks were professional and so would simply know what to do and so there was no need to institute the recommended measures.
[48] Importantly, the jurisdiction to have maintained an SIS over longest period is in New South Wales, Australia. There the SIS forms a central part of a wider Judicial Information Research System, which is a key element of the work of the NSW Judicial Commission.

interest in having access to the SIS or having a similar information system for their own purposes.

Moreover, the systematic information dedicated to sentencing contained in a census could also offer wider uses. Below we outline three examples: informing public knowledge of sentencing; a resource for judicial training and ongoing judicial education; and a bank of information to enable analysis of sentencing trends over time and the impact of policies.

5.2.3.1 Public knowledge: inform public perceptions of sentencing

It is now well established both in Scotland and comparable jurisdictions that judicial sentencing is often believed by the public to be too lenient (e.g. Black et al 2019; Biggs et al 2021; Hockaday et al 2025). Other jurisdictions have been able to show that this perception of lenience may not be borne out by the evidence of patterns of sentencing. By focusing on specific kinds of cases (vignettes), research has asked members of the public what their preferred sentence would be; what they believe to be the typical sentence for this kind of case. Typically, the public tends to think of the typical sentence as considerably more lenient than the sentence they would prefer to see passed given the same information.

Crucially, this can then be triangulated with the actual patterns of sentencing in that kind of case. Where data on the patterns of sentencing given the specifics of that case is available, it is possible to assess the extent to which the preferred sentence suggested by members of the public aligns with the reality of judicial practice. Where high-quality data is available to enable this triangulation, a powerful 'good news' story can be told. Members of the public may express some surprise and relief that the sentence they think is appropriate is more or less what is being done, and that they tended to imagine that the reality was characterised by far greater lenience than it is.

The problem which Scotland faces is that, by and large, it is very difficult to tell any good news story because triangulation between public preference and expectation cannot be completed with the reality. There is not sufficient high-quality data showing the normal pattern of the reality of sentencing for specific kinds of cases. In their report about public perceptions of sentencing of a range of scenarios (including areas of heightened public concern like serious sexual assault and causing death by driving offences), Black et al had no alternative but to try to obviate this problem by asking members of the research team[49] and officials what they supposed the typical sentence to be:

> 'Likely sentences for these specific scenarios have been agreed based upon sentencing data, where available, and the professional experience of the report authors, their

[49] This included one of the authors of the present report.

consultants, members of the Scottish Sentencing Council, and the Council's secretariat. *The likely sentences should therefore not be viewed as confirming how these or similar offences would in fact be dealt with.*' (Black et al 2019: 7, emphasis added)

As knowledgeable as those individuals may be about what possible sentences could be, this is no substitute for systematic and verifiable data showing the actual pattern of sentencing in this kind of case. If this information about the typical patterns of sentencing for such scenarios were available, triangulation between the public's preferred sentence ranges, the expected typical range of sentences, and the actual reality of the typical range of sentences could be completed. This would allow it to be clearly demonstrated to a range of audiences that the public's feeling of excessive leniency in sentencing may not, in fact, be borne out in reality.[50] Regrettably, due to the severe limits in terms of the quality of data about sentencing patterns, this triangulation is not achievable at present.

5.2.3.2 A useful resource relevant to judicial training and ongoing judicial education

A census approach, especially where (as in an SIS) the user interface allows for different criteria to show different distributions and even individual cases, can be a valuable resource in judicial training and education. Such an approach can be a way to provide data that is both meaningful to and respected by sentencers. It can also assist in the systematic analysis of features sentencers may be encountering in the real world.

5.2.3.3 A bank of information to enable analysis of sentencing trends over time and the impact of policies

Census approaches mean that it can be possible to investigate the impact of key changes (e.g. sentencing guidelines) on practices on the ground. It can also allow for consistencies and unwarranted disparities to be gauged and tracked over time, thus enabling further study and the appropriate action to be taken. Using census and other research techniques[51], specific topics of public interest can also be investigated and papers (with user-friendly summaries) can be

[50] It could also enable bodies like the SSC to examine more closely and with greater precision and engage with those specific issues where there may be significant differences between public preferences and actual sentences. Indeed, as the SSC's own commissioned research on public attitudes to sentences following a guilty plea demonstrates, communication can have a significant impact. See Gormley, Roberts and Tata (2025).

[51] Where useful, analysis of census data can be carried out in conjunction with other methods such as bespoke targeted studies – which we discuss briefly later in this chapter.

published on 'hot topics' to help to inform public knowledge and debate on contentious matters (Tata and Hutton 2003).[52]

The Sentencing Advisory Council (SAC) in Victoria, for example, takes an active role in the dissemination of information about sentencing patterns. Since its creation in 2004, the SAC has published many Sentencing Bulletins. These documents provide regular snapshots of current sentencing practices for an offence or offence category (e.g. Sentencing Advisory Council 2022). They are widely used by the media, official organisations, advocacy groups, and have also been cited by sentencers in their sentencing decisions.[53]

An illustration of how census information can be used can be seen in the recent work of the SCEW deploying the rich data from its CCSS to help advance understandings of questions about race/ethnicity and sentencing. The SCEW has published analyses of racial variations for some drug offences. Together with other reports by the MoJ, this sort of work can contribute to a better understanding of the issue in England and Wales. These data are a result of collaboration between the SCEW and the MoJ.[54]

Under Section 95 of the Criminal Justice Act 1991, the MoJ has a duty to publish statistics documenting any differences between groups of offenders at all stages of the criminal justice system, including sentencing. To seek to address this, the MoJ publishes a biennial statistical report.[55] However, these section 95 reports are unable to control for a range of case characteristics which may (or may not) legitimately explain differential patterns of sentencing. For example, differences in pleading trends (not guilty or guilty) or criminal histories might point to discriminatory practices 'up-stream' (e.g. policing, prosecution practices) or variable levels of trust in the courts, but they may not necessarily show discrimination at the point of sentencing. As we saw in the example of the Scottish Government (2023a) report purporting to show differential rates of imprisonment, comparisons without such controls can be misleading. Indeed, the section 95 reports caution against drawing direct inferences of discrimination:

> 'No causative links can be drawn from these summary statistics… Differences observed may indicate areas worthy of further investigation, but should not be taken as evidence of bias or as direct effects of ethnicity.' (Ministry of Justice 2019).

To improve on these section 95 reports, one-off studies have deployed additional statistical techniques to control for case similarity.[56] Recently, the MoJ published a report that sought to

---

[52] Recording key sociodemographic characteristics of the offender, would enable expanding the exploration of intersectional disparities using MAIHDA beyond what has been made possible so far, and in so doing identify with unprecedented level of detail specific subgroups of the population that appear to be more disadvantaged.
[53] An extract from a snapshot is provided in Appendix B.
[54] For a review, see Roberts and Bild (2021).
[55] For example, Ministry of Justice (2019).
[56] These studies have used different ethnic classifications, temporal periods, and data sources.

identify the points in the criminal justice system where there appeared to be unwarranted disparities (Uhrig 2016).[57]

However, an important step forward was taken with research published by the SCEW in 2020. It drew on the more in-depth sentencing information in its *Crown Court Sentencing Survey* (CCSS) which it combined with ethnicity information from MoJ databases. As we discussed above, the CCSS was generally a far superior dataset. The data was provided directly by each individual sentencer. It captured information that is recorded by official data from administrative sources.

5.2.4 Concluding remarks about census approaches

Census approaches are a way of collecting and retrieving systematic and comprehensive information about sentencing practices. Dedicated to the task of representing sentencing, they are far superior to the use of official administrative data that is collected for other operational purposes. Indeed, it is a testament to the quality of the CCSS that a decade on, it provides an invaluable baseline and continues to inform sentencing policy and research.

Yet, while providing a largely unparalleled insight, the SIS and CCSS encountered difficulties in maintaining and updating the data long-term. The attempt to be wholly comprehensive can (as we have seen) also be a potential Achilles Heel. Without authoritative backing and some resourcing, a census approach may encounter reluctance, even resistance, to the collection of information. Both the SIS and CCSS were eventually discontinued because of a lack of willingness (or inability) to keep them updated.

The fates of the SIS and CCSS suggest that census approaches should be either focused (e.g. perhaps time-limited) or have resourcing and/or clear and authoritative leadership to operate long-term. Either way, it would also seem prudent to communicate to sentencers the benefits that census approaches provide. Indeed, while a census may not take much time in and of itself, one could hardly fault busy sentencers who are disinclined to engage in what could appear to be a futile exercise that is extraneous to their functions. In other words, as noted earlier, the census and its data would have to be meaningful to, and so respected by, sentencers.

If it was determined that a comprehensive census approach would be too sizeable and/or too difficult to ‘sell’ to those collecting the data, a more modest alternative could be to deploy a series of mini-censuses. These could be targeted to address specific issues, which SSC is planning to address in the future, and/or matters of particular concern. For example, the SSC could integrate a mini-census into its planning of guidelines. Presently, the main empirical

[57] That analysis followed similar analyses in the US and was recommended by the Lammy Review (2017).

research which SSC publishes in the development of guidelines is into public perceptions. Further baseline sentencing data presents two potential benefits.

First, and as discussed more fully above (at 5.2.3.1), this data on public preferences and expectations could be triangulated with data on actual sentencing patterns for specific kinds of cases. Such triangulation can potentially tell a powerful good news story by showing clearly that judicial sentencing is not as lenient or 'out of touch' as is often believed.

A second benefit is that this data would also provide a baseline of sentencing patterns to help to inform future work (e.g. guidelines or research). After a suitable period of time following the implementation of a guideline, further data could be collected to assess the impact of the guideline. In this way, a census approach, or mini-census approach could provide pre- and post-guideline data and thus enable a more or less direct evaluation of the effect of a guideline.

However, while perhaps the best data collection option in many ways, a census is not a panacea. The complex issues raised by sentencing mean there can also be vital questions (for example, about possible unwarranted disparities) which also need to be addressed through bespoke, targeted approaches – to which we now turn.

5.3 Uses of AI: Large Language Models

Sentencing data can also be derived from the coding of transcripts of remarks made by the judicial sentencer at the point of sentence. Assuming the AI operates as intended (e.g. it does not 'hallucinate' or make errors), then as with a census return completed by the sentencing judge (e.g. CCSS), this method has the advantage of being derived from 'the horse's mouth'. It allows recording case characteristics reported to be relevant by the sentencer at the moment of passing sentence.

To put it very simply and succinctly, the use of large language models can simply be seen as a mathematically sophisticated version of what legal commentators have done for many years: counting and noting the mention of certain words (e.g. key word indices and terms) and in certain contexts. However, the use of artificial intelligence techniques adds potential sophistication in observing patterns to which the model is sensitive.

Similar to court observations, the analysis of sentence transcripts has traditionally been limited by its lack of scalability, which made it a resource-intensive methodology. However, in the last decade this data collection method has gained renewed interest because of recent advancements in big data and artificial intelligence that allow automatising the data coding process (Drápal et al 2023; Pina-Sánchez et al 2019, see also ECR project JuDDGES: *Judicial*

Decision Data Gathering, Encoding and Sharing)[58]. This has practically eliminated the method's marginal cost after a certain volume of cases are reached and the coding algorithm well-calibrated. As a result, in research contexts where large samples are necessary, the automatised coding of sentence transcripts has the potential to be relatively inexpensive.[59]

In sum, there are both optimistic and pessimistic views about the potential of AI in this context. The truth probably lies somewhere between the utopian and dystopian visions and, as with other areas of technology in criminal justice, there are always better and worse ways of doing things: meaning caution is warranted. A limitation of this approach is again the inability to observe relevant offender or case characteristics that the sentencer, consciously or subconsciously, chooses not to discuss in open court, or which appears (at least to the model) to be ambiguous. Accounts of sentencing in open court are, of course, skilfully crafted, keeping in mind the various audiences and (often competing constituencies) which judges dexterously have to address and balance. A single account may be composed of multiple sub-explanations and acknowledgements (Tata 2002), which may or may not be easily absorbed through machine-learning. A further limitation of this approach is that, like other Anglo-American jurisdictions, in non-jury triable cases there may not be a written transcript of sentencing remarks.

5.4 Bespoke, Targeted Approaches

Approaches to collecting information in a targeted way can provide smaller but more in-depth data. This can be especially useful in addressing strategic issues or questions of contemporary importance. Internationally, such bespoke studies are commissioned by a range of bodies, including sentencing councils. Here we outline some approaches. These should not be seen as mutually exclusive: they can, and should, be combined with each other and indeed with larger data collection efforts (e.g. census and/or official data). In this way, the complementary virtues of breadth and depth can be realised.

5.4.1 Archival

It has been consistently noted that official data from administrative sources suffers from some serious flaws. This is because, historically and (quite understandably), administrative agencies

[58] JUDDGES project website: https://sites.google.com/pwr.edu.pl/juddges/home?authuser=0.
[59] As demonstrated by Drápal et al (2023), using large language models removes the need to be trained in data analysis or programming to be able to undertake the data coding, so this work could be carried out by social researchers. For all the above reasons, under the condition that sentence transcripts can be made available to the Scottish Sentencing Council, consideration could be given to a data collection strategy based on a combination of automated coding of sentence transcripts supplemented with an offender questionnaire.

have recorded and continue to record information which is relevant to the daily operational work of each agency. Nevertheless, court papers contain a wealth of information about cases.

For example, evaluation by justice social work reports[60] of the person to be sentenced, can be pivotal in assessing the suitability of that person for a community sentence (e.g. Tata et al 2008). Yet, this vital information is not recorded by the courts service and cannot be combined (cross-tabulated) in any useful way with information recorded by social work.[61] However, such information can be recorded from court papers according to an agreed pro-forma, along with other key information. This can enable specific depth studies to be conducted about specific kinds of cases and issues. Equally, it can be carried out in terms of a census approach discussed above.

5.4.2 Observations

Observation of court proceedings is often vital so as to gain a naturalistic understanding of decisions. Combined with other techniques (e.g. interviews), it can be a powerful means of describing the nuances of proceedings and observed experiences. However, it is of course potentially time-consuming and it is important to take a targeted approach to specific kinds of cases which need to be identified in advance. It requires goodwill from a range of agencies to help to make sense of proceedings, and, where possible, access to key papers.

5.4.3 Interviews and Focus Groups

Interviews with sentencing judges, and others involved in the sentencing process are an invaluable way of gaining views and perceptions. As is demonstrated in research with those sentenced, or victims and their supporters (e.g. Jacobson et al), interviews allow participants to discuss how they understood events both before and after proceedings. Combined with observations, this can offer a unique opportunity for triangulating experiences.

Focus groups can offer similar possibilities, while in the context of the dynamic of discussion within the group. This has been a particularly common technique in exploring public perceptions of sentencing (e.g. Gormley, Tata and Roberts 2025). Focus groups and interviews may also be combined with interviews, and indeed any of the other approaches outlined here.

[60] Previously known as ‘Social Enquiry Reports’ and sometimes known in different jurisdictions as for example pre-sentence reports.
[61] As we saw in Chapter 3, data is fragmented among different agencies each with their own recording practices.

5.4.4 Simulations

Simulations, such as the use of mock-sentencing exercises or vignettes. These simulations can be a valuable way of teasing out key issues as they offer the potential for a high degree of control of specific information in order to tease out how different features impact the meaning of a case and its likely sentence. Additionally, while space is limited here, we would also note that vignettes can focus on abstract debates and elicit more specific information from participants.

Simulations (possibly combined with focus groups) may also yield important data to help to improve accurate public awareness of sentencing. We noted above that there is a significant opportunity to identify key points of convergence in and divergence between the practices of judicial sentencers and the perceptions of the public, especially in specific kinds of cases or issues of interest. This can be achieved by triangulating the public's preferred sentences and their expectations of typical sentencing reality *with* the actual reality of sentencing patterns. Where there is significant alignment between the public's sentencing preference and the reality of sentencing patterns for specific case scenarios, it becomes possible to demonstrate clearly that public expectation of leniency may be misplaced. This is a potentially powerful good news story in terms of public confidence. Where there is not alignment it would enable a body like SSC to target engagement work and to consider reflecting on guidance to sentencers.

The problem, as has been noted throughout this report, is that currently the official data is not of the high-quality to enable the patterns of sentencing to be shown for these case scenarios. A census or mini-census is, as discussed above, one way to fill this information gap.

Another, possibly more agile, approach would be to use hypothetical case-scenarios, each given to separate groups of members of the public and judicial sentencers. Sentencing judges would be asked to consider different hypothetical case scenarios in areas where the Council is considering developing guidance. If these scenarios are also used to ask members of the public about their sentencing preferences and perceptions, it means that a more or less *direct three-way comparison* can be made between judicial sentencer practices *and* public preferences *and* public expectations. In this way, the opportunities presented by this direct triangulation between public preference, expectation of judicial sentencing, and the reality of sentencing patterns can be achieved.

Simulations, involving hypothetical case-scenarios, are unlikely to be seen as alien to judicial sentencers: they are already a fairly familiar technique to them. Scenarios could be combined with focus groups as a way of teasing out thinking. They are a relatively in-depth and effective way of ascertaining a sense of patterns of sentencing from a reasonably representative sample of sentencers, which can then be compared with the views and perceptions of the public. In this way, simulations using case scenarios can be used as an addition (or alternative) to full or mini-census exercises.

5.5 Conclusion

Official sources of data about sentencing derived from administrative agencies in Scotland suffer from some fundamental flaws. This is not to say that it has no value: it can offer helpful descriptions of frequencies at a very broad level, including signposting *possible* unwarranted disparities (although, as above, it may be over-inclusive or under-inclusive in some contexts).

That said, the limitations of official data in describing sentencing patterns in any more than the most superficial ways are not entirely surprising. Despite the admirable efforts of government statisticians and analysts and the application of advanced statistical techniques[62], ultimately official sentencing data is a product of its administrative sources.[63] Traditionally, different administrative agencies have collected, and largely continue to collect, data for their own daily operational purposes of processing people. Their recording practices were not established as sentencing research tools and the imperatives of daily operations may leave little scope for what can be perceived as extraneous resource-using extravagances. Furthermore, the picture is, as we suggested in Chapter 3, one marked by fragmentation. Each agency collects information in different ways, not least in how each defines ‘a case’, making the ability to marry-up datasets extremely challenging.

As a result, official data can provide a broad picture of frequencies against the broadest categories. However, other approaches are needed to adequately describe sentencing patterns and begin to assess whether or not there are unwarranted disparities.

This chapter has outlined and evaluated a range of approaches other than the use of official data derived from administrative sources. These include larger census studies (e.g. CCSS and an SIS). Because they are dedicated to the collection of data which is specifically useful to understanding sentencing patterns in genuinely similar cases, they produce much more in-depth and meaningful data, which can be far more useful to bodies such as sentencing councils.

Inevitably, a challenge raised by larger census studies is how such information is collected (especially if it is to be collected contemporaneously, rather than from court archives). As much as the question of resourcing, it also underlines the importance of leadership. Clerks of court or judicial sentencers are far more likely to be willing to record information if they can see the point of doing so (namely that it will add benefit to the system in which they work, including, for example, helping to improve public trust in and knowledge and understanding of sentencing).

That said, even where census studies have had a limited life, they can represent a step change, leaving a positive legacy. For instance, the CCSS data has proved to be an invaluable depth

---

[62] See Chapter 4.
[63] As above, it is possible, as Chapter 4 suggested, to seek to use advanced statistical techniques to try to improve the utility of that official data, by making a series of probabilistic assumptions about how cases may or may not be randomly allocated. Of course, the extent to which such techniques produce accurate results depends in significant part on the quality of those assumptions.

resource, long after its discontinuation, enabling the SCEW to investigate and answer contentious topics of heightened public concern.

It might be concluded that a full census approach would be too large and would be likely to encounter resistance from those asked to record that data. If so, a more modest and practicable alternative would be to employ smaller, targeted mini-census exercises. The could be focused on finding out patterns of sentencing according to selected criteria in the sorts of cases which may be the subject of SSC's future work. This approach offers at least two possible benefits. First, it would also enable SSC to collect baseline data before any change (e.g. a guideline) came into effect. This would make it easier to observe possible post-guideline change. A second benefit is that the reality of sentencing patterns in specific kinds of cases could be compared *directly* with research findings about public preferences in the same kinds of cases and with what they expect is the normal reality of patterns of sentencing in those same kinds of cases. This direct comparison could show clearly whether or not judicial sentencing really is out-of-line with public preferences. It may transpire that sentencing patterns are not as lenient as the public imagines and closer to what they themselves would prefer. This is a potentially potent 'good news story', which could be disseminated to wider policy and public audiences.

An additional approach would be to deploy two sets of simulation exercises (deploying hypothetical case scenarios) run in parallel: one with members of the public and the other with judicial sentencers. These simulation exercises would ask the two groups to consider the same case scenarios. The public would be asked for their preferred sentences and their expectations of what they think the typical range of sentences actually passed by sentencers would be. By asking the judicial sentencer group what they would in fact pass for the same sort of case, it will become possible to draw a direct comparison. This would help the Council to identify areas of convergence and divergence between the two groups so as to inform future public engagement work and potential guidance to sentencers.

Simulations may have the advantage of being more agile than a large (or even mini) census approach. It will be easier to tailor and vary (or 'titrate') different hypothetical case conditions so as to observe the differences those variations make in the minds of both members of the public and judicial sentencers.

In sum, it should be borne in mind that none of these methods has to be used in isolation. All approaches have their strengths and weaknesses. Official data, for example, has the strength of (relative) breadth over time, but is weak in terms of depth. Census and bespoke research studies may be much stronger in terms of depth and nuance, but can be drawn from very specific, smaller samples. Using a combination of approaches will tend to provide a more complete understanding.[64] In other words, there is no need for efforts to improve the quality of sentencing data to put all their eggs in one basket. For instance, smaller census studies could

[64] The use of AI (such as Large Language Models) may also be worth consideration, but caution also is needed at this early stage.

themselves be targeted periodically to the collection of information about specific cases. These can be fruitfully combined with qualitative approaches, which answer different kinds of in-depth questions. Quantitative data can tell us about *what* is happening overall, whereas qualitative approaches help to understand the *how* and *why*. One is not better than the other: they simply have different complementary strengths. The two should go hand in hand and are mutually informative.

Chapter 6: Conclusions

6.1 Introduction

According to its statutory duties, the Scottish Sentencing Council 'must': promote consistency in sentencing; assist the development of sentencing policy; and, promote greater awareness and understanding of sentencing. To fulfil its requirements, the SSC is obliged to publish information about sentencing in Scotland. Accordingly, we were commissioned by the SSC to examine the possibility of unwarranted disparities in sentencing. Our aim was, therefore, to scrutinise potential indicators of disparities in sentencing, while also attending closely to complexities, including possible confounding characteristics which might help to distinguish warranted from unwarranted disparities. Following instructions from SSC, we concentrated on the possibility of sociodemographic disparities (e.g. race/ethnicity, gender, etc) and did not seek to examine the absence or presence of unwarranted disparities between judicial sentencers.[65]

Having surveyed the contemporary international literature[66], the initial aspiration for this project was to secure access to disaggregated administrative data to conduct a novel and controlled analysis so as to shed new light on patterns of Scottish sentencing practices according to a range of specified and relevant criteria. However, due to limitations beyond our and the Council's control (see Chapter 3), this proved not to be possible. In essence, the evidence as to the absence or presence of unwarranted disparities is extremely limited because, in terms of sentencing at least, the quality of available official data appears to be inadequate.

As such, this report has documented attempts to explore the absence or presence of unwarranted disparities in sentencing practices in Scotland. The provision, use, and dissemination of high-quality data about typical patterns of sentencing is essential to a number of key functions which support the legitimacy of sentencing and trust in an independent and impartial judiciary.

6.1.1 Promoting the Consistency of Sentencing

Genuine consistency in sentencing does not mean uniformity: treating every case as the same regardless of all relevant case differences. Simply put, consistency means that like cases should be treated alike and unlike cases should be treated differently. Although every individual case is in some respects unique, sentencing is also necessarily a comparative exercise, requiring the weighing of (among other things) relative harm and culpability compared to other cases. Comparison requires being able to assess what is meant by 'similarity' from the perspective of sentencing.

---

[65] That said, most of the methodological challenges we discuss are common to both.
[66] See Chapter 2.

To be able to promote the consistency of sentencing (and thereby reduce unwarranted disparities such as on grounds of race/ethnicity (or, if required, on the basis of the individual judicial sentencer), it is essential to be informed by high-quality data about sentencing patterns for different kinds of cases. Without that, it becomes impossible to know whether or not there is consistency or unwarranted disparity. Nor is it possible to identify a solution if one does not know where and how those consistencies and unwarranted disparities appear. To put it another way, one cannot provide a remedy to a problem without a sound diagnosis. High-quality data about patterns of sentencing according to a range of key criteria (e.g. relative seriousness of the offending, criminal history, etc) is essential to the pursuit of consistency in sentencing.

Like any area of policy, the development of sentencing policy requires regard to systematic information about patterns of sentencing according to different criteria. One obvious area is the development of sentencing guidelines. Monitoring the impact (or non-impact) of guidelines is essential. Guidelines may be intended to have some effect on everyday practices. The extent and nature of any impact of a guideline cannot, of course, be answered by simply restating the content of the guideline. By its nature, a guideline is a rule-like instrument setting out what ought to happen. It cannot measure the empirical reality of the patterns of sentencing on the ground.

Without the recourse to high-quality data, it becomes very difficult to know (or indeed demonstrate) what the measurable effect (if any) a guideline has had on actual practices.[67] This cannot be done simply through the assertion that one or other individual 'knows' what the impact may or may not be.[68] While such accounts should be carefully attended to, individualised knowledge, (valuable as it is), is inevitably fractional, anecdotal, and impressionistic. What may be seen as 'obvious' to one experienced individual may not necessarily be so to another. Furthermore, individualised views cannot demonstrate robust evidence of share sentencing practices.

The subjective experiences of those applying guidelines should be studied and form *part* of the picture, *together with* the provision of systematic and high-quality data showing the changing patterns of sentencing according to different criteria. This can then help to inform the development of future guidelines or other policy work.

[67] For example, we saw in Chapter 3 that the effect (or non-effect) of changes to policy (e.g. a reduced or 'discounted' sentence because of a guilty plea) cannot be gauged by official data derived from administrative sources.
[68] See for example the discussion at 5.2.4.1.

6.1.3 Promoting Public Knowledge and Understanding of Sentencing

The provision of high-quality data about sentencing patterns is vital to work seeking to promote more accurate knowledge of sentencing and thus public confidence in sentencing (and by extension to the independence of the judiciary).

It is now well established both in Scotland and comparable jurisdictions that judicial sentencing is often perceived by the public to be too lenient (e.g. Black et al 2019; Biggs et al 2021; Hockaday et al 2025). Other jurisdictions have been able to begin to show that this perception of lenience may not be borne out by the evidence of patterns of sentencing. By focusing on specific kinds of cases (vignettes), researchers can ask members of the public what their preferred sentence would be and what they believe to be the typical sentence for this kind of case. Researchers can then compare this public preference and public expectation with the reality of actual sentencing practices. Often, the public tends to think the typical sentencing patterns are considerably more lenient than the kind of sentence they would prefer to see passed, given the same information. Upon that realisation, people may express surprise and more positive feelings about sentencing than they had felt before they took part in the research.[69]

To be able to assess whether or not, and in what ways, judicial sentencing patterns are more aligned with public preferences for the same kinds of cases than members of the public (and media) tend to imagine requires the vital third element: high-quality data about sentencing patterns for different kinds of cases. Where such high-quality data is available to allow a triangulation between public preference, public expectation of sentencing patterns, and the actual reality of sentencing patterns, a potentially potent ‘good news’ story can then be disseminated to policy audiences and through the media. Furthermore, a post-research de-brief discussion with participants may note the change in their feelings: that the reality of sentencing patterns is not nearly as lenient as they had initially believed is reassuring. Of course, as well as a picture of much broader alignment between the public and judicial sentencing than had been imagined, there may also be some specific areas where there is less alignment. Identifying those areas will help to advance the research, public education and guidelines agenda of the SSC.

However, Scotland faces a major obstacle. As things currently stand, it is not possible to draw direct comparisons between, on the one hand, public sentencing preferences and public sentencing expectations of sentencing patterns *and,* on the other hand, the reality of sentencing patterns. So, the opportunity to tell the good news story that sentencing may not be as lenient as people imagine is thwarted. This is because, at present, the quality of the existing data is not

[69] See for example discussion in: Gelb (2008); Hough and Roberts (2023); Hough et al (2013); Roberts et al (2022).

able to tell us meaningful information about the reality of sentencing patterns in specific kinds of cases.

6.2 The Need to Develop New Sources of Data

This report has documented some of the fundamental problems with official data derived from administrative sources. We have seen that such data is marked by fragmentation, meaning that data collected by one agency cannot be easily married with data from another agency. Different agencies operate their own case-counting rules and, understandably, record information which is useful for their own operational purposes.

As we have seen, official data used to attempt to describe sentencing patterns suffers from some fundamental flaws. Ultimately, this should not be surprising. Notwithstanding the laudable efforts of government statisticians and the deployment of advanced statistical techniques[70] to try to extract some insights from it, the quality of that data is a consequence of the fact that it is collected for day-to-day operational purposes of processing individuals through the criminal justice system, rather than as a research tool to inform the development of policy.

As we have also observed, each individual criminal justice agency records information about the individuals it processes, and this can be relevant to sentencing. For understandable reasons, each agency (police, prosecution, social work, courts, prisons, etc) has developed different recording practices (including case-counting rules) according to its needs. The result is a fragmented picture, where it is difficult to marry-up the data from different agencies. It has been a long-held aspiration of successive Scottish administrations to seek to bring all of these disparate administrative data-sets together in a coordinated fashion. This would enable, for example, controlled analyses of sentencing according to specific criteria. It would allow for the effect (or non-effect) of policy changes to be identified. Even more promisingly, it might allow for the follow-through of cohorts of people (according to different key criteria) through the criminal justice system from arrest to charge, conviction, sentencing, sentence implementation, to exit and possible re-entry into criminal justice. This could enable a better understanding of longitudinal effects (including policing, prosecution, sentencing, etc), so as to track journeys of people, according to different criteria, through the criminal justice system.

However, despite the valiant efforts of government statisticians, progress has been (perhaps inevitably) very slow and limited.[71] Official data continues to be marked by fragmentation. Many of the deep-seated problems with official data appear to be long-standing and, in the absence of fundamental system-wide change, intractable. It seems likely that this fragmentation will remain

[70] See Chapter 4.
[71] See Chapter 2, 4 and especially Chapters 3, and 5.

a prominent feature of the Scottish criminal justice landscape for at the very least the medium-term, if not the long-term.

As such, official data tends to be able to offer simple frequency counts, but remains weak in terms of the ability to control for case seriousness (offence and offender features). This means that using such data to attempt to describe the existence of unwarranted disparities is at best weak and at worst could be misleading.[72]

All of this means that the SSC cannot currently draw on comprehensive and in-depth data about current sentence patterns on the basis of specific criteria of interest. Therefore, the SSC faces the unavoidable challenge of devising a means for the systematic collection and analysis of data to support its statutory duties in seeking to improve consistency in sentencing; inform the development of policy; and promote public knowledge of the reality of sentencing patterns. Similar data will later be needed for monitoring compliance with guidelines. While devising a way of gathering higher quality data about the reality of sentencing may be a challenge, it also presents the SSC with the opportunity to develop a method of recording and analysing data which meets its own information requirements.

Continuing to rely on existing official data may be tempting in terms of expediency. However, care should be exercised to communicate the weakness of that data in accurately and meaningfully representing patterns of sentencing. While bodies like the SSC will be well aware of this weakness, others may not be. A potential pitfall is that if others perceive data to better reflect sentencing than it in fact does, (or even incorrectly perceive the SSC as accepting it as an accurate and meaningful representation), then it could place the SSC in a dilemma of having to distance itself from weak and misleading data, which it may have been (even wrongly) perceived to earlier have accepted as valid. Viewed in this light, the development of a bespoke means of collecting data about sentencing patterns, (while still being open to what official data might one day be able to offer), offers the opportunity to provide high-quality data that is essential for: the development of informed policy; the improvement of consistency; and more accurate public knowledge.

6.2.1 Sources of Information the SSC Could Consider

While current official data can provide a very broad picture of frequencies against the broadest of categories, to carry out its statutory duties the SSC could consider additional and complementary approaches to describe sentencing patterns in a more in-depth and meaningful way.

[72] See for example the discussion in Chapter 2 and 3.

These approaches include larger, or smaller, census studies (e.g. CCSS and an SIS).[73] Because they are dedicated to the collection of data which is specifically useful to understanding sentencing patterns in genuinely similar cases, census studies produce much more in-depth and meaningful data. This can be far more useful to bodies such as sentencing councils (Gormley et al 2023).

Key questions which are raised by census approaches are how and by whom the data should be collected. This is less of an issue where there may be archival data collection (e.g. from court papers) than if data collection is to be contemporaneous. If contemporaneous data recording is to be done by judicial sentencers (as was the case with the CCSS) or by clerks (as was the case with the SIS when handed over to SCTS), it is likely there may be some resistance. Here, as well as resourcing, leadership is important in explaining and promoting the reason for requiring the recording of this information (even for a time-limited duration). It is more likely that clerks, and especially judicial sentencers, will be persuaded to take on this additional task if they can see the point of it (i.e. adding benefit to the system in which they work and public confidence in sentencing). Even where a census study is time-limited, it has the potential to leave a positive legacy. For example, as we saw in Chapter 5, the CCSS data continues to be used by SCEW as an invaluable resource.

A more modest alternative to a comprehensive census approach would be to deploy mini-census studies. These could be targeted to help to inform the evaluation and development of policy about specific subjects which SSC is planning to tackle and/or areas of particular public and media concern. This would allow the collection of baseline data against which the impact of particular changes (e.g. guideline) could then be assessed. It would also allow for the patterns of sentencing in specific kinds of cases to be compared directly with public preferences and expectations of what is thought to be typical sentencing practices.

Whichever approaches are adopted, it would be inadvisable for those undertaking efforts to improve the quality of sentencing data to put all their eggs in one basket. While the SSC may wish to continue to encourage improvements in the quality of official data to describe sentencing patterns, at the same time, a range of other approaches that the SSC controls will be needed as well. Consideration should be given to, for example, comprehensive census approaches; targeted and time-limited census approaches (whether archival and/or contemporaneous); court observations about specific kinds of cases; hypothetical sentencing exercises; surveys; interviews; and small simulation exercises. None of these approaches is better than another. They each have their strengths and limitations. Quantitative studies and qualitative studies should be combined to answer ‘what’ and ‘how’ questions.[74]

[73] See Chapter 5.
[74] The use of AI (e.g. Large Language Models) is also worthy of consideration, although with caution given that such work is at a relatively early stage in its development.

In conclusion, none of these approaches has to be, or should be, used by itself. Their strengths are mutually complementary. For example, official data (if very significantly improved) has the potential to be comprehensive and describe sentencing-relevant trends over time. Census studies, surveys and smaller bespoke studies (e.g. hypothetical sentencing exercises, surveys, focus groups and interviews) offer the SSC greater control in addressing specific questions in greater depth and comparability with data about public perceptions. While it is necessary to pursue efforts aimed at the long-term improvement in the quality of official data about sentencing, it is also important, especially in the meantime, to pursue a research strategy which combines the strengths of different approaches, and which would be within the more immediate control of the SSC.

References

- Alberton, A. and K.M. Gorey (2021), 'Structural violence perpetrated against Indigenous peoples in Canadian criminal courts: Meta-analytic evidence of longstanding sentencing inequalities.' *Critical Social Work*, 22(1): 2-22.
- Anthony Doob and Park, N. (1987), 'Computerised Sentencing Information for Judges: An Aid to the Sentencing Process.' *Criminal Law Quarterly* vol. 30.
- Archer, N. Butler, M., Avukatu, G. and Williams Savanta, E. (2022), 'Public knowledge of and confidence in the criminal justice system and sentencing: 2022 research.' *Sentencing Council for England and Wales*.
- Audit Scotland (2015), 'Efficiency of prosecuting criminal cases through the sheriff courts*'* *Audit Scotland*, available at https://audit.scot/uploads/docs/report/2015/nr_150924_sheriff_courts.pdf.
- Bagaric, M. (2016), 'Three things that a baseline study shows don't cause Indigenous over-imprisonment; three things that might (but shouldn't) and three reforms that will reduce Indigenous over-imprisonment.' *Harvard Journal on Racial and Ethnic Justice* 32: 103-151.
- Bales, W.D. and Piquero, A.R. (2012), 'Racial/Ethnic Differentials in Sentencing to Incarceration.' *Justice Quarterly* 29(5): 742–773.
- Biggs, H. Reid, S, Attygalle, K, Vosnaki, K, McPherson, R, Tata, C (2021), 'Public perceptions of sentencing in Scotland: Qualitative research exploring sexual offences.' *Scottish Sentencing Council*.
- Black, C. R. Warren, R. Ormston, and C. Tata (2019), 'Public perceptions of sentencing: National survey report.' *Scottish Sentencing Council*, available at https://www.scottishsentencingcouncil.org.uk/media/2ctkgnpj/20190902-public-perceptions-of-sentencing-report.pdf.
- Bontrager, S. K. Barrick, and E. Stupi (2013), 'Gender and sentencing: A meta-analysis of contemporary research.' *Journal of Gender Race and Justice* 16: 349-372.
- Brandon, A.M. and M. O'Connell (2018), 'Same crime: Different punishment? Investigating sentencing disparities between Irish and non-Irish nationals in the Irish criminal justice system.' *British Journal of Criminology* 58(5): 1127-1146.
- Bushway, S.D, Johnson, B.D. and Slocum, L.A. (2007). 'Is the magic still there? The use of the Heckman two-step correction for selection bias in criminology.' *Journal of Quantitative Criminology,* 23(2):151–178.
- Canadian Department of Justice (2023), 'Overrepresentation of Indigenous people in the Canadian criminal justice system: Causes and responses.' *Canadian Department of Justice,* available at https://www.justice.gc.ca/eng/rp-pr/jr/oip-cjs/toc-tdm.html.
- Carson, E.A. (2023), 'Prisoners in 2002 – Statistical Tables, Bureau of Justice Statistics.' *Bureau of Justice Statistics*, available at https://bjs.ojp.gov/library/publications/prisoners-2022-statistical-tables.
- Chen Q. M. Vuk, C. Kuppuswamy, and D. Kirsch (2023)*, '*Equality and diversity in the work of the Sentencing Council*.' Sentencing Council for England and Wales*, available at

https://www.sentencingcouncil.org.uk/wp-content/uploads/Equality-and-Diversity-Report-FINAL.pdf.

- Cinelli, C. and Hazlett, C. (2020). 'Making sense of sensitivity: Extending omitted variable bias.' *Journal of the Royal Statistical Society Series B: Statistical Methodology* 82(1): 39-67.
- Cinelli, C. Forney, A. and Pearl, J. (2022), 'A crash course in good and bad controls. *Sociological Methods and Research.*' *Sociological Methods & Research*, *53*(3), 1071-1104.
- Crown Prosecution Service (2003), 'CPS charging decisions - examining demographic disparities in the outcomes of our decision making.' *CPS*, available at: https://www.cps.gov.uk/publication/cps-charging-decisions-examining-demographic-disparities-outcomes-our-decision-making.
- Davis, L. (2010), 'Rock, powder, sentencing-making disparate impact evidence relevant in crack cocaine sentencing.' *Journal of Gender, Race and Justice* 14 (2): 375-404.
- Drápal, J. Westermann, H. and Savelka, J. (2023), 'Using Large Language Models to Support Thematic Analysis in Empirical Legal Studies.' *Legal Knowledge and Information Systems* 197-206.
- Engen, R.L. and Gainey, R.R. (2000), 'Modelling the effects of legally relevant and extralegal factors under sentencing guidelines: The rules have changed.' *Criminology,* 38(4):1207–1230.
- Eren, O. and Mocan, N. (2018), 'Emotional judges and unlucky juveniles.' *American Economic Journal: Applied Economics*, 10(3): 171-205.
- Evans, C.R. Leckie, G. Subramanian, S.V. Bell, A. and Merlo, J. (2024), 'A tutorial for conducting intersectional multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA).' *SSM - Population Health,* available at https://doi.org/10.1016/j.ssmph.2024.101664.
- Feilzer, M. and Hood, R. (2004), 'Differences or Discrimination?' *US Department of Justice*, available at: https://www.ncjrs.gov/App/abstractdb/AbstractDBDetails.aspx?id=233677.
- Ferguson, C.J. and S. Smith (2024), 'Race, class, and criminal adjudication: Is the US criminal justice system as biased as is often assumed? A meta-analytic review.' *Aggression and Violent Behavior* 75: 101905.
- Field, S and Tata, C (2023) *'Criminal Justice and the Ideal Defendant in the Making of Remorse and Responsibility.' Hart, Bloomsbury*.
- Franklin, T.W. (2018), 'The state of race and punishment in America: Is justice really blind?' *Journal of Criminal Justice* 59: 18-28.
- Gaebler, J. Cai, W. Basse, G. Shroff, R. Goel, S. and Hill, J. (2022), 'A causal framework for observational studies of discrimination.' *Statistics and Public Policy* 9(1): 26–48.
- Gaudet, F. Harris, G. St John, Ch, 1933. 'Individual differences in the sentencing tendencies of judges. *Journal of Criminal Law and Criminology* 23: 191-208.
- Gelb, K (2008), 'Myths and misconceptions: public opinion versus public judgment about sentencing' in Freiberg, A and Gelb, K (eds) *Penal Populism, Sentencing Councils and Sentencing Policy* (Willan/Routledge).

- Gormley, J and Tata, C (2020), 'To plead? or not to plead? 'Guilty' is the Question. Re-thinking plea decision-making in Anglo-American countries' in Spohn, C and Brennan, P (eds) *Handbook on Sentencing Policies and Practices in the 21st Century* Routledge pp208-234.
- Gormley, J, McPherson, R, Tata, C (2020), '*Sentence Discounting: Sentencing and Plea Decision-Making* Scottish.' Scottish *Sentencing Council*.
- Gormley, J and Tata, C (2022), 'Remorse and Sentencing in a World of Plea Bargaining' in S Tudor, R Weisman, M Proeve, K Rossmanith (Eds) *Remorse and Criminal Justice: multidisciplinary perspectives* Routledge, Taylor and Francis Group, 40-66.
- Gormley, J. Hamilton, M. and Belton, I. (2022). 'The Effectiveness of Sentencing Options on Reoffending.' *Sentencing Council of England and Wales*.
- Gormley, J. Roberts, J.V. Bild, J. and Harris, L. (2020), 'Sentence Reductions for Guilty Pleas: A review of policy, practice and research.' *Sentencing Academy*.
- Guilfoyle, E. and I.D. Marder (2021), 'Using data to design and monitoring sentencing guidelines: The case of Ireland.' *Common Law World Review*, 50(2-3), 103-119.
- Gormley, J. O'Malley, T. Roberts, J. Spohn, C. Tata, C (2023), 'Assessing Approaches to Sentencing Data Collection and Analysis: Final Report Sentencing Guidelines and Information Committee.' *Judicial Council of Ireland*, available at https://judicialcouncil.ie/publication-of-sentencing-data-research-report/.
- Gormley, J. Roberts, J, Pina-Sánchez, J. Veiga, A. Tata, C (2022), 'The methodological challenges of comparative sentencing research.' *Scottish Sentencing Council*.
- Gormley, J, Roberts, J, Tata, C (2025), 'Public Attitudes to Sentences following a Guilty Plea: findings from a mixed methods research project.' *Scottish Sentencing Council*.
- Guilfoyle, E. and Pina-Sánchez, J. (2024), 'Racially determined case characteristics: Exploring disparities in the use of sentencing guideline factors in England and Wales.' *The British Journal of Criminology 65(2), 241–260.*
- Halevy, T. (1995), 'Racial discrimination in sentencing: a study with dubious conclusions.' *Criminal Law Review*, 267–271.
- Healy, D. and D. Griffin (2023), 'Unnesting the Matryoshka doll: An ecological model of probation and parole decision-making in Ireland.' *Journal of Contemporary Criminal Justice*, *39*(1), 75-93.
- Hockaday, C., Lemesheva, Y, Omstone, R, McPherson, R, O'Loughlin, A, Tata, C (2025), 'Public Perceptions of Sentencing Survey 2025: Main Findings Scottish Sentencing Council.' Scottish Sentencing Council.
- Hofer, P. J. Loeffler, C. E. Blackwell, K. and Valentino, P. (2004), 'Fifteen years of guidelines sentencing: An assessment of how well the federal criminal justice system is achieving the goals of sentencing reform.' *Washington, DC: United States Sentencing Commission*.
- Hogarth, J (1988), 'Sentencing Database System: User's Guide.' *University of British Columbia.*
- Hood, R.G., Cordovil, G. (1992), 'Race and Sentencing: a Study in the Crown Court.' Oxford University Press.

- Hopkins, K. N. Uhrig, and M. Colahan (2016a), 'Associations between being male or female and being sentenced in England and Wales in 2015.' *Ministry of Justice*.
- Hough, M. and Roberts, J.V. (2023), 'Public opinion, crime, and criminal justice' in A. Liebling, S. Maruna, L. McAra (eds), *The Oxford Handbook of Criminology*. Oxford: Oxford University Press, 223-242.
- Hough, M, Bradford, B, Jackson, J, Roberts, J (2013), 'Attitudes to Crime and Trust in Justice: findings from the Crime Survey for England and Wales.' *Ministry of Justice*.
- Hunt, J.S. (2023), 'Injustice in the courtroom: How race and ethnicity affect legal outcomes.' in D. DeMatteo and K. Scherr (eds) *The Oxford Handbook of Psychology and Law*, 742-765.
- Hutton, N and Tata, C (2010), 'A Sentencing Exception? Changing sentencing policy in Scotland.' *Federal Sentencing Reporter*, 22 (4): 272-278.
- Hutton, N. Paterson, A, Tata, C, Wilson, J (1996), 'A Sentencing Information for the Scottish High Court of Justiciary: Report of the Study of Feasibility.' *HMSO*.
- Isaac, A. (2020), 'Investigating the association between an offender's sex and ethnicity and the sentence imposed at the Crown Court for drug offences.' *Sentencing Council for England and Wales*.
- Isaac, A. Pina-Sánchez, J. and Montane, A. (2021), 'The impact of three guidelines on consistency in sentencing.' *Sentencing Council for England and Wales*.
- Ivan Potas (2005), 'The Sentencing Information System.' *Australian Law Reform Commission - Reform Journal* 86: 17-23.
- Jacobson, J, Hunter, G, Kirby, A (2016), 'Inside Crown Court: Personal Experiences and Questions of Legitimacy.' *Bristol University Press*.
- Jamieson, F.E. (2013), 'Narratives of crime and punishment: A study of Scottish judicial culture.*' Edinburgh University,* available at: https://era.ed.ac.uk/bitstream/handle/1842/9745/Jamieson2013.pdf.
- Johnson, B. D. (2006), 'The multilevel context of criminal sentencing: Integrating judge-and county-level influences.' *Criminology*, 44(2), 259-298.
- Johnson, B.D. and Lee, J.G. (2013), 'Recent disparity under sentencing guidelines: A survey of recent research and emerging perspectives.' *Sociology Compass* 7(7): 503-514.
- Kemp, S. and Varona, D. (2023), 'Foreign and dangerous? Unpacking the role of judges and prosecutors in sentencing disparities in Spain.' *The British Journal of Criminology* 63(4): 984-1002.
- King, R.D. and M.T. Light (2019), 'Have racial and ethnic disparities in sentencing declined?' *Crime and Justice: A Review of Research* 48: 365-438.
- Knox, D. Lowe, W. and Mummolo, J. (2020), 'Administrative records mask racially biased policing.' *American Political Science Review* 114(3): 619-637.
- Lammy, D. (2017), 'The Lammy Review an Independent Review into the Treatment of, and Outcomes for, Black, Asian and Minority Ethnic Individuals in the Criminal Justice System.' *UK Government*.

- Leclerc, C. and Tremblay, P. (2016), 'Looking at penalty scales: How judicial actors and the general public judge penal severity.' *Canadian Journal of Criminology and Criminal Justice*, 58(3): 354-384.
- Lehmann, P.S. (2022), 'Racial, ethnic, gender, and economic sentencing disparity.' in E. Jeglic and C. Calkins (eds.) *Handbook of Issues in Criminal Justice Reform in the United States* 127-142.
- Lymperopoulou, K. (2024), 'Ethnic Inequalities in Sentencing: Evidence from the Crown Court in England and Wales.' *The British Journal of Criminology*: 64(5): 1189–1210.
- McManus, D (2024), 'Direct Measures in Scotland: A Conflict with the Principles of Justice? PhD Thesis.' *University of Glasgow*, available at: https://theses.gla.ac.uk/84566/4/2024mcmanusphd.pdf.
- Merlo, J. and Wagner, P. (2013), 'The tyranny of the averages and the indiscriminate use of risk factors in public health: a call for revolution.' *European Journal of Epidemiology*, 28(1), 148.
- Ministry of Justice (2019), 'Statistics on Race and the Criminal Justice System: A Ministry of Justice publication under Section 95 of the Criminal Justice Act 1991.' *Ministry of Justice*.
- Ministry of Justice (2022), 'Women and the Criminal Justice System.' *Ministry of Justice.*
- Ministry of Justice (2022), 'Data First: An Introductory User Guide – Harnessing the Potential of Linked Administrative Data for the Justice System.' *Ministry of Justice.*
- Ministry of Justice (2024), 'Statistics on Ethnicity and the Criminal Justice System 2022.' *Ministry of Justice.*
- Mitchell, L. (2018), 'The continuing evolution of race and sentencing research and reviews of this research.' *Journal of Criminal Justice* 59: 29-31.
- Mitchell, O. (2005), 'A meta-analysis of race and sentencing research: Explaining the inconsistencies.' *Journal of Quantitative Criminology,* 21(4):439–466.
- Mitchell, O. and D.L. MacKenzie (2004), 'The relationship between race, ethnicity, and sentencing outcomes: A meta-analysis of sentencing research.' *US Department of Justice,* available at: https://www.ojp.gov/pdffiles1/nij/grants/208129.pdf
- NI Department of Justice (2023), 'Court Prosecutions Convictions and Out of Court Disposals Statistics for Northern Ireland 2022.' *NI Department of Justice,* available at: https://www.justice-ni.gov.uk/publications/court-prosecutions-convictions-and-out-court-disposals-statistics-northern-ireland-2022.
- NI Department of Justice (2023), 'Findings from the 2021/22 Northern Ireland Safe Community Telephone Survey.' NI Department of Justice, available at: https://www.justice-ni.gov.uk/sites/default/files/publications/justice/2021-22%20ni%20safe%20community%20telephone%20survey%20tables.ods.
- Northern Ireland Statistics and Research Agency (2023), 'The Northern Ireland Prison Population 2022/23.' *NI Department of Justice*, available at: https://www.justice-ni.gov.uk/sites/default/files/publications/justice/Northern-Ireland-Prison-Population-2022-23.pdf.

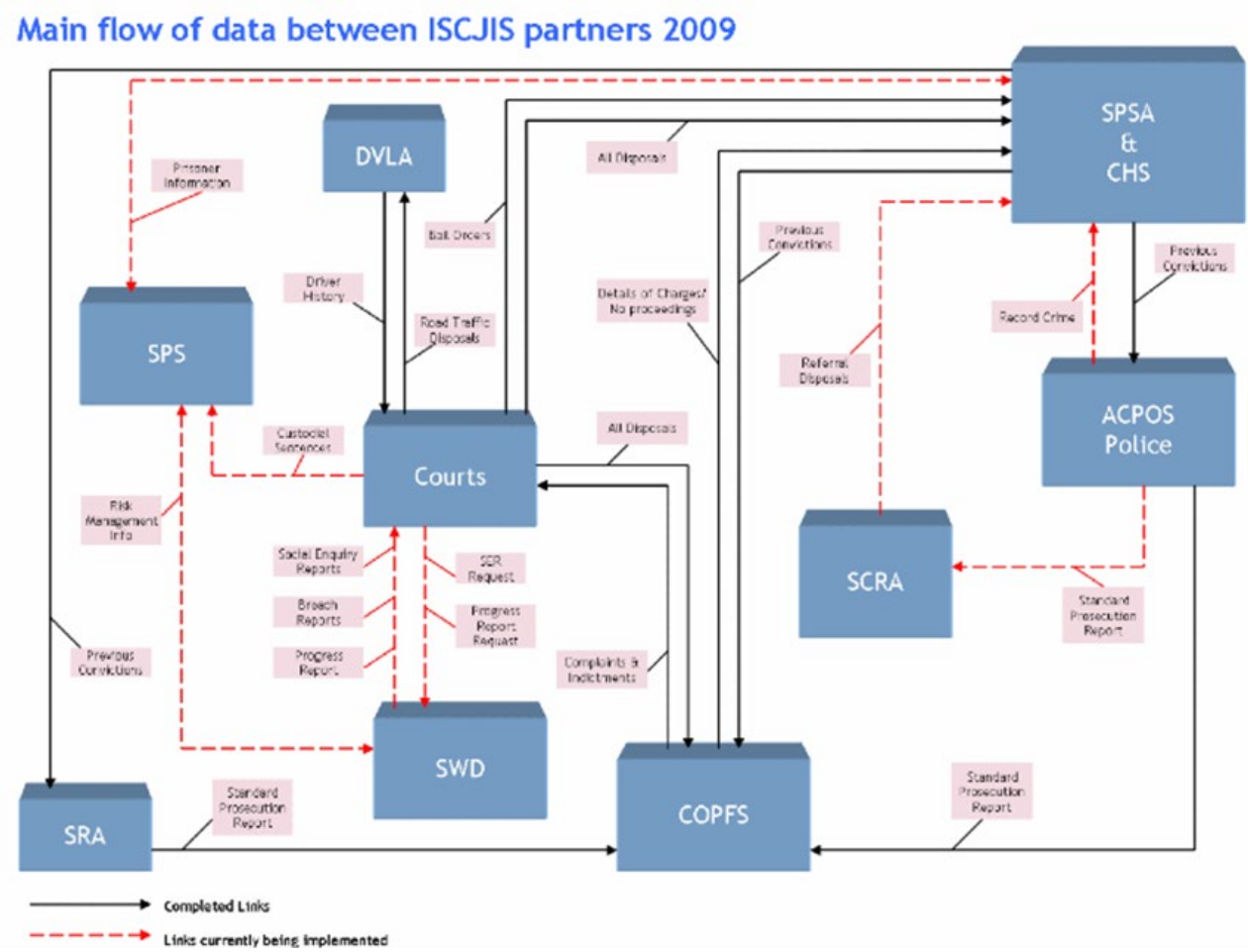
- Nowacki, J.S. (2020), 'Gender equality and sentencing outcomes: An examination of state courts.' *Criminal Justice Policy Review* 31(5): 673-695.
- O'Hara, K. and M. Rogan (2015), 'Examining the use of community service orders as alternatives to short prison sentences in Ireland.' *Irish Probation Journal* 12: 22-45.
- Phillips, C., Bowling, B., and Parmar, A. (2003), 'Ethnicities, Racism, Crime, and Criminal Justice' in A. Liebling, S. Maruna, L. McAra (eds) *The Oxford Handbook of Criminology*. Oxford: Oxford University Press, 329-349.
- Pina-Sánchez, J. and Linacre, R. (2016), 'Refining the measurement of consistency in sentencing: A methodological review.' *International Journal of Law, Crime and Justice*, 44:68–87.
- Pina-Sánchez, J. and L. Harris (2020), 'Sentencing gender? Investigating the presence of gender disparities in Crown Court sentences.' *Criminal Law Review*, 1, 3-28.
- Pina-Sánchez, J. and Gosling, J. P. (2020), 'Tackling selection bias in sentencing data analysis: a new approach based on a scale of severity.' *Quality and Quantity*, 54(3), 1047-1073.
- Pina-Sánchez, J. and Linacre, R. (2013), 'Sentence consistency in England and Wales: Evidence from the Crown Court Sentencing Survey.' *British Journal of Criminology* 53(6): 1118-1138.
- Pina-Sánchez, J. D. Grech, I. Brunton-Smith, and D. Sferopoulos (2019), 'Exploring the origin of sentencing disparities in the Crown Court: Using text mining techniques to differentiate between court and judge disparities.' *Social Science Research* 84: 102343.
- Pina-Sánchez, J. Geneletti, S. Veiga, A. Morales, A. and Guilfoyle, E. (2024), 'Can ethnic disparities in sentencing be taken as evidence of judicial discrimination?' *Journal of Legal Research Methodology* 3(1): 54-82.
- Pina-Sánchez, J. Gosling, J. P. Chung, H. I. Bourgeois, E. Geneletti, S. and Marder, I. D. (2019), 'Have the England and Wales guidelines affected sentencing severity? An empirical analysis using a scale of severity and time-series analyses.' *The British Journal of Criminology* 59(4): 979-1001.
- Pina-Sánchez, J. Grech, D. Brunton-Smith, I. and Sferopoulos, D. (2019), 'Exploring the origin of sentencing disparities in the Crown Court: Using text mining techniques to differentiate between court and judge disparities.' *Social Science Research*, 84: 102343
- Pina-Sánchez, J. Morales, A. Geneletti, S. Veiga, A. and Guilfoyle, E. (2025), 'Testing the interrelationship between area deprivation and ethnic disparities in sentencing.' *Analyses of Social Issues and Public Policy* 25(1), available at: https://doi.org/10.1111/asap.12446.
- Pina-Sánchez, J. Tura, F. and Gómez-Bellvís, A. B. (forthcoming), 'Intersectional disparities in sentencing: An application of MAIHDA to Federal Courts data' in Welch, K. (ed.) *Research Handbook on Race, Crime, and Justice*. Edward Elgar Publishing.
- Pina-Sánchez, J. Dhami, M. and Gosling, J. P. (forthcoming), 'Finding the key case characteristics considered by judges: An application of Bayesian model averaging to theft offences sentenced in the England and Wales Magistrates' Courts' in Fix, M. P. and Montgomery, M. (eds) *Research Handbook on Judicial Politics.* Edward Elgar Publishing.

- Pina-Sánchez, J. Roberts, J. V. and Sferopoulos, D. (2019), 'Does the Crown Court discriminate against Muslim-named offenders? A novel investigation based on text mining techniques.' *The British Journal of Criminology*, 59(3), 718-736.
- Poppleton, S, E. Wedlock, A. Isaac, and E. Marshall (2021), 'A review of consistency in sentencing. Sentencing Council for England and Wales.' *Sentencing Council for England and Wales*.
- Pratt, J. (2017), 'New Zealand penal policy in the twenty-first century' in A. Deckert and R. Sarre (eds.) *Palgrave Handbook of Australian and New Zealand Criminology, Crime and Justice* 347-361.
- Provine, D (2002), 'Sentencing Policy and Racial Justice' in C Tata and N Hutton (eds) (2002) *Sentencing and Society: International Perspectives,* Ashgate/Routledge 483 – 503.
- Scottish Government (2021), 'Recorded Crime in Scotland, 2020-2021.' *Scottish Government*, available at: https://www.gov.scot/publications/recorded-crime-scotland-2020-2021.
- Ri, S. and Cheng, K. K. Y. (2024), 'Criminal record and sentencing: A comparative perspective between England and Wales and Hong Kong.' *Journal of Criminal Justice*, 92(2): 102193.
- Richard S. Frase, 'Sentencing Principles in Theory and Practice.' *Crime and Justice* 22: 366.
- Roberts, J and Bild, J (2021), 'Ethnicity and Custodial Sentencing: A Review of the Trends 2009-2019.' *Sentencing Academy*.
- Roberts, J. V. and Pina-Sánchez, J. (2014), 'Previous convictions at sentencing: Exploring empirical trends in the Crown Court.' *Criminal Law Review* 8: 575-88.
- Roberts, J, Bild, J, Pina- Sánchez, J, Hough, M (2022), 'Public Knowledge and Public Opinion about Sentencing.' *Sentencing Academy*.
- Rossmanith, K (2023), 'Cultural Sensitivity Training, Judicial Feeling and Everyday Practice: conversations at the edge of research' in S. Field and C. Tata (eds) *Criminal Justice and the Ideal Defendant in the Making of Remorse and Responsibility* Hart, 111-126.
- Scottish Government (2022), 'Women in the Justice System: Evidence Review.' *Scottish Government*, available at https://www.gov.scot/binaries/content/documents/govscot/publications/research-and-analysis/2022/01/women-justice-system/documents/women-justice-system/women-justice-system/govscot%3Adocument/women-justice-system.pdf.
- Scottish Government (2023a), 'Occasional Paper: Analysis of the ethnicity of individuals subject to hearings in Scottish Courts.' *Scottish Government.*
- Scottish Government (2023b), '*Criminal Proceedings in Scotland, 2021-22.*' *Scottish Government,* available at https://www.gov.scot/publications/criminal-proceedings-scotland-2021-22/pages/17/.
- Scottish Government (2024), 'Scottish Prison Population Statistics 2023-24.' *Scottish Government,* available at: https://www.gov.scot/publications/scottish-prison-population-statistics-2023-24/.

- Scottish Sentencing Council (2018), '*Sentencing Guideline: Principles and purposes of sentencing* 1–5.' *Scottish Sentencing Council.*
- Scottish Sentencing Council (2021), 'Sentencing Guideline: The Sentencing Process.' *Scottish Sentencing Council.*
- Sentencing Council for England and Wales (2015), 'Crown Court Sentencing Survey Annual Publication: January to December 2014.' *Sentencing Council for England and Wales*.
- Sentencing Council for England and Wales (2020), 'Investigating the Association between an Offender's Sex and Ethnicity and the Sentence Imposed at the Crown Court for Drug Offences.' *Sentencing Council for England and Wales.*
- Sentencing Council for England and Wales (2023), 'Equality and diversity in the work of the Sentencing Council.' *Sentencing Council for England and Wales, available at https://sentencingcouncil.org.uk/publications/research-reports/equality-and-diversity-in-the-work-of-the-sentencing-council/.*
- Shankley, W. and P. Williams (2020), 'Minority ethnic groups, policing and the criminal justice system in Britain' in B. Byrne et al (eds.) *Ethnicity, Race and Inequality in the UK: State of the Nation* Bristol University Press 51-71.
- Sklansky, D.A. (1995), 'Cocaine, Race, and Equal Protection.' *Stanford Law Review* 47(6): 1283–1322.
- Spohn, C. (2017), 'Race and Sentencing Disparity' in E. Luna (ed.) *Reforming Criminal Justice* (vol.4), available at: https://academyforjustice.asu.edu/wp-content/uploads/2022/04/Reforming-Criminal-Justice_Vol_4.pdf
- Sporer, S.L. and J. Goodman-Delahunty (2009), 'Disparities in sentencing decisions' in Oswald, M.E. S. Bieneck, and J. Hupfeld-Heinemann (eds.) S*ocial Psychology of Punishment of Crime* 379-401.
- Stockton, B. Strange, C.C. and Harel, O. (2023), 'Now you see it, now you don't: A simulation and illustration of the importance of treating incomplete data in estimating race effects in sentencing.' *Journal of Quantitative Criminology* 40(3): 1-28.
- Sturge, G. (2023), 'UK Prison Population Statistics.' *House of Commons Library*, available at: https://researchbriefings.files.parliament.uk/documents/SN04334/SN04334.pdf.
- Tata, C. (1997), 'Conceptions and Representations of the Sentencing Decision Process.' *Journal of Law and Society* 24(3): 395-420.
- Tata, C. (2000) 'Resolute Ambivalence: Why Judiciaries Do Not Institutionalize Their Decision Support Systems' *International Review of Law, Computers and Technology* 14(3): 297-317.
- Tata, C (2020), 'Sentencing: A Social Process – Re-thinking Research and Policy.' *Palgrave Springer*.
- Tata, C (2023), 'Punishment and the 'blind symbiosis' of legal and rehabilitation work in the making of the 'ideal' defendant'' in S Field & C Tata (eds), *Criminal Justice and the Ideal Defendant in the Making of Remorse and Responsibility*. Hart 213-237.

- Tata, C, and Hutton, N (2003), 'Beyond the Technology of Quick Fixes. Will the Judiciary Act to Protect Itself and Short Up Judicial Independence? Recent Experience from Scotland.' *Federal Sentencing Reporter* 16(1): 67–75.
- Tata, C, Burns, N, Halliday, S, Hutton, N, McNeill, F (2008), 'Assisting and Advising the Sentencing Decision Process: the pursuit of 'quality' in pre-sentence reports.' *British Journal of Criminology* 48(6): 835-855.
- Tonry, M. (1997), 'Ethnicity, crime, and immigration.' *Crime and Justice* 21: 1-29.
- Uhrig, N. (2016), 'Black, Asian and Minority Ethnic Disproportionality in the Criminal Justice System in England and Wales.' *Ministry of Justice.*
- Ulmer J.T. and Bradley, M.S. (2019), 'Criminal justice in Indian country: A theoretical and empirical agenda.' *Annual Review of Criminology* 2: 337-357.
- VanderWeele, T. J. and Staudt, N. (2011), 'Causal diagrams for empirical legal research: a methodology for identifying causation, avoiding bias and interpreting results.' *Law, Probability and Risk* 10(4): 329-354.
- VanderWeele, T.J. and Ding, P. (2017), 'Sensitivity analysis in observational research: introducing the E-value.' *Annals of internal medicine*167(4): 268-274.
- Veiga, A. Pina-Sánchez, J. and Lewis, S. (2023), 'Racial and ethnic disparities in sentencing: What do we know, and where should we go?', *The Howard Journal of Crime and Justice*, 62(2) : 167–182.
- Velasquez , J. (2018), 'Doing justice: Sentencing practices in Scottish sheriff courts.' *Glasgow University*.
- Velasquez, J (2024), 'Sentencing individuals on cusp-cases: The use of offenders' backgrounds by Scottish Sheriffs.' *Howard Journal of Crime and Justice* 63(4): 304-320.
- Ward, J.T. Hartley, R.D. and Tillyer, R. (2016), 'Unpacking gender and racial/ethnic biases in the federal sentencing of drug offenders: A causal mediation approach.' *Journal of Criminal Justice,* 46:196–206.
- Wilbanks, W. (1987), 'The Myth of a Racist Criminal Justice System.' *Brooks/Cole, Monterey.*
- Wu, J. and Spohn, C. (2009), 'Does an offender's age have an effect on sentence length? A meta-analytic review.' *Criminal Justice Policy Review* 20(4): 379-413.
- Zatz, M. S. (1987), 'The changing forms of racial/ethnic biases in sentencing.' *Journal of Research in Crime and Delinquency* 24(1): 69-92.

Appendix A: Data Flows Between Scottish Criminal Justice Organisations



Scottish Government, 'Main flow of data between ISCJIS partners 2009.' <https://www.webarchive.org.uk/wayback/archive/20150218221752mp_/http://www.gov.scot/Resource/Doc/254431/0097170.pdf>

Appendix B: Extract from a Data 'Snapshot'

Sentence types and trends

Figure 2 shows the proportion of people who received a custodial or non-custodial sentence for the principal offence of sexual penetration of a child under 12.

A custodial sentence involves at least some element of immediate imprisonment or detention.[7] The rate of custodial sentences was lowest in 2016–17 (85.7%) and highest in 2015–16 (94.4%). Over the five-year period, 90.6% of people were given a custodial sentence.

Figure 2: The percentage of people who received a custodial sentence and non-custodial sentence for sexual penetration of a child under 12 by financial year

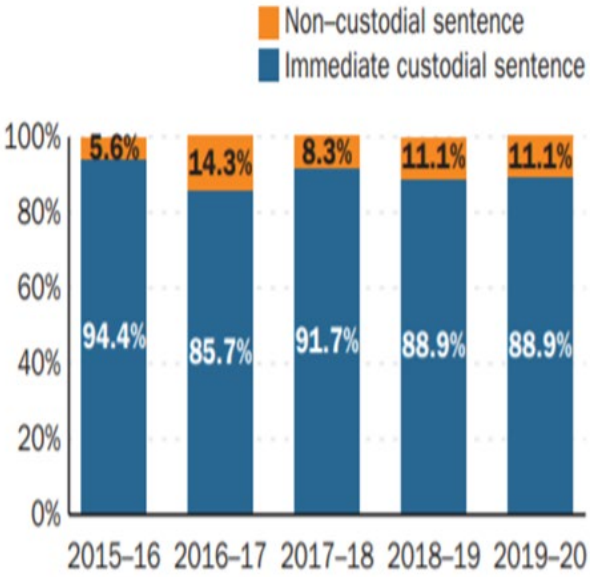


Table 1 shows the principal sentence imposed for sexual penetration of a child under 12 from 2015–16 to 2019–20.[8] The *principal sentence* is the most serious sentence imposed for the charge that is the principal offence.[9] The availability of different sentence types has changed over time. Most notably, wholly and partially suspended sentences have now been abolished.[10] Changes to community correction orders may have also influenced the sentencing trends over the five years covered by this Snapshot.[11]

Over the five-year period, most people sentenced for sexual penetration of a child under 12 received a principal sentence of imprisonment (87.5% or 56 of 64 people). The rate of imprisonment sentences was highest in 2015–16 (94.4%) and lowest in 2016–17 (71.4%). All offenders whose offence attracted standard sentence classification received a sentence of imprisonment.

Source: Sentencing Advisory Council (2021) Sentencing Advisory Council, Victoria, Australia

Appendix C: the Authors of this Report

Prof Cyrus Tata

Professor Cyrus Tata, PhD, FRSA is Professor of Law and Criminal Justice at the Law School, University of Strathclyde, Scotland. Prof Tata has conducted single and multi-jurisdiction studies on a wide range of sentencing and criminal justice issues including: the measurement of sentencing consistency; the relationship between plea and sentencing; the impact of legal aid reforms on case outcomes; plea decision-making and negotiation; executive release from prison; pre-sentence reports and mitigation practices; and the role of emotion in criminal justice and sentencing. In recent years, he has conducted several studies of public perceptions of sentencing (including in cases of sexual assault, death by driving, guilty pleas etc).

He has considerable experience of conducting empirical research studies, both in Scotland and abroad. Cyrus has a particular interest in and has published several works about the quality of data about sentencing and issues in attempts to measure sentencing practices. For example, he recently led a comparative research study examining the quality of sentencing data in different jurisdictions funded by *the Irish Judicial Council.*

Cyrus is founder and chair of the *European Group on Sentencing and Penal Decision-Making* (a network of some 130 academic, policy and practitioner members in over 25 countries. Regularly invited to speak to judicial, governmental policy and practice audiences around the world, he has served as Adviser to governments and judiciaries in several countries.

Cyrus' recent academic works on sentencing include: *Sentencing: a Social Process – Re-Thinking Research and Policy* (Palgrave Springer, 2020) and (with S Field) *Criminal Justice and The Ideal Defendant in the Making of Remorse and Responsibility* (Hart Bloomsbury, 2023)

www.strath.ac.uk/staff/tatacyrusprof/

Dr Jay Gormley

Dr Jay Gormley is a Lecturer in Law and Criminal Justice at the University of Glasgow; an *Adviser* to the Sentencing Academy; and a Member of the Scottish Sentencing Council. He is an expert in sentencing and socio-legal research who works regularly with guideline-creating bodies and key policy influencers: including the *Scottish Sentencing Council*; *the English and Welsh Sentencing Council*; and the *Sentencing Guidelines and Information Committee of the Judicial Council* (Ireland).

Jay is a skilled empirical researcher with a strong understanding of criminal and evidential matters applicable to multiple jurisdictions. He is trained in both advanced qualitative and quantitative methods and is an *ONS Accredited Researcher*. His empirical research provides him with distinctive experience working with various stakeholders: such as sentencers and legal professionals; offenders and victims/survivors; and members of the public.

Jay has published numerous journal articles and book chapters on the topic of sentencing and his work has been used to evaluate critical aspects of sentencing, data, and case proceedings. He was an expert witness in the High Court whose work was cited by Lady Dorrian in *HM Advocate v B(L)* 2023 SCCR 64. His recent research on the effectiveness of sentencing has aided the work of the Sentencing Council for England and Wales.

www.gla.ac.uk/schools/law/staff/jaygormley/

Prof Melissa Hamilton

Professor Melissa Hamilton, JD, PhD is Professor of Law and Criminal Justice at the University of Surrey. She is a Surrey AI Fellow with the *Surrey Institute for People-Centred Artificial Intelligence*, a Fellow with the *Royal Statistical Society*, a member of the *American Psychological Association*, a member of the *International Corrections and Prisons Association*, and a licensed attorney. Professor Hamilton has a Juris Doctorate (advanced degree in law) and a PhD in Criminology. Skills include quantitative methods, qualitative research, legal analysis, and forensic sciences. She has published papers in dozens of high- quality law journals and scientific periodicals.

Hamilton has served as an ongoing consultant/expert on various issues related to sentencing and about studying sociodemographic disparities across criminal justice decision points. She has been an invited expert speaking on these issues before the *Ministry of Justice*, *Sentencing Academy*, U*.S. Department of Justice*, *U.S. Congress*, *U.S. Sentencing Commission*, and the *Korean Sentencing Commission*. Hamilton has regularly worked with international and interdisciplinary teams to deliver impactful projects to commissioning stakeholders in criminal justice.

https://www.surrey.ac.uk/people/melissa-hamilton

Prof Jose Pina-Sánchez

Professor Jose Pina-Sánchez, PhD, is a Professor in Quantitative Criminology, Coordinator of the Social Research Methods centre at the University of Leeds, and Director of Advanced Quantitative Methods at the White Rose Doctoral Training Partnership. He has a PhD in Social

Statistics from the University of Manchester, and before joining Leeds he worked as an ESRC intern at the Sentencing Council for England and Wales and as a Fellow in Statistics at the London School of Economics.

Jose has co-authored more than 30 peer review articles on the subjects of sentencing and quantitative methods, most of them published in the top journals like: *Justice Quarterly*, *the Journal of Quantitative Criminology*, or *Survey Research Methods*. He has also led research consortiums alongside partners from the *Crown Prosecution Service*, *the Parole Board*, *the Race Unit at the Cabinet Office*, *the Sentencing Academy* and *the Sentencing Council for England and Wales*.

In his latest project funded by the ESRC Jose led a team of statisticians, criminologists and criminal lawyers to explore the presence of ethnic and class disparities in sentencing in England and Wales. Through this project Jose's team has been able to identify a series of key findings that are being considered in future efforts to reduce ethnic disparities through modifications of the sentencing guidelines. For example: i) although ethnic disparities are detected in the Crown Court, these are markedly concentrated in drug offences, hence a blanket approach is not justified; ii) similarly, class disparities are also identified, but only for sex and assault offences; iii) some of the ethnic disparities observed in sentence outcomes could be redressed by reforming key personal mitigating factors like remorse, good character, or potential for rehabilitation, which appear to be inappropriately applied in favour of White offenders.

https://essl.leeds.ac.uk/law/staff/222/dr-jose-pina-sanchez

Scottish Sentencing Council

Scottish Sentencing Council
Parliament House
Parliament Square
Edinburgh
EH1 1RQ

sentencingcouncil@scotcourts.gov.uk

scottishsentencingcouncil.org.uk

ISBN: 978-1-912442-79-9