Please cite the Published Version

Developing Methods to Measure Sentence Severity and Assess Disparity in Sentencing in the Criminal Courts of England and Wales

Stephanie Julia Wallace, MRes, BA(HONS)

Submitted for the degree of Doctor of Philosophy

Applied Social Statistics

Lancaster University

December 2014

Declaration

I declare that the work in this thesis has been done by myself and has not been submitted elsewhere for the award of any other degree.

Stephanie Julia Wallace

Acknowledgements

Abstract

This doctoral research focuses on sentencing, and more specifically, disparity in sentencing in the courts of England and Wales. Previous research in this area has introduced bias into the samples by focusing solely on custodial sentences and therefore failing to consider the more common non-custodial sentences. This thesis overcomes this methodological problem.

There are two parts to this research: the first part develops a sentence severity scale based on the principle of proportionality, and includes the full range of sentences meted out by the courts in England and Wales. The second part of this research then uses the new scale to assess disparity based on the extra-legal factors of the case. Data from the Offender Assessment System (OASys) was merged with data from the Police National Computer (PNC), which provided information on the offence and resultant sentence, as well as information relating to the socio-demographic characteristics of the offender.

Goodman Row Column Association analysis was used to model the association between the offences and sentence categories. This model was then extended to control for three legal factors - offence plea, previous history of offending, and the number of offences the offender was sentenced for - which must all be taken into consideration when sentencing offenders. This analysis provided a series of scores for each of the sentence categories allowing them to be arranged by magnitude, ranging from the least to most severe sentence. To anchor the scale and make the scores less arbitrary, linear

interpolation was used. This estimated the equivalent number of days in custody for the non-custodial sentences using the modal days in custody for each of the custodial sentence categories.

The new sentence severity scale was treated as the dependent variable in which to model sentencing disparity using multilevel modelling. Here a number of legal and extra-legal variables were considered to try and explain the large amount of sentencing variation between offenders.

After controlling for the legal factors of the offence this research found there to be significant sentencing disparity. Consistent with previous research, female offenders were sentenced less severely than male offenders, and White British and Irish offenders were also sentenced less severely than 'like-situated' Black offenders. However, offenders in the White other category were sentenced less severely than the White British and Irish group. This research also found that unemployed offenders were sentenced more severely than those in employment prior to sentencing, although the research did find that those on benefits were sentenced less severely. Finally, offenders who do not live in permanent accommodation were also sentenced more severely.

The results of this research suggest significant disparities in the sentencing of offenders in England and Wales.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1   Introduction - Setting the scene

One of the basic precepts of justice is to 'treat like cases alike and treat different cases differently' (Ashworth, 2010b). This can be said to be based on a model of parity which is founded on the underlying notions of equality and fairness. Treating like cases differently may result in disparity which, can be argued, leads to discrimination, inequality, and injustice (Nagel and Johnson, 1994). Disparity can undermine a criminal justice system, as well as lead to unpredictable sentences (Tak, 1995). Some would then argue that the increase in the severity of sentencing is partly responsible for the rising prison population (Hough and Jacobson, 2008).

However, this basic precept of justice is widely debated, and so too is this pathway to achieving justice. For example, what resonates as equality and parity is widely debated, especially by some feminist criminologists who argue that treating men and women the same does not ensue equality or justice (Hudson, 1998). Additionally, different sentences can affect offenders differently (Tonry, 1996). Therefore what is required is a better understanding of these terms (Hudson, 1998), such as taking parity in sentencing to mean an equivalent sentence rather than the same sentence

for example.

For some, disparity can result in discrimination because legally irrelevant characteristics of the offender affect the sentence imposed after the legally relevant characteristics of the offence(s) are taken into consideration (Spohn, 2009). I would argue that disparity occurs when, after controlling for the legal factors of the case, offenders receive sentences which are not of equivalent severity. I argue that certain sentences overlap in terms of their severity and are therefore equivalent sentences, even though they possess different punitive elements. Therefore I would argue that disparity occurs when legally irrelevant characteristics of the offender affect the severity of sentences differently. For example, when an offender is sentenced more severely on the basis of being Black, compared to an offender that is White, having controlled for the legal characteristics of the case. These legally irrelevant factors (or extra-legal factors as they are generally referred to) are at the centre of a vast array of literature and research, and can be grouped into four main areas; gender, ethnicity, characteristics of the presiding judge, and the geographic location of the court. However, there is a lack of research which looks at the effects of additional extra-legal characteristics of the offender, such as the effects of employment and accommodation status on sentencing decisions.

A number of studies have looked at disparity with regards to the sentencing of male and female offenders, such as Carlen (1988), Hedderman and Gelsthorpe (1997), and Spohn (2009). There is evidence to suggest that, on the whole, men are sentenced more severely than women (Ministry of Justice, 2010): men are more likely to be incarcerated and face longer custodial sentences than like-situated female offenders (Dowds and Hedderman, 1997), (Albonetti, 1997). There is also evidence to suggest that judges are reluctant to give women offenders a fine due to their inability to pay and therefore more likely to give a discharge or community sentence (Hedderman and Gelsthorpe, 1997). This is an area of great contention among those interested in sen-

tencing and equality, as the Equality Act (2010) states no-one should be discriminated against because of their gender. However, many feminist academics such as Barbara Hudson (1998) have argued that this disparity between the sentencing of men and women is fair. Her argument is that equality in sentencing does not mean sameness, she views fairness and equality as treating men and women differently to reflect their existence within society.

Race and ethnicity is also a focus in the literature on sentencing disparity. However, this literature is less contentious as there is a general consensus that people should not be treated differently based on their race or ethnicity. The over-representation of ethnic minority offenders in prison is the first 'warning signal' of sentencing disparity. However sentencing is just one stage of the criminal justice system in which Black offenders are discriminated against (Ashworth, 2010b). Additional research has shown that ethnic minority offenders are sentenced more severely e.g. Hood (1992) in England and Wales and Bushway and Piehl (2001) in the US.

In addition, research also finds that the characteristics of the judge, specifically their age and ethnicity, leads to sentencing disparity. Johnson (2006) found older judges and judges from ethnic minorities tend to be more lenient towards offenders. Judges can also vary in their philosophical approaches to sentencing, which is likely to impact on the sentencing of offenders, especially when discretion is still an integral part of the English justice system.

The final area in which there is ample research is the geographic location of the court. Studies in England and Wales have shown disparity based on the geographic location of the court in which the offender is sentenced. Mason et al. (2007) reported large variation in custody rates and average custodial sentence lengths across the 42 Criminal Justice Areas in England and Wales. For example, the average custodial sentence length in the magistrates' court was 3 months and this varied between 2.4 and 3.6 months, and in the Crown Court the average custodial sentence length was

24.1 months and this varied between 19.5 and 28.3 months (Mason et al., 2007). Research from the U.S. also suggests sentencing variation, based on the location of the court as well as the available jail capacity (Johnson, 2006).

All these studies provide empirical evidence of disparity in sentencing and are described in detail in Chapter 3. However, there are methodological problems with these studies, which this research seeks to rectify in order to more accurately measure the amount of disparity occurring in the courts in England and Wales.

1.2   Problems with measuring severity

To quantitatively assess disparity in sentencing some form of measurement or scale is required, but up to now there has been no widely accepted standard measure of sentence severity (Sweeten, 2012). A number of methodologies have been employed - the most common measures being; either the 'in-out decision' to incarcerate or the length of the custodial sentence. These methods introduce bias into the samples: for the first measure by excluding the detail of sentences, which is the result of using a binary outcome of receiving a custodial sentence (in) or not receiving a custodial sentence (out), and the second measure by using just the length of that custodial sentence, in which case all the detail is lost. This results in using only a very small proportion of the available data. For example, according to the 2009 sentencing statistics, only 7% of sentences resulted in a custodial sentence (Ministry of Justice, 2010). By employing either method just mentioned, a vast amount of detail in the data is excluded. The presence of such detail (non-custodial sentences) is essential in looking at measuring sentence severity and disparity, as these sentences are just as important (von Hirsch et al., 1989).

A third way that sentence severity has been measured in the past is through the use of scales. On the whole these scales generally incorporate all types of disposals

available to the courts. However, these scales have often been developed on the basis of a vignette approach - presenting survey participants with hypothetical sentences or scenarios and asking them to attach scores relating to their perceived severity. This makes these types of scales highly subjective. The scales used in this area will be discussed fully in Chapter 4.

This research adopts von Hirsch's (1992) concept of ordinal proportionality, and becomes the basis of creating a scale to measure sentence severity and facilitate an assessment of sentencing disparity. This assumes that the range of sentencing outcomes can be placed on a unidimensional scale which measures sentence severity. The different types of sentences are therefore aligned on one continuum to allow differences in the severity of sentences to be measured. Marinos (2005) opposes this unidimensional approach, arguing that a more multidimensional approach is needed to capture the multifaceted nature of sentencing, and which incorporates the many qualitative, as well as, quantitative factors that are taken into consideration when sentencing offenders, such as the age and the "appropriate level of denunciation"(Marinos, 2005, p. 451). However, using a unidimensional scale does not to deny the multifaceted nature of sentencing - sentence severity is just one aspect or dimension of sentencing which allows measurement to occur.

1.3   Thesis aims and objectives

Issues with previous ways of measuring sentence severity and assessing sentencing disparity have partly fuelled the motivation to conduct this research. This research fills this methodological gap by using sentencing information on all types of sentences to reduce sample bias in the data used to measure sentence severity. In contrast to the majority of the literature and research, a scale of sentence severity is constructed for all of the disposals currently in use in the England and Wales justice system,

and thus provides a better measure of sentence severity. This scale then enables the investigation into sentencing disparity.

This research investigates the sentencing practice of judges by analysing real sentencing data on offenders sentenced at the magistrates' and Crown Courts in England and Wales between March 2008 and April 2010. Data from the Offending Assessment System (OASys) merged with downloads from Police National Computer (PNC) facilitate this work.

OASys data comes in the form of pre-sentence reports requested by the courts and produced by the Probation Service to aid decision-making. OASys is described as an actuarial tool for assessing the risks and needs of the offender and records information relating to the social and economic characteristics of the offender. The PNC data then provides information on the offence, disposal and criminal histories of the offender. Both datasets will be explained fully in Chapter 5.

1.3.1 Aims

There are two aims of this thesis. The first aim of this research is concerned with the distribution of sentences and whether the distribution of these sentences can be arranged on a single scale which measures sentence severity. In doing so it is also possible to assess the interchangeability of sentences: to identify sentences that carry the same penal bite or equivalence in terms of their severity. For example, an onerous community order and a short custodial sentence may be equivalent in terms of their severity.

The unidimensional scale is constructed using a form of log-linear model which models the association between the two categorical variables; offence and sentence. This is an extension of the Goodman Row Column Association Model which allows covariates to be added to the association parameters, therefore controlling for legal variables of the case. The outcome is a series of scores which form a unidimensional

scale of sentence severity.

The second aim of this thesis then uses this unidimensional scale as a tool which robustly measures sentence severity to identify disparity in sentencing. It provides evidence of disparity through the use of a multilevel model with a range of legal and extra-legal covariates.

These aims set out to answer a number of research questions:

1. Can sentences be arranged into a unidimensional scale of sentence severity?

    (a) Is it possible to assess the equivalence/interchangeability of sentences?

2. Do extra-legal characteristics of the offender affect sentence severity? And more specifically;

    (a) Are men sentenced more severely than women?

    (b) Does ethnicity impact on sentence severity?

    (c) What other extra-legal factors are affecting sentencing disparity?

3. If certain types of offenders are discriminated against, what are the implications of this?

1.3.2 Objectives

This thesis sets out to achieve a number of research objectives. The first research objective is methodological, and sets out to develop a methodology which allows for the construction of a severity scale which incorporates the full range of sentencing outcomes. It will specifically investigate the Goodman Row Column Association model and whether it is possible to create a unidimensional scale of sentence severity which also controls for legal covariates in constructing this scale, to begin to answer the first research question. The final objective will be to utilise the scale developed in the first stage of this research and use real sentencing data to identify disparity, and

thus answer the second research question. All of this will then facilitate discussion on the last research question.

1.4 Thesis structure

There are ten chapters in this thesis. The following chapter (Chapter 2) provides a brief overview of the theoretical approaches to sentencing which influence the current sentencing framework and its principles underpinning sentencing practice. It will provide an account of sentencing practices as described in the legislative statues and provisions, as well as provide a contextual account of sentencing - including the decision-making process and the types of disposals available to the courts. This information is used later on in the thesis to justify the decisions made during the modelling stages and also to assess if extra-legal factors are influencing sentencing decisions.

Chapter 3 reviews the literature on disparity in sentencing, drawing from literature and research from England and Wales, as well as, the U.S. Disparity can be seen to go against the notion of justice, and therefore it is essential that this is investigated. Research to date reveals a great deal of disparity in sentencing, identifying the socio-demographics of the offenders (particularly gender and ethnicity), characteristics of the judge, and location of the court as most associated with disparity.

I then go on to discuss the problems with the ways in which these studies measure severity and how severity is often measured; namely that bias is introduced into the sample by excluding masses of important information relating to non-custodial sentences. Chapter 4 then reviews the literature surrounding measuring severity and goes on to discuss how disparity is measured and why this needs to be made a focus.

The fifth chapter describes the data used to conduct the research. It explains this research uses data from the Offending and Assessment System (OASys), recorded by the Probation Service. Specifically, the data comes from pre-sentence reports

carried out by the Probation Service prior to sentencing, and by request of the court. The information included in this data includes the specifics about the offence, for example whether the offender used a weapon, and information relating to the offender, including socio-demographic information.

The OASys data was then merged with data from the Police National Computer (PNC) post-sentencing. The courts upload information about the outcome of cases - i.e. the sentences - to their court's computer systems and this information is then fed into the PNC. Therefore, the data used for this research uses the combined data from OASys and the PNC. The remainder of Chapter 5 provides a description of the data and the sample.

Chapter 6 discusses the methodology employed to create the unidimensional scale. It discusses the concept of log-linear modelling, followed by the various other methods that could have been used to create the scale. Finally, Chapter 6 discusses the Goodman Row Column Association analysis employed to create the scale. This models the association between the offence and sentence categories. I go on to explain how I extended the model to allow for the inclusion of legal covariates relating to the offender's plea, their prior offending history and the number of offences they were sentenced for at the time of sentencing. I then explain the use of linear interpolation on the scores which then serves as the scale in which to measure sentence severity.

Chapter 7 is the first results chapter, which discusses the results from the extended Goodman Row Column Association analysis. It reports the scores for the sentence categories.

Chapter 8 is the second methodology chapter which discusses multilevel modelling, and how it is used to identify disparity in sentence severity. The linear interpolated scores from the extended Goodman RC analysis are treated as continuous and this becomes the dependent variable within the random intercept multilevel model. The models included a three-level model with offences at level one, offenders at level two

and police force area at the third level, and a two-level model with police force area excluded. The second model is preferred to the first.

The results from the multilevel modelling are reported in Chapter 9. The final model found disparity in sentencing relating to the following legal factors; plea, conviction history, the number of offences they were sentenced for, and the type of court they were sentenced in, as well as the following extra-legal factors relating to the offenders sex, ethnicity, age, accommodation status, employment status and income.

Chapter 10 provides the discussion which relates the findings from this research with previous research. It then also re-addresses the research questions and draws the main conclusions from this work.

Chapter 2

Sentencing in England and Wales

2.1 Introduction

The core of this doctoral thesis is to assess the disparity in sentencing in the Criminal Courts of England and Wales. In order to do this, is it necessary to clearly set-out what the correct course of action should be (Ashworth, 1987). This chapter therefore provides the background and context of sentencing in England and Wales. However, it will become clear that the legislation and guidelines are somewhat less than clear and at times ambiguous, which may contribute towards disparity (ibid). The following chapter will return to the literature on disparity.

In recent years, great emphasis has been placed on consistency in sentencing as a means of achieving justice and fairness. Consistency can be viewed as optimal and is a central feature of sentencing made apparent within the statutory provisions and Sentencing Council objectives. The main focus being on consistency in the approach to sentencing as opposed to consistency in the outcomes (Hutton, 2013). By establishing a consistent approach to sentencing this should in turn reduce any disparity and subsequent discrimination (ibid). Often there is confusion between these two forms of consistency, as well as a lack of definition as to what is meant by consistency.

In the following chapter I will explain that although the term consistency is used repeatedly in the literature and by, for example the Sentencing Council, they are essentially referring to parity. However, I will return to these notions of consistency and parity in terms of what this means for assessing disparity in the next chapter. This chapter will first provide the legislative context for sentencing and how it should be undertaken.

2.1.1   The role of sentencing

Today punishment is described as "the legal process whereby violators of criminal law are condemned and sanctioned in accordance with specified legal categories and procedures" (Garland, 1990, 17). This is administered by professional and lay judges in a court of law, by way of sentencing. However, the ways in which offenders are sentenced has changed. The opening scene from Foucault's (1977) book "Discipline and Punishment: The Birth of the Prison" describes the brutal mutilation of Robert-Franois Damiens in 1757, the last man to be publicly executed in France by drawing and quartering. From here on, Hudson (1996) explains we have seen the "transition from penalties inflicted on the body (execution, torture and mutilation, as well as less drastic physical punishments such as the stocks) to those directed at the mind and the character (labour and penance in the eighteenth - and nineteenth-century prisons, education and therapy in twentieth-century prisons), punishments which are designed to produce not the physically incapacitated citizen but the right thinking citizen" (Hudson, 1996, p. 7). This highlights the shift in the types of punishment that are deemed acceptable within society, and also how sentencing has moved away from the spectacle that it once was to become an almost hidden part of the penal system, taking place behind the closed doors of the courts (Ashworth, 2010b).

It is generally assumed that the right to punish is one aspect of the modern sovereign state (Ashworth, 2010b), and this is now extended to international bodies,

such as the International Court of Justice. Ashworth (2010b) notes, the justifications for assigning the central role to the state is generally derived from social contract theories. The crux of such theories involve a contractual agreement between the state and its citizens: citizens give up their right to use force against those who attack their interests, and give it over to the state. In return, the state will protect its citizens by maintaining law and order (ibid). The role of the state is to respond to wrongs in the form of crimes and offences.

Sentencing is also a form of public censure, whereby the sentences imposed should convey the relative censure for a particular offence(s) (Ashworth, 2010b).

2.2 Sentencing framework

There are a number of rationales for sentencing which can be adopted as an approach and generally these rationales will underpin the sentencing framework of a particular jurisdiction. I will briefly illustrate a general theoretical sentencing schema which demonstrates how these rationales can shape sentencing frameworks. I will then go on to discuss the current sentencing framework in England and Wales. The following section will then give a brief background of the sentencing legislation which led to the current framework being implemented.

2.2.1 Sentencing schema

The diagram in Figure 2.1 displays what I have called a 'sentencing schema' because it outlines the various aims and objectives of sentencing, which once adopted, provides a sentencing model or framework that can be utilised in sentencing.

At the centre of this schema are two distinct approaches to sentencing. The first approach goes by the umbrella term of utilitarianism. Utilitarianism encompasses a number of sentencing concepts and objectives, such as rehabilitative ideal - popular
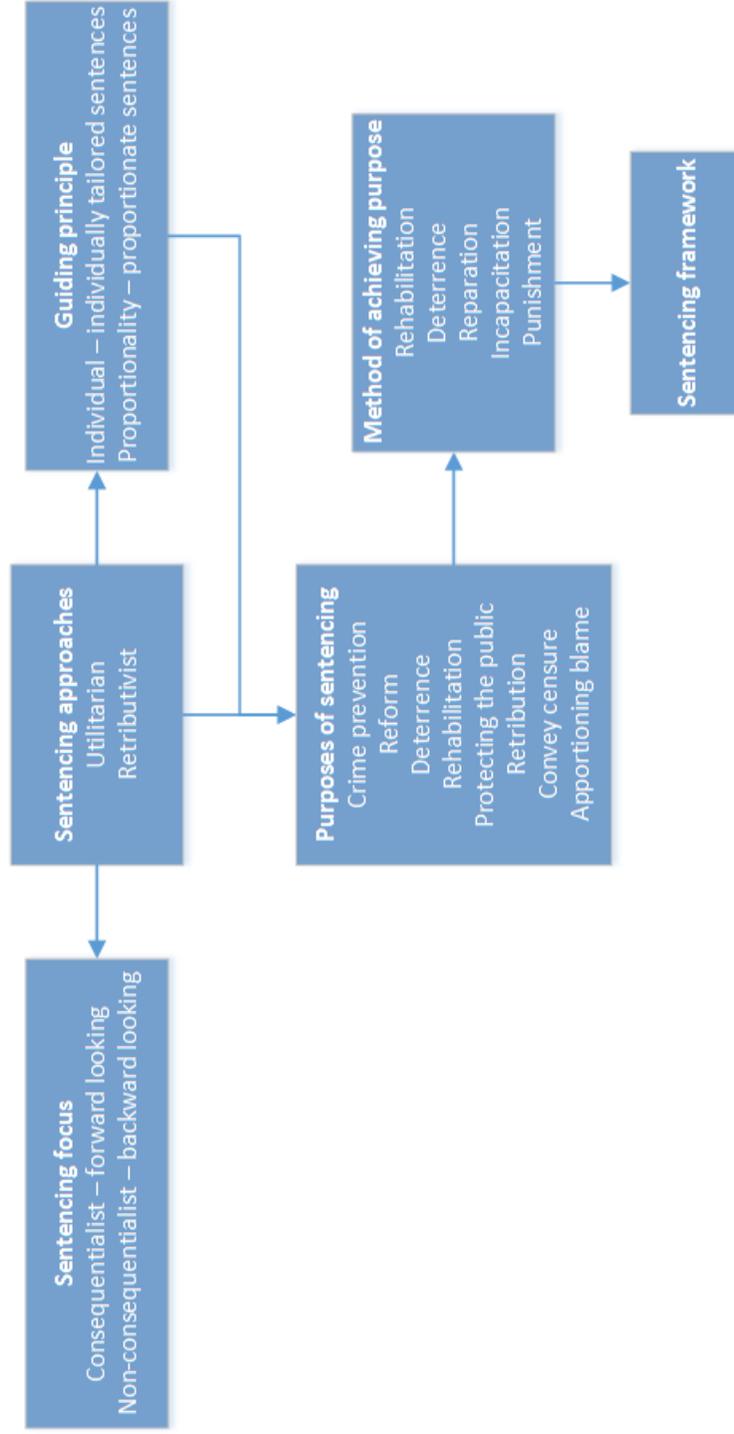
Figure 2.1: Sentencing schema diagram

in the 1960's - and restorative justice. The second is a retributivist approach, often referred to as the 'just-deserts' model or approach.

Each approach has a different sentencing focus and guiding principle. The utilitarian approach is concerned with preventing future crimes and so has a consequentialist, forward-looking sentencing focus (Francis et al., 2001). In this approach the guiding principle is the individual offender - or victim in the case of restorative justice - and therefore, these approaches argue that, sentences should be tailored towards the individual offender to prevent them from future offending (von Hirsch et al., 2009). Conversely, the retributivist approach is backward-looking - non-consequentialist (Francis et al., 2001). Proponents of these approaches believe that offenders should be punished proportionally for the offences they have already committed, and therefore, offenders should receive their 'just-deserts'. This approach argues that sentencing should be based on the seriousness of the offence, and not on the individual offender and what they may do in the future. Proportionality requires crimes to be ranked by their seriousness - determined by the harm caused by the offence and degree of culpability. I will say more about this later in the chapter.

These various approaches also have different views on the purposes of sentencing and the methods in which to achieve these purposes. As I have already explained, one of the fundamental principles of the utilitarian approach is crime prevention. Sentencing should therefore reform and rehabilitate the offender, and act as a deterrent from future offending. This approach argues that this can be achieved through, for example, rehabilitation and reparation. On the other hand, retributivists view sentencing as a form of retribution. Sentencing is also a way to convey censure for a particular offence, as well as a way of proportioning blame. To achieve justice, offenders must be punished proportionately for the offence they committed.

Most sentencing frameworks will be based on this sentencing schema, although generally, they will use a combination of these two approaches as I will explain for

England and Wales.

2.2.2   Sentencing in England and Wales

In 1974 a sub-committee of the Council of Europe, which had long been concerned with sentencing policy - namely inequalities and disparity in sentencing - produced an influential report on sentencing (Ashworth, 1994). Work on sentencing continued and this eventually led to the creation of a working group which comprised of a number of sentencing experts from around Europe. In 1989 the working group produced a report 'Consistency in Sentencing' and this became the basis of Recommendation No. R (92) 17, which was later adopted by the Council of Europe Committee of Ministers in 1992 (Council of Europe Committee of Ministers, 1993).

The report produced a number of recommendations (see Council of Europe Committee of Ministers (1993)). One of the recommendations stated that a primary sentencing rationale should be declared and that whatever rationale was adopted, it should avoid offenders being sentenced disproportionately. Thereby the seriousness of the offence should be the most important factor in determining the severity of a sentence (Council of Europe Committee of Ministers, 1993).

Furthermore it recommended that no discrimination should be made on the basis of the offender or victim's race, colour, gender, nationality, religion social status or political belief (Council of Europe Committee of Ministers, 1993). Additionally, factors such as unemployment, cultural or social conditions of the offender should not influence the sentence as to discriminate against the offender (ibid).

These recommendations have influenced sentencing in England and Wales, as well as the rest of Europe, and are somewhat responsible for the sentencing reforms which took place in a number of jurisdictions around the early 1990s. In England and Wales the recommendations have filtered into the sentencing framework by way of the sentencing statutes of 1991 and 2003 (von Hirsch et al., 2009).

The 1991 Criminal Justice Act laid down a sentencing framework which made proportionality the main principle of sentencing. From this point on, proportionality was deemed to be the main criterion of sentencing: putting emphasis on retribution as opposed to rehabilitation. However, Rex and von Hirsch (1997) highlighted that there was a lack of legislation or guidance on 'seriousness' - a key concept in the 1991 Act - and how this was to be implemented in practice.

In May 2000, the then Home Secretary - Jack Straw, announced a review of the sentencing framework which had been established in the Criminal Justice Act 1991. John Halliday led the review and identified some limitations of the 1991 framework (Halliday, 2001). The main problem identified was the erosion of the basic principle of proportionality: namely that the severity of a sentence should only reflect the seriousness of the offence. His review also noted the importance of improving public confidence in sentencing, which he argued could be achieved through a principled sentencing framework.

Following the Halliday review, an updated sentencing framework was introduced by the Criminal Justice Act 2003: part of the Act set out a new principled sentencing regime. The Criminal Justice Act 2003 reinforced the sentencing framework set out by the previous Act (Criminal Justice Act 1991), again placing emphasis on offence seriousness and proportionality, and for the first time, it laid down the principles of sentencing. However, heed was also given to the notion of risk, and in situations where an offender was believed to be dangerous it became mandatory to sentence out-with the realms of proportionate sentencing. I will come back to this point in a later section.

Sentencing principles

The Criminal Justice Act 2003 states sentencers must have regard for each of the five principles when deciding on the sentence to be passed down to the offender. The five

principles of sentencing according the Act are to:

- punish offenders

- reduce crime (reduction through deterrence)

- reform and rehabilitate the offender

- protect the public (by segregating offenders from the rest of the law abiding society)

- provide reparation to the persons affected by the offenders

The Act states all five principles are to be given equal weight: one principle should not be given precedence over another. This reinforces a more 'hybrid' sentencing approach, essentially covering all bases. However, this is met with some scepticism, for example Ashworth (2010b) views this as inviting inconsistency into the sentencing process and I will say more about this in the next chapter.

Offence seriousness

The Act also reinforced the principle of proportionality. In considering the appropriate sentence for the offender the sentencer must begin by considering the seriousness of the offence in respect to the culpability of the offender and the harm caused or risked being caused by the offence (Sentencing Guidelines Council, 2004).

Four levels of culpability are identified; intention, recklessness, knowledge and negligence (Sentencing Guidelines Council, 2004). The most serious form of culpability is where the offender had intention to cause harm, whereas the least serious form of culpability is harm that has been caused by negligence on the part of the offender. Along with the culpability, the actual harm caused by the offence must also be taken into consideration. This extends from individual victims to the community, and other types of harm which are difficult to categorise, such as cruelty to animals (ibid).

2.3   Sentencing guidelines

For most of the 20th century, the sentencing framework in England and Wales barely changed: other than setting maximum penalties and introducing and abolishing different forms of sentences, discretion over that time had been left largely to the judges and magistrates (Ashworth and Roberts, 2013). However, recognition of the importance of consistency and ensuring fairness in sentencing led to the implementation of sentencing guidelines in a bid to structure judicial discretion and provide greater accuracy in projecting the prison population (Ashworth and Roberts, 2013). Thus guidelines are viewed as one way of achieving greater consistency and predictability in sentencing (ibid).

Sentencing guidelines are not new to the criminal justice system in England and Wales: they have been used in practice for the last 25 years or so. Sentencing guidelines were initially used by the Court of Appeal in the form of judgements before they were incorporated into everyday sentencing. The Court of Appeal initially drew up some guidelines but were limited as to what they could achieve due to the lack of staff.

As a result, the Crime and Disorder Act 1998 created the Sentence Advisory Panel (SAP)[1] The Panel was established to draft and consult on proposals for guidelines and refer them back to the Court of Appeal. The Court of Appeal was not obliged to accept the Panel's recommendations but in most cases it did with some minor modifications.

An additional feature of the 2003 Act was the creation of the Sentencing Guidelines Council (SGC)[2]. The Sentencing Guidelines Council was responsible for setting up

---

[1]The Sentence Advisory Panel was constituted in July 1999. The panel consisted of sentencers, academics, people involved in the criminal justice system (prisons, prosecutions, police and probation) and lay people who had no connection to the criminal justice system (Ashworth, 2010b).

[2]The SAP still continued to draft and consult on guidelines as before, but the SGC, rather than Court of Appeal took responsibility for the creation and form of the guidelines. The SGC essentially filled the gap between the Court of Appeal and the SAP.

guidelines for the sentencing of a full range of offences, which Ashworth and Roberts (2013) acknowledge accelerated the development of definitive sentencing guidelines. Guidelines were produced for a number of offences, such as offences relating to the Sexual Offending Act 2003 and Robbery.

The Carter review in 2007 proposed that a permanent sentencing commission should be set up in England and Wales, and this was also the conclusion drawn by Hough and Jacobson (2008). It was felt that in a bid to address the "prison crisis" of a growing prison population, sentencing disparity and the increased politicisation of sentencing it was necessary to set up a commission to oversee sentencing in England and Wales (Hough and Jacobson, 2008). The overall objective of the commission was to achieve greater consistency and stability in sentencing practice thereby preventing any further upward drift in sentencing severity[3].

Following these recommendations the Sentencing Council (SC) was set up under the Coroners and Justice Act 2009. The SC replaced the SAP and SGC, and came into force on the 6th April 2010. The Sentencing Council[4] is responsible for issuing sentencing guidelines, as well as assessing the impact of the guidelines and promoting awareness amongst the public. The previous guidelines remain in force until they are replaced by the SC guidelines.

2.3.1 Sentencing guidelines in practice

For each type of offence the sentencing guidelines provide three categories which will reflect the varying degrees of offence seriousness together with the range of sentences for that offence. Consideration is given to the offender's culpability in committing the offence, the harm caused or intended to be caused by committing the offence, and other factors the Council considers to be relevant to the seriousness of the offence.

---

[3]Between 1993 and 2002, the prison population in England and Wales increased by around 36,000. This was due to increases in both the number of cases going before the courts and the proportion of those cases receiving a custodial sentence (Ministry of Justice, 2013b)

[4]The SC is an independent non-departmental body of the Ministry of Justice

The guidelines list a number of aggravating and mitigating factors that the court are required to take into account when considering the seriousness of the offence.

By virtue of the Coroners and Justice Act 2009, every court in England and Wales "must" follow the relevant sentencing guidelines unless the court is satisfied that it is not in the interests of justice to do so. This wording replaced the previous legislation which stated that the courts "should have regard" for the guidelines. Therefore the most recent provision now makes it obligatory to follow the guidelines in a bid to ensure consistency. However, judges are still allowed to depart from the guidelines where they feel they are justified to do so in which case they are required to state the reasons for sentencing out-with the guideline recommendations.

This wording has been criticised for being rather ambiguous. For example, Ashworth (2010a) has referred to the wording of the Act as being "pitifully loose". Not being explicit in these guidelines leaves room for not only differences in interpretation but also disparity. Although others have argued that it can also be seen as a way of maintaining some judicial discretion (Roberts, 2011b).

Aggravating and mitigating factors

The guidelines also produce a list of aggravating and mitigating factors that the court should take into consideration when considering the seriousness of a given offence. Aggravating factors are seen to increase the seriousness of the offence, whereas mitigating factors reduce the seriousness of the offence.

There are a number of aggravating factors and one mitigating factor that must be considered according to the Criminal Justice Act 2003 (Ashworth, 2010b). The aggravating factors are;

- previous convictions for relevant and recent convictions

- offence committed while on bail

- racial or religious aggravation

- aggravation related to disability or sexual orientation

- offences with a terrorist connection

The 2003 Act states that previous convictions must be treated as an aggravating factor in determining offence seriousness. Sentences are therefore cumulative in that as the numbers of previous convictions increase, so too does the seriousness of the current offence. This is partly due to the belief that the offender is not learning from their past experiences and therefore should be treated more severely in order to deter them from future reoffending. Judges should also consider the nature of the previous offence(s) in relation to the current offence to establish if there is a pattern of offending or whether this current offence is an escalation of prior offending. Also the time elapsed since the previous offence should be considered: a ten year old offence will be less relevant than an offence committed within the last six weeks.

However, this is met with some antagonism from some retributive theorists. Roberts (2009) explains that with regard to the use of prior convictions, retributive theorists form two distinct groups. The first group is the 'Exclusionary' group who believe previous convictions should not impact on sentence severity. The other group are those who adhere to the Progressive Loss of Mitigation (PLM) doctrine. They are of the view that first-time offenders should receive a discounted sentence. This mitigation should continue until they reach approximately five convictions. Roberts (2009) endorses a different model with regard to prior convictions; arguing for 'enhanced culpability', whereby previous convictions do indeed add to the culpability of an offender. However, the biggest increase is between a first-time offender and an individual with two previous convictions (Roberts, 2008). Therefore greater leniency should be shown to a first-time offender (ibid).

The second statutory requirement relates to being out on bail at the time of com-

mitting another offence. If the offender was out on bail at the time of committing the offence, it should be used as an aggravating factor therefore increasing the seriousness of the offence.

The 2003 Act explicitly makes reference to offences which target specific groups within society: this is viewed as an aggravating factor when assessing offence serious-ness. These include where the offence was racially or religiously aggravated or where the offence related to a disability[5] or the sexual orientation of the victim. Finally, the Counter Terrorism Act 2008 states that if an offence has a terrorist connection that that should be treated as an aggravating factor (Ashworth, 2010b).

There is only one statutory mitigating factor - a guilty plea. The Sentencing Guidelines Council (2007) state the reduction for a guilty plea is appropriate because it avoids the need for the case to go to trial, shortens the gaps between charge and sentence, reduces costs, and in the case of an early plea, saves victims and witnesses from having to appear in court to give evidence. The amount of reduction is gauged on a sliding scale ranging from a recommended one third where the plea was entered at the earliest possible opportunity (i.e. during police interview), reducing to one quarter (where trial date is set), and reducing to one tenth (for a guilty plea entered at the 'door of the court' or once trial has begun). The court should also take into consideration the circumstances in which the indication was given. These recommen-dations are fairly ambiguous and require the judge to use discretion in determining the amount of reduction.

Additional aggravating and mitigating factors

The Sentencing Guidelines Council (2004) produced a definitive guideline - 'Overar-ching Principles: Seriousness - which provides a non-exhaustive list of the aggravating factors which indicate higher culpability and harm which in turn, increase the serious

---

[5]'Disability' refers to any physical or mental impairment.

of the offence (see appendix A). This list applies to all cases sentenced on or after 4 August 2008. It also lists mitigating factors which indicate lower culpability, such as a greater deal of provocation than normally expected; mental illness or disability; youth or age, where it affects the responsibility of the offender; and the fact the offender played a minor role in the offence. Other offender mitigation is also mentioned including where the offender shows genuine remorse; made admissions to the police in interviews; and readily co-operated with the authorities (Sentencing Guidelines Council, 2008).

Youth or age (being under the age of 21) is viewed as a mitigating factor indicating lower culpability (Sentencing Guidelines Council, 2009). The guidelines state "a court sentencing a young offender must be aware of the obligations under a range of international conventions which emphasise the importance of avoiding 'criminalisation' of young people..." (Sentencing Guidelines Council, 2009, p. 3). Over the age of 21, age should not be considered as a factor relevant to sentencing.

The guidelines have also come under criticism by Roberts (2011a) for example, because they do not specify how much weight should be given to these aggravating and mitigating factors and also for not differentiating between the statutory requirements and the additional factors which are to be taken into consideration. I will say more about this is the following chapter.

Personal mitigation

The Criminal Justice Act 2003 makes provision for sentencers to also take into account any matter which they deem as relevant in the mitigation of a sentence, referred to as personal mitigation. The Sentencing Guidelines Council (2004) lists remorse and admissions to the police as personal mitigation, and any other factors that contribute to the smooth running of the criminal justice system (Ashworth, 2010b). Jacobson and Hough (2007) refer to personal mitigation as any factors which relate to the offender

as opposed to the offence, such as the offender's past or current circumstances, the response to the offence by way of showing remorse, and the offender's prospects. I will return to the discussion on the use of personal mitigation in the following chapter.

2.4   Risk and the New Penology

Around the same period that the new sentencing framework was being implemented in England and Wales, there was also a fundamental shift in crime control strategies particularly in England and Wales (Feeley and Simon, 1992). They refer to this shift as the 'New Penology' based on actuarialism and managing risk rather than trying to treat and reform offenders (the 'Old Penology'). The focus moved away from the individual offender, to priority being given to managing groups of offenders. This period marked the move away from clinical assessments based on professional judgements towards actuarial assessments based on calculated predictions of future risk. The actuarial approach aims to manage risk and therefore fits in with the future-orientated theories of sentencing. This in essence reaffirms the hybrid sentencing framework underpinning the current sentencing practice in England and Wales. Given that emphasis is given to both past and predicted/future behaviour.

Jacobson and Hough (2007) explain that Criminal Justice Act 2003 introduced new risk-based sentencing measures for those deemed as "dangerous" and these measures attempt to both manage risk and predict future criminality. Under the Criminal Justice Act 2003, if an offender was over the age of 21 and is convicted of a 'specified offence' (violent or sexual) after April 4th 2005, and the court considers the offender to be a risk to the public due to the likelihood of them committing a further specified offence, the court must impose a sentence of life imprisonment, imprisonment for public protection (IPP) or an extended sentence of imprisonment (Thomas, 2008). These mandatory sentences received mixed reactions from judges, but many felt that

their discretion had been curtailed (Jacobson and Hough, 2007). However, following the Criminal Justice and Immigration Act (2008) it became no longer mandatory to impose one of the sentences (Ashworth, 2010b). This gave some discretion back to the judiciary.

Another effect of the 'New Penology' was the introduction of the Offender Assessment System (OASys). In England and Wales, the Probation and Prison Services use the Offender Assessment System as an actuarial tool to assess the risk of recidivism and the potential harm the offender poses to others, as well as, themselves. The information recorded in OASys is provided to the court in the form of pre-sentence reports, which are discussed in section 2.5.1. The assessment records information on a number of static[6] (criminal history and demographics) and dynamic[7] (social and personal) risk factors to produce the OASys General reoffending Predictor (OGP) score of their potential risk of reoffending. Risk, as measured by OASys, can therefore be viewed as a product of the interaction between individuals and their environments (Grimshaw, 2002). A more detailed discussion of the OASys data is provided in section 5.2.1.

However, it would seem that the sentencing guiding principle of proportionality underpinning the current sentencing framework which is based on a retributivist or desert-orientated approach, and the actuarial approach used to predict future reoffending are somewhat conflicting (Jacobson and Hough, 2007). As previously discussed, the current sentencing framework places proportionality at the forefront of sentencing: offenders should be sentenced based on the seriousness of their offence and prior behaviour. Judges' can use their discretion using the aggravating and mitigating factors which will increase or decrease the seriousness of the offence. However, OASys then provides the court with information relating to the offender, and a recommended sen-

---

[6]The static risk factors are fixed (e.g. age at first offence and previous offending histories), and are said to assess potential long-term recidivism.

[7]The dynamic risk factors are included because research has found these factors to be closely associated with reoffending, whether directly or indirectly. These factors can change and therefore can have a positive or negative effect on the likelihood of reoffending.

tence that is based on predicted future behaviour. Ashworth (2010b) comments that this focus on risk and on 'non-legal' variables, such as upbringing, family size, income and housing, will lead to discrimination. This discussion is continued in Chapter 3.

2.5   The decision-making process

This section provides an overview of the process of sentencing decision-making in England and Wales.

2.5.1   Pre-sentence reports

Prior to sentencing, the court can request a pre-sentence report (PSR) to be provided by the Probation Service. This report can assist with the decision-making process in determining the most suitable sentence for an offender. PSR's are conducted by the Probation Service using the Offending Assessment System (OASys) (see Section 5.2.1) to record information about the offender and the nature and seriousness of the offence. Every report will also contain basic information about the offender and analysis of the offence, assessment of the offender and their risk of recidivism, assessment of the risk of harm they pose to themselves and the public, as well as a sentencing proposal, which will contain a recommendation of the type of sentence the Probation Service deems appropriate.

The court would generally request a report before deciding:

- if the custody thresholds[8] have been passed

- what is the shortest term of custody that is commensurate with the seriousness of the offence

---

[8]A court must not pass a custodial sentence unless the offence was serious enough to warrant such an offence (Sentencing Guidelines Council, 2004)

- what restrictions on liberty within a community order are commensurate with the seriousness of the offence

- whether requirements attached to a community order are suitable for the offender

Pre-sentence reports are a form of risk assessment of the offender and therefore relate back to the control strategy as discussed above in the New Penology.

2.5.2 Sentencing guidelines - user guide

The Sentencing Guidelines Council (2008) produced a user-guide for judges to follow, outlining the step-by-step instructions on the decision-making process. This user-guide has subsequently been updated by the Sentencing Council (2011), who have produced a nine-step guideline on sentencing assault cases and for all new guidelines produced. It is argued that compliance with the guidelines and therefore adopting consistency in the approach should reduce any unwarranted disparity, and any differences in sentences should be based on the legally-relevant factors of the case (Hutton, 2013). These steps will be briefly outlined.

Step one - Defining the offence category

The first step involves considering the offence seriousness and identifying an appropriate offence category. Judges are required to match the characteristics of the offence with those in the guidelines to identify the correct level of seriousness. Category 1 is the most serious and category 3 is the least serious. The court should determine the offenders culpability and harm using the list of factors provided in the relevant guideline.

Step two - Starting point and category range

Attached to each of the three offence seriousness categories is an appropriate sentence starting point as well as a sentencing range which is acceptable for that particular level of seriousness. Having determined the relevant offence seriousness category the court should use the corresponding starting point to decide on a sentence that falls within the given sentence category range. It then provides the additional list of aggravating and mitigating factors which the court should take into consideration, including previous convictions, whether the offence was committed whilst on bail and other non-statutory aggravating and mitigating factors. Hutton (2013) explains that consideration of factors at stage 2 may mean that it is appropriate for the court to move out of the category range having re-assessed the seriousness of the offence. The judge should now have an idea of the type of sentence they are going to give the offender before moving on to step 3.

However, as discussed in the previous section (Section 2.3.1), a lack of adequate guidance or weight to be applied to the different aggravating and mitigating factors (Roberts, 2011a) may result in judges reaching different decisions. This may then detract from the consistency in approach.

Step three - Consider any other factors which indicate a reduction, such as assistance to the prosecution

The court should consider if the offender should receive a discounted sentence if the offender provided any assistance while in custody.

Step four - Reduction for guilty pleas

As explained previously in Section 2.3.1, the court can reduce an offender's sentence if they plead guilty. The amount of the reduction in sentence will very depending on when the plea was entered. To receive the most discount, the offender would need to

have entered their guilty plea very early on in the proceeding, for example during their Police interview. This will mean the sentence can be less severe than is commensurate with the seriousness of the offence.

Step five - Dangerousness

Where the court deems the offender as 'dangerous' having committed a 'specified offence' as stipulated by the Criminal Justice Act 2003 (see Section 2.4), the court can award a more severe sentence, such as a life sentence or imprisonment for public protection. Where the offender meets the 'dangerous' criteria a minimum term sentence can also be set.

Step six - Totality principle

After the 11th June 2012, where an offender is being sentenced for more than one offence, the court must impose a sentence which reflects the overall behaviour of the offender (Sentencing Council, 2012a). All courts in England and Wales are required to apply the principle of totality (Sentencing Council, 2012b). The sentencing guidelines explain that when sentencing an offender for multiple offences, the court should pass a total sentence which reflects all the offending behaviour and is just and proportionate. This is to avoid some of the extreme sentences that we have seen given in the U.S., for example in the case of the United States of America versus Quartavious Davis. In this specific case, the defendant - Quartavious Davis, a 20 year-old man was sentenced to 1,941 months (162 years) in prison by a court in Florida for several counts of robbery, conspiracy, and possession of a firearm. The case is now under appeal (U.S. Appeal Court, 2014).

Step seven - Compensation and ancillary order

The court should decide whether the offender should give any compensation and/or receive any ancillary orders[9] .

Step eight - Reasons

The court has a duty to provide an explanation for any given sentence where it is outside the recommended range provided by the guidelines.

Step nine - Defining the offence category

Finally, consideration should be given for any time the offender has already spend on remand.

This last section should have provided the reader with a general understanding of the sentencing decision-making process and how it should be in principle according to the guidelines provided by the Sentencing Council.

2.6   Courts in England and Wales

In England and Wales there are two levels of court which deal with criminal cases, and there are substantial differences between them. The magistrates' court handles the least serious cases involving summary offences (e.g. most motor offences, minor criminal damage offences) and some triable-either-way offences which can be tried at the magistrates' or Crown Court depending on the seriousness of the offence (offences such as burglary or drugs offences). In 2009 the magistrates' court dealt with 93% of criminal cases (Ministry of Justice, 2010). The Crown Court then deals with the most serious cases involving indictable offences (e.g. rape and robbery offences) and

---

[9]Ancillary orders encompass a wide range of sentences, including; punitive orders, such as deprivation orders and disqualification from driving; reparation orders, such as restitution order; and preventative orders, such as anti-social behaviour orders and binding over orders (Ashworth, 2010b)

some triable-either-way offences. There is currently around 330 magistrates' courts and 77 Crown Courts throughout England and Wales.

All criminal cases will initially be heard in the magistrates' court and some will remain there for sentencing. There are generally four reasons why a case will go to Crown Court. Firstly, if an offender has committed an indictable offence the case will automatically be sentenced at the Crown Court - offenders will register their initial plea at the magistrates' court and then a date will be set for trial at the Crown Court. Secondly, the Powers of the Criminal Court (Sentencing) Act 2000 states that if the court is of the opinion that an offence(s) is so serious that greater punishment is necessary, that is out-with the punishment range of the magistrates' court, the court may commit the offender to be sentenced at the Crown Court. In the third instance for triable either-way offences, where the offender pleads not guilty and elects to be sentenced by trial by jury, the trial will be held at the Crown Court. The fourth reason a case will be held in the Crown Court is when an offender appeals against decisions made in the magistrates' court which will automatically be heard in the Crown Court.

2.6.1 Judges

The other distinction between the two courts is the type of judge which presides over them. Crown Courts have professional judges that sit with a jury, compared to magistrates' courts who have lay judges called magistrates' (usually three sit on the bench), or a District judge sitting alone (Ashworth and Redmayne, 2010).

Magistrates' have no formal legal training and will therefore be assisted in court by a legally qualified clerk. Upon becoming a magistrate, they will receive a compulsory programme of practical training preparing them for court (Courts and Tribunals Judicary, 2014). They will also be equipped with copies of sentencing guidelines which will aid their decision-making. Most magistrates' are unpaid members of the commu-

nity who have been appointed by the Lord Chancellor on the advice of local advisory committees. They work part-time, sitting in court a half day every fortnight but usually they will do more to meet the demands of cases going before the courts. There are approximately 21,500 magistrates' in England and Wales (Courts and Tribunals Judicary, 2014).

District Judges are paid members of the judiciary who usually sit alone on the bench. They deal with the longer and more complicated cases in the magistrates' court. District judges are also known as Justices of the Peace, formerly known as Stipendiary judges. They have had training in law and are usually practising barristers and solicitors.

Crown Court judges are professional judges who have not only trained in law but have also spent most of their working lives as practising barristers. The most serious offences are heard by a High Court Judges for 'Class 1' offences like murder; 'Class 2' offences including rape will be heard by a circuit judge, under the authority of the Presiding Judge; and 'Class 3' offences will be tried by a Circuit judge or Recorder (Courts and Tribunals Judicary, 2014).

A High Court judge will usually sit in London but can travel to other major court centres. They are trained lawyers who have held a 'right of audience' (the right to appear in court as an advocate) and have served as a recorder or district judge before they can be appointed (ibid).

Circuit judges sit in the Crown Courts. As with High Court judges, Circuit judges are trained lawyers who have held a 'right of audience' and have served as a recorder or district judge.

Recorders are fully qualified barristers or solicitors with at least ten years practice. The post is fully paid and is often the first step in the judicial path to the circuit bench. Recorders are similar to circuit judges but generally handle less serious cases.

2.7   Types of disposal

There are a number of disposals that the judge can choose to impose on an offender. As mentioned previously, the severity of the sentence should reflect the seriousness of the offence. These disposals will be summarised in turn along with how they were used in 2009 and the years running up to 2009. I focus specifically on 2009 because the data used in the subsequent research relates to offences between March 2008 to April 2010, and therefore 2009 is the full year.

2.7.1   Discharges

The Powers of Criminal Courts Sentencing Act 2000 states that in instances where an offender is convicted of an offence and there is no sentence fixed by the law or the court is of the opinion that the circumstances and nature of the offence and character of the offender does not warrant a punishment, the court can choose to either discharge the offender completely resulting in an absolute discharge, or the judge can place conditions on the discharge, in which case they will receive a conditional discharge (Richardson, 2011). Conditional discharges are subject to the offender not committing any further offences within the time period specified in the order, which cannot exceed three years from the date of the order (ibid). A conditional discharge thus carries the threat of future punishment (Ashworth, 2005). If an offender commits another offence within the time period specified, the offender can be re-tried for the original offence as well as the new offence (Richardson, 2011).

A sentence of discharge will cease to apply to the conviction (Richardson, 2011), therefore, a conviction resulting in a discharge does not count as such for most future purposes (Ashworth, 2005). Where a discharge has been given, it is not necessary to report the conviction when asked in the future (Richardson, 2011) but the same does not apply if they are asked if they have been found guilty in a court of law (ibid).

A discharge may not be combined with a fine for the same offence or a confiscation order but can be combined with a compensation order, a restitution order, an order to pay the costs of the prosecution, an order depriving the offender of their rights in property, or a recommendation for deportation (Richardson, 2011) . It can also be combined with a disqualification from driving order, a banning order or an anti-social behaviour order made under the Crime and Disorder Act 1998 (Thomas, 2008).

In 2009 1% of all sentences given in England and Wales were absolute discharges and a further 6% were conditional discharges (Ministry of Justice, 2010). The use of absolute discharges has remained stable over the ten year period between 1999 and 2009 (ibid). The same can be said for conditional discharges which also remained stable between 6-7% over the same period (ibid).

2.7.2 Fines

When an offender has been found guilty of an offence the judge can impose a financial circumstances order (Thomas, 2008). This order requires the offender to give the court information on their financial circumstances, which will then be taken into account when deciding an appropriate fine amount. The amount of the fine should reflect the seriousness of the offence and the circumstances of the case, which also includes the financial circumstances of the offender.

The Crown Court can impose any fine amount and they must fix a term of imprisonment to be served in default of the whole payment (Thomas, 2008). The magistrates' court can also impose a fine but there is a limit of £5000 for any one triable either-way offence unless special provisions are made for a larger fine. However there is no limit on the amount of the aggregate fines imposed for more than one offence (Thomas, 2008). Magistrates' courts can also impose fines for summary offences. Although there is again no aggregate limit for more numerous offences, there is a limit for any one summary offence which is £5000 (Thomas, 2008).

Fines can be combined with other sentences (not discharges) including custodial sentences. Where an offender has been on remand, this time should be taken into consideration when issuing them with a fine (Richardson, 2011).

Fines make up the majority of disposals issued by the courts in England and Wales. Between the years of 1999 and 2007, fines accounted for between 74-76% of all disposals given to adult offenders (Ministry of Justice, 2010). After 2007 the use of fines have fallen slightly to 71% in 2008 and 72% in 2009 (ibid).

2.7.3 Community orders

Community orders came into force on April 4th 2005 under the Criminal Justice Act 2003 but must not be passed where there is a sentence fixed by law, or a custodial sentence is required by the Firearms Act 1968, the Powers of Criminal Courts Sentence Act 2000, or the Criminal Justice Act 2003 (Thomas, 2008). This new order was viewed as an improvement on the previous community penalties, such as; curfew orders, community punishment orders, community rehabilitation orders, community punishment and rehabilitation orders, supervision orders, drug treatment and testing orders (DTTOs), as well as, order encompassing all or some these previous orders. Although the community order was introduced in 2005, the previous orders are still being used by judges although they are starting to be used less frequently now. In this research we only include community orders.

Where the offender is aged over 18 and is convicted of an offence the court may make them accountable to a community order imposing one or more requirements:

a. unpaid work requirement

b. activity requirement

c. programme requirement

d. prohibited activity requirement

e. curfew requirement

f. exclusion requirement

g. residence requirement

h. mental health treatment requirement

i. drug rehabilitation requirement

j. alcohol treatment requirement

k. supervision requirement

l. attendance centre requirement

A full explanation of the requirements is provided in Appendix B.

The court must make sure the restrictions on liberty are commensurate with the seriousness of the offences, the combinations of requirements are compatible and they are the most suitable for the offender. Community orders cannot last more than three years after the start of the order and time spent in remand should to be taken into consideration when passing the sentence (ibid). Additionally, community orders can only run concurrently so offenders who receive a community order for more than one offence will only actually serve one community order although the number of requirements will likely reflect the totality of offences. Consequently the severity of the sentence depends on the length of sentence as well as the onerousness of the requirements making up the order.

The use of community penalties on the whole has increased between 1999 and 2009. In 2009 14% of all disposals were community penalties, compared with 11% back in 1999 (Ministry of Justice, 2010). This increase in use of community penalties may be due to the introduction of the new community order in 2005. Over time judges may see the merits of the new orders and therefore start to use this disposal over another.

2.7.4 Suspended sentence orders

Suspended sentence orders were introduced in the 2003 Criminal Justice Act and were brought into force on the 4th April 2005 for offences committed on or after this date. The 2003 Act states that a court which imposes a custodial sentence for a term of at least 14 days but not more than 1 year (or 6 months in the case of the magistrates' court) can suspend the sentence for up to a period of two years[10]. The offence must be serious enough to warrant a custodial term and the length of the term should also be decided upon and only then should the decision be taken to suspend the sentence (Sentencing Guidelines Council, 2004). In doing so the offender will then serve their sentence in the community and will avoid prison providing they comply with the requirements within the 'supervision period' and do not commit a further offence within the 'operational period' as specified in the order (ibid).The supervision and operational periods must each be at least six months and the operational period is usually longer than the supervision period: it cannot be shorter. A suspended sentence order cannot be combined with a community penalty but the requirements that can be attached are the same as those in the community order.

A court cannot suspend a sentence without imposing any requirements but they can impose a suspended sentence order with modest requirements, such as a supervision requirement (Thomas, 2008). If the offender fails to comply with the requirements set out in the supervision period or the offender commits another offence within the operational period, the suspended sentence can be activated in full or in part or the terms of the supervision made more onerous. Although a suspended sentence order closely resembles a community order, offences which are likely to warrant a suspended sentence should generally be more serious than for those receiving a community order. The threat of future punishment (imprisonment) makes the suspended sentence order

---

[10]Prior to suspended sentence orders there were suspended sentences under the Powers of Criminal Courts Act 2000, s.118. The two differed in that the custodial sentence could be up to two years in length and could be suspended for between 1 and 2 years (Thomas, 2008). Only suspended sentence orders are included in this research.

a more severe punishment and with this in mind, the onerousness of the requirements should be less than those requirements attached to a community order (Sentencing Guidelines Council, 2004).

Prior to the introduction of the suspended sentence order in 2005, suspended sentences accounted for less than 1% of disposals given to adult offenders between 1999 and 2005 (Ministry of Justice, 2010) . From 2005 the number of offenders receiving this disposal increased, and by 2009 3% of disposals were suspended sentence orders (ibid). Suspended sentence orders are therefore growing in popularity with more judges now using this type of disposal.

2.7.5   Immediate custody

For a custodial sentence to be given the seriousness of the offence must cross the custody threshold. This threshold relates The seriousness of an offence must warrant a custodial sentence. However, if the offender fails to express his willingness to comply with requirements proposed by the court under a community order, and which the order requires an expression of willingness (e.g. an alcohol treatment requirement), or the offender fails to comply with pre-sentence drug testing then nothing in the Act prevents the court from passing a custodial sentence. However, as with all discretionary custodial sentences, Section 153 of the Criminal Justice Act 2003 states the length of the discretionary sentence should be the shortest possible commensurate with the seriousness of the offence (Richardson, 2011).

Custodial sentences make up a relatively small proportion of sentences: between 1999 and 2009 custodial sentence use ranged from 6% to 8% of disposals, most recently in 2009 this figure stood at 7% (Ministry of Justice, 2010).

Where an offender is sentenced for multiple offences that are serious enough to warrant a custodial sentence, the judge can order for the sentences to be run concurrently or consecutively. Sentences which are run consecutively must still be proportionate so

that the sentence is not overly excessive. This is generally referred to as the principle of totality (see Section 2.5.2).

Under the Magistrates' Court Act 1980, and the Powers of the Criminal Courts (Sentencing) Act 2000, magistrates' courts do not have the power to issue custodial sentences of over six months in respect to any one offence. If two or more terms are imposed in respect of a triable either-way offence, the maximum aggregate term is 12 months. Furthermore, these restrictions on aggregate terms do not apply to suspended sentences or restrictions to default terms fixed in respect of fines imposed on the same occasion as a custodial sentence is imposed (Thomas, 2008). For summary offences the maximum term of imprisonment is 6 months. Amendments were made in the Criminal Justice Act 2003 to extend the powers of the magistrates' court but at the time of this research, they are still to be brought into force (Richardson, 2011).

Early release provisions under the Criminal Justice Act 2003 (for offences committed after April 4th 2005) state, offenders sentenced to 12 months or more immediate custody will usually be released on licence after serving half of their sentence and will remain on licence until the end of the whole term of the sentence.

The Criminal Justice Act 2003 states the general restriction on imposing a discretionary custodial sentence. Discretionary custodial sentences cannot be given where there is a sentence already fixed by law, or where the offence(s) are associated with the use of firearms, is a third class A drug trafficking offence, a dangerous weapon was used, and where the offender is deemed dangerous and therefore a risk to the public (Richardson, 2011).

A court in dealing with an offender for an offence under the Firearms Act 1968 must impose at least the required minimum sentence unless there are exceptional circumstances (Thomas, 2008). The required minimum sentence is five years where the offender is aged 18 or over. Furthermore the court cannot give a discount for a guilty plea if the effect would take the term to below the required minimum (ibid).

Under the Powers of Criminal Court (Sentencing) Act 2000, a court which sentences an offender for a third Class A drug trafficking offence must pass a sentence of at least seven years. This is providing they are over 18 years of age at the time of committing the offence and the offence was committed on or after October 1st 1997. If the offender pleads guilty to the offence, the court can apply the discount for a guilty plea but the reduction cannot be less than 80% of the seven years (2045 days) (Thomas, 2008).

Dangerous offenders

The judge can impose life imprisonment, Imprisonment for Public Protection (IPP) or an extended sentence for any offender deemed to be "dangerous" having committed a specified violent or sexual attack.

Life imprisonment is imposed where the court considers the seriousness of the offence justifies the imposition of a sentence of imprisonment for life (Thomas, 2008). Usually a minimum term of custody will be set and following their release the offender will remain on licence for the rest of their life.

An IPP must be imposed if the offence is punishable by 10 or more years imprisonment, but not serious enough to warrant a life sentence. The offender will not be released until they have served the minimum term. IPPs are indeterminate sentences; therefore offenders can remain in prison for a period longer than the maximum term normally given for the offence if the Parole Board still views the offender as dangerous, and poses a threat to society. Upon release an offender will remain on licence for a minimum of 10 years. After this specified time, the Parole Board can end the licence but if they do not give the direction to do so the offender will remain on licence for the rest of their life. The only difference between the life imprisonment and IPPs is that the Parole Board can after 10 years choose to cease their licence after prison.

Extended sentences are usually given where a specified offence is not punishable

with 10 years' imprisonment. These sentences have two components: 'appropriate custodial term' and 'the extension period'. Both components combined cannot exceed the maximum sentence for the offence. Therefore the custodial sentence would be the same as it would be had the sentence not been extended. The extension period is the period the offender will be on licence. For a specified violent offence it cannot be more than five years and for a specified sexual offence it cannot be longer than eight years. Offenders serving an extended sentence can be released after serving one half of their sentence unless the Parole Board thinks they are still a threat to public protection and in that case will have to serve their entire sentence in custody.

2.7.6   Other disposals

This is a miscellaneous category of disposals, which includes compensation orders and confiscation orders. In official statistics, this category is recorded as 'otherwise dealt with' and accounts for approximately 2% of all disposals. Generally this type of disposal can be given for a primary sentence, but often this disposal is given as well as another disposal, such as a fine. However, as is explained in Chapter 4, official statistics only ever report the primary offence and with that, the most severe sentence where there are multiple offences or sentences.

2.8   Concluding remarks

This chapter provides a context to the subsequent research which assesses disparity in sentencing, and sets out how sentencing should be 'conducted' in theory, or as stated by the legislative provisions and guidelines in place in England and Wales.

This chapter also explains the theoretical underpinnings of sentencing, and the principles that judges should abide by when sentencing offenders. Proportionality is adopted as the primary rationale, and the severity of the sentences should reflect the

seriousness of the offence.

Additionally, where relevant and according to the guidelines the aggravating and mitigating circumstances of the offence should be taken into consideration. However, as I have highlighted, these guidelines have received some criticism due to ambiguity in meaning. Therefore, a consequence of this could be disparity if judges are interpreting guidelines differently or applying different weights to factors they feel are important which may differ amongst judges.

Also potentially confusing the philosophy underlying sentencing decisions is the development of the 'New Penology' and the concept of risk-management that this entails. The Probation Service use offender's socio-demographic characteristics to predict the probability of reoffending, which is then provided to the courts to assist with the decision-making process. The risk assessment is a fundamentally forward-looking endeavour which contradicts the principle of proportionality, which is essentially concerned with prior behaviour and the retributivist approach to sentencing.

Finally this chapter provides a brief explanation of the criminal court system in England and Wales, the judiciary, and the range of sentences that are available to the courts when sentencing offenders.

The decisions made in the modelling process in subsequent chapters are based on the information provided in this chapter. Therefore this chapter acts as a reference point throughout the remainder of this thesis.

The next chapter will provide more of a discussion on disparity in sentencing, drawing from the literature surrounding equality, consistency and fairness. It then goes on to discuss the existing literature on disparity and the research that has been conducted in this area.

Chapter 3

Review of the literature on sentencing disparity

3.1   Introduction

Up to now this thesis has concentrated on providing a background to sentencing in England and Wales, to provide a context to the discussion that follows in this chapter. In the previous chapter, I referred to the point that Ashworth (1987) makes that in order to assess disparity, it is necessary to stipulate how sentencing should be conducted in theory, or by way of the legislation that is in place for that jurisdiction. Having done so in Chapter 2, I will now provide the context in relation to sentencing disparity.

One of the basic precepts of justice is that "like cases should be treated alike and different cases differently" (Ashworth, 2010b, p. 255). This resonates with the principle of parity, suggesting that 'like-situated' offenders should be treated equally or equivalently. The general rule in legislative provision states that "when two or more offenders are convicted of the same offence, and their individual responsibility is the same, and there are no relevant differences in their personal circumstances,

they should receive the same sentence" (Thomas, 2008, p. 100). However, there is a growing body of evidence that suggests that some offenders are not being treated in this way and that actual sentencing behaviour can diverge quite dramatically from what is intended (Dhami, 2013).

This chapter will consider and review the sentencing literature in relation to disparity. Firstly by considering what is meant by disparity, then considering the legislation in relation to sentencing by way of sentencing guidelines and equality legislation. The remainder of this chapter will then review the literature and research conducted into sentencing disparity and highlight factors that are commonly considered to be associated with disparity.

3.2 Disparity

Disparity as a concept is widely discussed in the literature but formal definitions are rare. Blumstein (1983) defined disparity as a form of unequal treatment that is often of unexplained cause and is at least incongruous, unfair and disadvantaging in consequence. It is important to consider legal variables when considering disparity, and also take into account that sentences on different scales may have equivalent severity. This leads me to my definition of disparity. In this thesis the definition of disparity is as follows: disparity in sentencing occurs when offenders are sentenced not as a result of the offence they have committed but as a result of their extra-legal or personal characteristics. Disparity in sentencing occurs when two offenders, who have committed the same offence, and have the same legal characteristics (for example they have committed the same offence with the same level of seriousness, have entered the same plea, and have the same conviction histories) are sentenced and their sentences are not of the equivalent severity. Therefore disparity occurs when offenders are sentenced as a result of their extra-legal personal characteristics

and not the seriousness of the offence they have committed. Sentences do not need to be the same (for example, both offenders must receive a three month custodial sentence) but the sentences should be equivalent in terms of their severity.

Disparity in sentencing relates to instances where legally-irrelevant or extra-legal factors impact on the severity of sentences (Albonetti, 1997), although this is often referred to as 'unwarranted disparity' (Bushway and Piehl, 2001); (Maguire, 2010). This is the definition of disparity used in this research but this definition differs to that of (Spohn, 2009), who would argue that disparity is a difference in the treatment or outcome that does not necessarily result from intentional bias. Spohn (2009) argues that disparity occurs when two offenders are convicted of the same offence and who have identical offending histories, are sentenced differently, or when very different offenders receive the same sentence, or when the sentence depends on the judge and/or where the court is situated.

However, I am inclined to disagree with Spohn (2009), as in this thesis I argue (see Section 3.3.2 and Chapter 4) that sentences do not need to be the same for there to be parity, and likewise sentences do not have to be different for there to be sentencing disparity. Sentences require the same level of severity for there to be parity. I will argue that different types of sentence can actually have the same penal bite, and thus disparity occurs when like-situated offenders receive sentences which differ in terms of the severity rather than the actual sentence given, once the legal characteristics of the offence are taken into consideration.

I also argue that disparity in sentencing is contrary to the notion of justice, and introduces discrimination into the sentencing process. Disparity in sentencing can bring the criminal justice system into disrepute, as well as leading to unpredictable sentences (Tak, 1995), which relates back to the notion of fuelling the "prison crisis" (Hough and Jacobson, 2008) as referred to in Section 2.3. Furthermore, Cavadino and Dignan (2007) state that as disparities in sentencing violate the basic precepts

of justice it therefore adds to the penal systems crisis of legitimacy. The public also need to believe that justice has been served. Disparity in sentencing can lead the public to question the legitimacy of the criminal justice system. Sentencing outcomes that appear to be too lenient or severe can lead to a sense of injustice, which again undermines the legitimacy of the judiciary and the criminal justice system (Easton and Piper, 2008).

It is therefore imperative that research is conducted into disparity in sentencing to firstly assess whether it is occurring in England and Wales and secondly, where and how disparity is affecting sentencing decisions.

(In)consistency

Disparity and inconsistency are also often used interchangeably (Maguire, 2010) as both terms imply undesirable sentencing outcomes and are associated with injustice. However, there is a lack of a clear definition or explanation of what is meant by (in)consistency. In my view, the term 'consistency' is used when actually they mean parity. However, I will return to this discussion in Section 3.2.1.

Ashworth (1992) would suggest that it is in fact easier to agree that inconsistency in sentencing has been a problem in many jurisdictions than it is to agree on a definition. The Sentencing Council would also agree that there is no universally excepted definition, however the concept of consistency is clear: similar offenders who commit similar offences in similar circumstances would be expected to receive similar sentencing outcomes (Sentencing Council, 2010, p. 2). Therefore inconsistency implies offenders will receive dissimilar sentences.

Consistency in sentencing has been described as a"fluid and slippery concept" (Padfield, 2013, p. 32) because it is often unclear whether the focus should be on consistency in the approach or consistency in the outcome. Nonetheless, consistency is viewed as a way of achieving justice (Ashworth, 2010b) and has also been described

as a "fundamental constitutional principle of equality before the law" (see Section 3.3) (Samuels, 1987, p. 66). Samuels (1987) explains that this does not mean giving uniform sentences but being consistent in the approach. So although sentences may differ in terms of the types of sentences, sentences would be equivalent in terms of the severity of the sentence, and as a result, offenders are treated fairly. As I have said in the Chapter 2, one of the primary aims of the Sentencing Council is to promote consistency in the approach to sentencing, and one way of doing this is by producing the step-by-step guide to sentencing to ensure consistency in the approach.

Maguire (2010) insists that inconsistency occurs when there are differences in the legally relevant factors of the case which may result in offenders receiving different sentences. An example of this would be the result of applying the discount for a guilty plea. We would expect sentences to reflect this and differ accordingly between an offender who pleads guilty at the earliest possible opportunity, an offender who pleads guilty on the day of the trial, and an offender who pleads not guilty but is then found guilty, once we control for the offence type and the other legal factors of the case. This could be seen as one form of inconsistency, as the sentencing outcomes will differ to reflect the discount. Maguire (2010) suggests this difference is warranted, due to the fact that the legislative provision states that judges should apply a sentencing discount on this basis. We can then say that unwarranted disparity occurs when there are differences in sentence severity based on legally irrelevant factors of the case after the legal factors have been taken into consideration. These differences should unequivocally not differentiate the severity of the sentence.

In general there is a lack of a clear and agreed definition of consistency and inconsistency. Maguire (2010) is explicit in her definition of inconsistency but others do not explain what they mean when they use the term. As a result there is lot of ambiguity surrounding these terms which make them rather subjective. This has led me to believe that in fact when the term consistency is used with respect to sentencing

the term parity would be more accurate. Parity is less subjective because there is a clear definition of what these terms mean in relation to sentencing. It is also then possible to quantitatively measure parity, and with that disparity.

Discrimination

It can be said that too much discretion in sentencing can lead to disparity as it allows 'space' for discrimination (Easton and Piper, 2008). There is some confusion and uncertainty within the literature regarding disparity, inconsistency and discrimination: often these terms are used interchangeably. In Spohn's (2009) view, disparity and discrimination do not mean the same thing: discrimination in sentencing occurs when legally irrelevant factors, such as gender, race and class, etc, are taken into the sentencing decision after the legally relevant factors of the case are taken into consideration (Spohn, 2009).

For the purposes of this research, the focus is on disparity in sentencing and this can be said to occur when the legally-irrelevant factors are taken into consideration at the decision-making stage. I would argue that in most instances this disparity leads to discrimination, if for example an offender is sentenced less severely on the basis of them having a permanent job. As I will discuss in later sections, this disparity in sentencing then discriminates against those offenders that, for example, do not have a permanent job. Therefore, I would generally argue that disparity in sentencing essentially leads to discrimination, which resonates with Easton and Piper's (2008) point that was made at the start of this section, that disparity leaves room for discrimination. However, there are some that would disagree with this notion of disparity and discrimination which I will discuss in the following sections.

Additionally, (Spohn, 2009) goes on to say that there are two forms of discrimination; direct and indirect. Direct discrimination occurs when for example a judge sentences Black and Asian offenders more severely than White offenders even after

controlling for the legal characteristics of the case. Whereas indirect or subtle discrimination occurs when, for example, race has no direct effect on sentencing but those who spend time on remand are more likely to receive a custodial sentence, if Black offenders have a higher likelihood of being kept on remand than White offenders, then indirectly this will discriminate against Black offenders. Spohn (2009) also refers to an interaction effect, and this is where for example, women with children are treated more leniently than men with children, where both offenders are the primary carers of the child.

Discrimination occurs when the extra-legal factors are taken into consideration, which is the result of the disparate sentencing.

3.2.1 Model for achieving justice - where does disparity fit in?

The diagram in Figure 3.1 depicts a conceptual framework for achieving justice. As I have already explained, consistency is viewed as an objective of sentencing which if implemented should result in justice being achieved. This then follows the left-hand pathway of the diagram. However, as I mentioned in Section 3.2, I would argue that consistency is used more as a euphemism for parity. Parity requires offenders to be treated similarly (by not taking into consideration the legally irrelevant factors) - although I argue not necessarily the same as I explained earlier. Adopting the principle of proportionality should ensure that offenders are sentenced on the basis of the seriousness of their offence and the legally relevant factors of the case. It is generally believed that this will essentially lead to equality because offenders are not treated differently due to their extra-legal characteristics. Continuing to follow this pathway should then ensure that equality and justice are achieved.

It would then be generally agreed that if this pathway is not followed, i.e. the principle of parity is not exercised, this will produce disparity in sentencing. In the
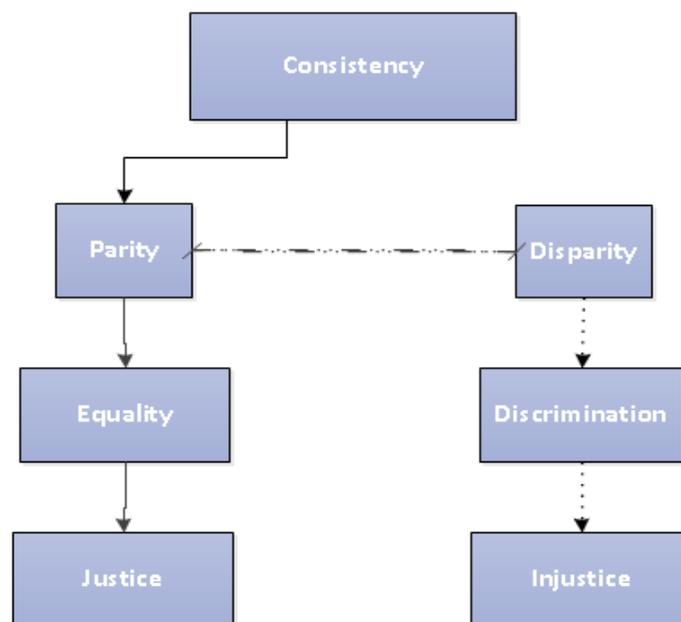
Figure 3.1: Diagram of the conceptual framework for achieving justice

diagram in Figure 3.1, this is illustrated by the broken line spurring from parity and resulting in disparity. As I have already explained, disparity occurs when legally irrelevant factors are taken into consideration during the sentencing decision-making process and offenders are therefore treated differently on the basis of these extra-legal factors, such as their gender or ethnicity. This disparity in sentencing leads to some form of discrimination, if for example, Black offenders are sentenced more severely than like-situated White offenders, controlling for the offence and other legal factors of the case. In turn, these offenders are being discriminated against which ultimately leads to injustice.

3.2.2 Alternative model for achieving justice - is disparity is necessary?

However, this conceptual model of achieving justice is generally contested by those who would disagree with the idea that consistency requires parity to facilitate equality and in turn achieve justice. For example, Hudson (1998) largely criticises this notion of consistency or rather parity, arguing instead that to accept this idea that equality and justice are achieved by treating offenders the same is to collude in the acceptance of the inequalities that co-exist with such treatment. Arguing from a feminist perspective, Hudson (1998) amongst others, such as Eaton (1986), Daly (1989), Chesney-Lind and Pasko (2012), argues that to assume justice means treating women like men, is to ignore the different existences of men and women. Figure 3.2 illustrates an alternative conceptual framework, which illustrates that in order for the idea of consistency in sentencing to lead to justice, men and women need to be treated differently (disparately), and only through this disparity can equality be achieved, which will lead to more 'just' sentencing.

I will say more about this model and the literature surrounding this framework in Section 3.5.1 which looks at disparity based on gender.
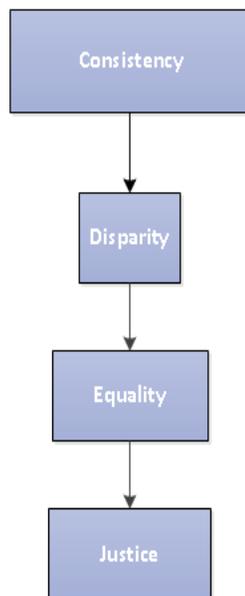
Figure 3.2: Diagram of the alternative conceptual framework for achieving justice

3.3 Equality before the law

Equality before the law states that offenders should not be discriminated against on the grounds of their gender, ethnicity, income or social class (Warner, 2012). These factors are deemed legally irrelevant and should not be taken into consideration at sentencing. Consequently, disparity in sentencing may then lead to discrimination, if for example, offenders are being treated differently due to their social characteristics or circumstances. Disparity in sentencing generally detracts from the principle of equality before the law.

There is no British Constitution as such which proclaims the principle of equality before the law (Ashworth, 2010b). However, the Human Rights Act 1998 brings into law most articles of the European Convention on Human Rights[11] (ECHR) (Ashworth, 2010b).

The Human Rights Act 1998 states that all public authorities must ensure that their actions are compatible with the Convention rights. Article 14 of the ECHR relates to freedom from discrimination and this seeks to protect individuals from different treatment - in exercising the other Convention rights such as the right to a fair trial (Article 6) and no punishment without law (Article 7) - on the grounds of sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status, as well as sexual orientation, whether born inside or outside of marriage, disability, marital status and age (Department for Constitutional Affairs, 2006). Therefore individuals should not be treated differently in similar situations without proper justifications. To do so is to discriminate against that person.

---

[11]The European Convention on Human Rights (ECHR) is an international agreement made by the member states of the Council of Europe and was ratified by the UK and entered into force in September 1953 (Department for Constitutional Affairs, 2006). However, for many years the convention was not part of the law in England and Wales until the Human Rights Act 1998 came into force in 2000. From here on the rights from the ECHR are enforceable in the courts of England and Wales (ibid).

The 2006 Equality Act established a general legal anti-discrimination framework monitored by the Equality and Human Rights Commission[12] (EHRC). The Act states that it is unlawful for any public authority when exercising a function (i.e. sentencing) to discriminate against a person and that the Commission must promote understanding of the importance of equality and diversity, amongst other objectives (Ashworth, 2010b).

The Equality Act 2010 then replaced the Equality Act 2006, and is an amalgamation of other previous discrimination legislation, including the Sex Discrimination Act 1975, Race Relations Act 1976 and the Disability Discrimination Act 1995. The Act came into force in October 2010 and set out to ban unfair treatment and provide equal opportunities within the wider society. This also includes sentencing which is considered an exercise of public function. The Act makes it illegal to discriminate against people on the basis of nine protected characteristics - age, disability, gender reassignment, marriage, and civil partnership, pregnancy and maternity, race, religion or belief, sex and sexual orientation (Equality and Human Rights Commission, 2011).

Yet in spite of this legislation there is evidence to suggest sentencing disparity and discrimination is prevalent in the England and Wales justice system.

3.3.1 Equal treatment or treatment that is equal?

It is argued that for justice to be achieved, treatment that is equal is required and not equal treatment per se (Ashworth, 2010b). This sentiment is shared amongst those who argue that in order to achieve fair and just sentencing, it must be recognised that the same sentence can have different effects on offenders (Morris and Tonry, 1990), which is a result of living in an unjust society (Aas, 2005). Many would refer to this problem as "just deserts in an unjust society", therefore punishment is unlikely to be 'just' when not everyone is equal from the outset (ibid). Tonry (1994) cites poverty

---

[12]The Equality and Human Rights Commission was set up under the Equality Act 2006, as a non-departmental public body, and came into force in October 2007.

and social disadvantage as just some of the ways in which society is not equal, and which will impact on those being punished.

Hudson suggests rejecting the idea of 'sameness' when considering equality in sentencing. This is not to reject equality, but what is needed is to ask for a "more complex formulation of equality" (Hudson, 1998, p. 246).

Morris and Tonry (1990) argue that parity can never be achieved due to the difficulty of comparing penalties among individuals.

> "In subjective terms...two years imprisonment in a single setting will have very different meanings to different offenders who have committed the same crime. Two years imprisonment in a maximum security prison may be rite of passage for a Los Angeles gang member. For an attractive, effeminate twenty-year old, it may mean the terror of repeated sexual victimisation. For a forty-year old head of household, it may mean the loss of a job and a home and a family. For an unhealthy seventy-five-year old, it may mean a death sentence" (Tonry, 1996, p. 19).

Individuals will therefore experience penalties differently. The concept of equality - relating to treatment that is equal as opposed to equal treatment - would suggest sentencing should be individualised and relative to each offender.

A way round this equality debate may be to adopt the principle of equal impact as opposed to equal treatment (Ashworth, 2010b). Equal impact requires sentences to be calibrated to create equal penal impact on the offenders subjected to them (Warner, 2012). Kolber (2009) would agree with this individualised approach to sentencing, as in his view, offenders should be sentenced depending on their baseline conditions, that is their conditions prior to being sentenced. This, in his opinion, is the only way to achieve fair and just sentencing. Therefore offenders should be sentenced subjectively, taking into consideration the true severity of incarceration which depends on the ways in which incarceration changes an offender's life, as illustrated in the above

quote by Tonry (1996). This alternative 'comparative' approach would examine the offender's life or their 'baseline' conditions prior to incarceration and sentence offenders based on these factors. This approach that Kolber (2009) suggests would take into consideration the fact that not all offenders are equal to begin with. Therefore the offenders referred to in the above quote would be given different sentences which take into consideration their different 'baseline conditions' (Kolber, 2009).

However, Nagel and Johnson (1994) would argue that if we were to allow for these "intersubjective comparisons" this would lead to a "free-for-all of sentence individualisation" and the end result of this would be disparity and discrimination (Nagel and Johnson, 1994, p. 205). In Nagel and Johnson's (1994) view, structured sentencing may help to control this, however the argument becomes even more difficult when female offenders are considered. I will return to this discussion in Section 3.5.1.

Von Hirsch (1992) also comments that the law only deals with standard cases, and therefore to take into account extraneous factors or extra-legal factors when sentencing then undermines its objectivity. Sentencing should only take into consideration the seriousness of the offence. However, deviations can be made for special situations and thus parity is more of an approximation that an exact measure (ibid).

In England and Wales, fines are the only penalty that are currently means tested which means the judge will take the financial circumstances of the offender into account when passing a financial penalty. No other penalties in England and Wales operate in this manner, although personal mitigation is taken into consideration prior to sentencing (see Section 3.4.2).

3.3.2   Equivalent sentence severity

However, there may be another way in which parity and equality can be achieved which in turn leads to just sentencing. That is to reconsider what is meant by these terms in sentencing - and with that disparity, if we consider the approach suggested

by Hudson (1998). Rather than taking equality or parity to mean 'the same sentence' (i.e. the same sentence or the same length of custodial sentence), we can take parity in sentencing to mean equivalent severity or leniency (Hudson, 1998). This relates to the earlier point made by Samuels (1987), where he suggests that consistency (or rather parity) in the approach to sentencing means that the types of sentences can vary but not the severity of the sentence, as they must be equivalent.

This resonates with von Hirsch (1992) on the grounds of ordinal proportionality, who argues that

> "Persons convicted of crimes of comparable seriousness should receive punishments of comparable severity...Persons convicted of crimes differing in gravity should suffer punishment correspondingly graded in their onerousness" (von Hirsch, 1992, p. 76)

If we then use the model for achieving justice as shown in Figure 3.1 and take parity to be considered in terms of sentence severity, then it is possible to ensure equality and achieve justice. In this respect, disparity occurs when sentences differ in their severity, as a result of extra-legal factors influencing sentencing decisions once the legal factors have been taken into consideration. I will say more about this in the following chapter where I review the literature on how to measure sentence severity.

3.4   Sentencing guidelines

A sentencing system that places emphasis on proportionality should avoid discrimination relating to the offender as the severity of that sentence should be primarily based on the seriousness of their offence (Ashworth, 2010b). However, in Section 2.2.2 I explained that Ashworth (2010b) is concerned that the hybrid sentencing framework adopted by the 2003 Criminal Justice Act is likely to undermine the principle of proportionality. Even though proportionality remains the primary sentencing rationale,

the fact that the 2003 Criminal Justice Act states judges must have regard to all five principles may detract from proportionality when judges attempt to pursue these ends. It may be very difficult to balance all five principles especially when a judge may be more inclined to pursue one particular end if, for example, the judge believes in a more utilitarian model of sentencing. I will return to this discussion on judicial preference later in the chapter.

However, sentencing guidelines are said to reduce judicial discretion and make the decision-making process more transparent and predictable (Aas, 2005). Aas (2005) explains that in the past judges were not subjected to any rules and this lack of rules could result in unfair and unequal treatment. Therefore guidelines are seen to minimise this.

Likewise, Hutton (2013) would argue that guidelines do not necessarily generate consistency in practice measured by outcomes but they should enable an account of consistency to be produced. The step-by-step sentencing approach discussed in Section 2.5.2 provides a more consistent approach to sentencing, which should then reduce unwarranted disparity. Roberts (2011b) refers to the step-by-step sentencing guidance as a sentencing "algorithm". By following this algorithm consistency in sentencing should be possible.

However, the usefulness of sentencing guidelines may be undermined by the lack of guidance provided for the use of aggravating and mitigating factors in the sentencing decision-making process.

3.4.1 Aggravating and mitigating factors

In Section 2.3 I briefly mentioned Roberts' reservations regarding aggravating and mitigating factors; namely the lack of adequate guidance provided on them (Roberts, 2011a). This lack of guidance or structure can be viewed as a means that allows the judiciary to maintain some discretion with only minimal direction from the legislature

(Roberts, 2008). However, this lack of guidance may also cause the judiciary to use the aggravating and mitigating factors inconsistently, ultimately undermining the concepts of equality and fairness which the sentencing guidelines set out to achieve (ibid). A consequence of this may then lead to sentencing disparity.

Roberts (2011a) also draws attention to the list of aggravating and mitigating factors: the definitive offence seriousness guidelines list 31 aggravating factors and only 4 mitigating factors, which instantly creates an imbalance towards increasing the seriousness of the offence. As well as this, within the list of statutory aggravating and mitigating factors, there are 4 statutory aggravating factors and 1 statutory mitigating factor. Firstly Roberts (2011a) points out that the guidelines do not differentiate between those which are statutory as stated in the Criminal Justice Act 2003, and those that are not: secondly, the Criminal Justice Act 2003 does not provide any reference to the weight that each of these factors should carry in determining an appropriate sentence, including no distinction between the statutory and the non-statutory factors, where in essence the statutory factors should be given more weight than the non-statutory factors (Roberts, 2008). Judges are left to make these decisions without any indication as to how much these factors should influence offence seriousness and consequentially, sentence severity. Dhami (2013) argues that too much discretion appears to lead to unwarranted disparity and this is because judges are given no guidance on how to use these aggravating and mitigating factors.

However, there is research to suggest this might not actually be the case. Pina-Sánchez and Linacre (2013) conducted research which assessed the use of sentencing guidelines for assault cases and how they were applied by judges in the Crown Courts in England and Wales, using data from the Crown Court Survey. Their results suggest that guideline factors are being taken into account in a way that is intended by the sentencing guidelines, and this is consistent across courts. Their results revealed that pleading guilty at the first opportunity, showing remorse, being a carer of a

dependent person, or a public worker, reduced the length of the sentence, where as having previous convictions, perpetrating on a vulnerable victim, sustained offending over time, and being under the influence of drugs increased the length of the sentence (Pina-Sánchez and Linacre, 2013). This work ties in with the definition of consistency/inconsistency offered by Maguire (2010), which relates to the legal factors of the case.

On the basis of this research, it may seem that sentencing is carried out as it should be - informed by the guidelines produced by the Sentencing Council, and that judges are applying the aggravating and mitigating factors consistently. However, these results do not consider the possibility that extra-legal factors may also be taken into consideration. Johnson (2005) would argue that sentencing does not occur in a social vacuum and therefore by not taking the extra-legal factors of the case into consideration we will not get an accurate representation of sentencing.

3.4.2 Personal mitigation

Roberts (2011a) is not alone in his opinion that not enough guidance is given in relation to aggravating and mitigating factors. Jacobson and Hough (2007) would argue there is also not enough guidance on personal mitigation. In Section 2.3.1 I explained that the Criminal Justice Act 2003 is very ambiguous in relation to personal mitigation, which can be "any matter" that the court feels is "relevant in mitigation of sentence". Research conducted by Jacobson and Hough (2007) found that many different personal mitigation factors were taken into consideration by judges, such as an offender's employment status, which was frequently expressed as the deciding factor for cases on the borderline of the custody threshold.

Additionally, Roberts (2008) draws attention to the fact that the guidelines do not specify factors that should not be taken into consideration, unlike most other sentencing guideline systems. For example, the Minnesota Sentencing Guidelines Commis-

sion (2013) states that; race, sex, employment factors - such as being employed at the time of sentencing, and social factors - such as living arrangements and marital status should not be factors used to depart from the sentencing guidelines. Roberts (2008) refers to the offender's employment status at the time of sentencing, which is often cited as a mitigating factor where the offence is unrelated to the offender's employment[13]. Additionally, research by Jacobson and Hough (2007) found that offender's employment status was frequently expressed as a mitigating factor when cases were on the borderline of getting a custodial sentence. However, Roberts (2008) points out that mitigating a sentence on the basis of employment disadvantages offenders that are not employed at the time of sentencing or those that have poor employment histories.

Although it may be acceptable to some that a good employment record is a powerful factor in the mitigation of a case, this source of mitigation is not available to those who do not have a good work record or are unemployed (Roberts, 2008). Using employment status as a mitigating factor then discriminates against those who are already less fortunate and for whatever reason have experienced difficulties in obtaining or maintaining paid employment. Ashworth (2010b) would then argue that the only way to pursue equality in sentencing is to offer no concession at all on the basis of employment just like the state of Minnesota.

Furthermore Ashworth (2010b) suggests it is perhaps wrong to think of aggravating and mitigating factors as opposites, in fact we should think of some factors as being neutral. This may seems plausible when we consider the reduction for a guilty plea but this does not seem fair when considering the employment status of the offender. To mitigate on the grounds of being employed could be seen to discriminate against those who are unemployed (ibid). This discussion will be continued in Chapter 10.

---

[13]If the crime arises out of the offenders' employment, it would be regarded as an abuse of trust in which case it would be considered an aggravating factor.

3.5 The evidence of sentencing disparity

The following section summarises the literature on disparity.

3.5.1 Gender

In 2009, 314,627 women were sentenced in England and Wales - an increase of 9.1% (difference of 26,289 additional female offenders) from the previous year (Ministry of Justice, 2010). However, this figure is still much smaller in comparison with the male figures which stood at 1,046,281 in 2009 - a 0.1% increase from the previous year. Table 3.1 displays the proportions of persons (aged 21 and over) sentenced at all courts for indictable offences (these are more serious offences) in 2009 by gender and disposal. The statistics show that women appear to be sentenced more leniently than male offenders: women offenders are more likely to receive a fine (81%) compared to male offenders (69%), whereas a higher proportion of male offenders are given community orders (11%), suspended sentences (4%) and immediate custody (9%) compared to female offenders (7%, 2% and 3% respectively) (Ministry of Justice, 2010). Although it is worth noting that these figures only report the primary offences and do not control for offence type: they therefore cannot be taken at face value (Ashworth, 2010b).

	Male	Female
Absolute discharge	0.5%	0.5%
Conditional discharge	5.4%	5.8%
Fine	68.8%	81%
Community order	10.9%	6.9%
Suspended sentence order	3.9%	2.2%
Immediate custody	8.6%	2.5%
Otherwise dealt with	1.9%	1.1%

Source Ministry of Justice (2010) supplementary tables (Table 1d)

Table 3.1: Proportions of persons (aged 21 and over) sentenced for indictable offences in 2009 by gender and disposal

Ashworth (2010b) explains that different types of offences are usually committed by men and women and therefore these figures may not represent the true nature of sentencing. More women are sentenced for theft and related offences (a common offence being shoplifting) and less likely to be convicted for burglary and drug offences which carry heavier penalties (Flood-Page and Mackie, 1998); (Ashworth, 2010b). Furthermore, women tend to play more minor roles in offences and are therefore sentenced accordingly (Hedderman and Gelsthorpe, 1997).

The result of this is that a higher proportion of women are sentenced at the magistrates' court rather than the Crown Court, due to their involvement in less serious offences (Ashworth, 2010b). As explained previously, the Powers of Criminal Courts (sentencing) Act 2000 limits the sentencing powers of the magistrates' court and this will likely have an impact on the sentences meted out to women offenders.

Research has shown that between 1992 - 2002, some 37% of women coming before the courts had no previous convictions (Prison Reform Trust, 2000) which could explain the differences in the sentencing statistics. However, Dowds and Hedderman (1997) found that for shoplifting offences, regardless of whether the offender was a time offender (1% females and 7% males) or a repeat offender (5% females and 15% males), women offenders were less likely to receive a custodial sentence than male offenders.

Cavadino and Dignan (2007) acknowledge that by and large commentators agree females commit fewer offences than males and these offences are generally less serious than those committed by their male counterparts. Research conducted by Spohn (2009) found these claims were unsupported. Spohn's (2009) research found statistically significant gender differences: judges were 2.5 times more likely to sentence male offenders to prison than women controlling for the legal factors of the case. Albonetti (1997) also tested this by looking specifically at drug offences and controlling for the legal factors of the case. Albonetti (1997) found that not only are women 6% less likely

to get a custodial sentence compared to men, but when they do receive a custodial sentence, the length of the sentence is substantially shorter. Stacey and Spohn (2006) also found male offenders were sentenced significantly longer than female offenders. Using data from three district courts in the US where the offender was sentenced for a drug offence between 1998 and 2000, Stacey and Spohn found the average length of custodial sentence for female offenders was just under 60 months, compared to that of male offenders, which stood at just over 97 months.

One argument for the disparate sentencing of men and women is that the cultural and biological differences are relevant and legitimate considerations for sentencing. The Corston Report[14] called for a "distinct, radically different, visibly-led, strategic, proportionate, holistic, woman-centred, integrated approach" (Corston, 2007, p. 79). Baroness Corston and the Commission on Women Offenders in Scotland report (Scottish Government, 2012) both highlighted that much of women's offending is initiated and propelled by a number of underlying issues, such as; drug and alcohol addiction and mental health problems. Her Majesty's Inspectorate of Prisons reported that in Scotland's only all-female prison[15] 80% of women prisoners had mental health problems (Her Majesty's Inspectorate of Prisons, 2007). In another study, 60% of women prisoners said they had been under the influence of drugs at the time they committed their offence (Scottish Government, 2012).

It was agreed by the Probation Chiefs Association[16] , Prison Governors' Association and Michael Spurr of the National Offender Management Service (NOMS) that the Corston Report had brought to their attention the additional needs of women offenders and the evidence that "women require a different and distinct approach" (House of Commons Justice Committee, 2013, p. 9).

---

[14]In 2007, Baroness Corston conducted a review of women in the Criminal Justice System resulting in The Corston Report.

[15]Up until 2011 Cornton Vale was Scotland's only female prison but since then wings for female prisoners have been established at HMP Edinburgh and HMP Greenock.

[16]The Probation Chiefs Association (PCA) is the independent professional organisation that represents the professional voice of senior leaders in Probation Trusts in England and Wales

A source of great controversy is whether female offenders should be treated the same as male offenders (Spohn, 2009). As already discussed, one argument dominates: fairness and justice demand equal treatment not 'special treatment' (Nagel and Johnson, 1994). This is illustrated by the conceptual framework for achieving justice in Figure 3.1. However, Baroness Corston maintains "equality does not mean treating everyone the same" (Corston, 2007, p. 3) and therefore the alternative conceptual framework shown in Figure 3.2 should be adopted. Corston (2007) argues that the fundamental difference between male and female offenders indicates that a different and distinct approach is needed to compensate for the different lives and nature of men and women, which must mean that they cannot be treated the same.

It has been suggested that gender-based leniency is apparent, as female offenders tend to benefit from what they refer to as "a benign form of reverse discrimination" (Nagel and Johnson, 1994, p. 182). This is often referred to in the literature as the sentencing chivalry.

In a similar vein, Daly (1989) and Carlen (1988) conducted research looking at judge's motivations in sentencing of women. Daly (1989) found that judges were motivated by a desire to protect families (although not necessarily women). In the Scottish study by Carlen (1988) judges openly admitted their dislike of sentencing women. Carlen (1988) found judges tended to sentence women offenders on the basis of whether or not offenders are mothers. Thus if their child was already in care they would be more inclined to give a custodial sentence. These findings are supported by surveys of women in prison that highlight the disproportionate numbers of women with unconventional family backgrounds (Cavadino and Dignan, 2007).

It may also be possible that women are sentenced less harshly because of "gendered presuppositions of crime and justice" (Daly, 1994, p. 197). These more lenient sentences are a result of judges believing women are less blameworthy, less dangerous, more likely to be deterred from future offending and that the social cost of punishing

female offenders has increased (Steffensmeier et al., 1993). Johnson (2005) supports this argument as his results are consistent with this theoretical perspective, suggesting courtroom actors may succumb to stereotypical patterned responses which tie offender characteristics to courtroom assessments of dangerousness and culpability. Additionally research carried out by Hedderman and Gelsthorpe (1997) found that judges often view women offenders as 'troubled' as opposed to 'troublesome' because women tend to be first-time offenders, face less serious charges and behave more respectfully in court (Hedderman and Gelsthorpe, 1997, p. 26). Although when shoplifting was conducted for profit rather than necessity, they were seen to be more like their male counterparts (ibid).

Males and females differ in many aspects but special treatment of female offenders may carry risk (Nagel and Johnson, 1994). The main risk is that those who do not live up to the typical female stereotype are treated more harshly. Thus they are deemed as doubly deviant by going against social norms as well as the law (Nagel and Johnson, 1994; Cavadino and Dignan, 2007). Farrington and Morris (1983) found evidence of this: divorced and separated women, as well as those seen as 'deviant'(e.g. unemployed unmarried mothers), received more severe sentences than married 'normal' women. The traditional family unit is seen as the 'norm' and therefore having dependent children is often viewed as a mitigating factor.

However, the Fawcett Society (2009) draws attention to the fact that too many women are being imprisoned on short sentences for non-violent crime. They argue that remand is over-used for female defendants and there are too many foreign national women in prison (ibid). The Fawcett Society attributes the rise in the female prison population to the significant rise in the severity of sentences given out by the courts. They are not alone in this belief: the increase in the severity of sentencing in England and Wales has been the main cause of the rocketing prison population as a whole according to Halliday (2001) and Carter (2003). Sentencing decisions are seen as a

potent source of perceived injustices fuelling the crisis of legitimacy (Cavadino and Dignan, 2007).

In the US the rise in the number of women being incarcerated has been blamed on the development of drug policy. It is said the 'war on drugs' policy resulted in the increase in penalties for possession and sale of drugs and with this, the "'war of drugs' has translated into war on women" (Chesney-Lind, 1995, p. 111), with women being treated like their male counterparts.

The literature regarding female sentencing is somewhat mixed: on one side of the argument there are those that maintain the conceptual model for achieving justice where offenders need to be sentenced with parity and no 'special treatment' (Nagel and Johnson, 1994) and on the other hand there are the supporters of the alternative model, where women offenders are treated disparately to deal with their additional needs, as raised by Corston (2007). There is some evidence to suggest that men and women are sentenced differently, and this may be a result of some of the factors just highlighted. Further research is therefore necessary to identify if there is sentencing disparity based on the severity of the sentence and not just the type of sentence.

3.5.2 Race and ethnicity

The principle of equality before the law states that no person should be sentenced more severely on account of their race or ethnicity. However, there is evidence to suggest this may not be the case. The statistics show that Black and Minority Ethnic (BME) group offenders are over-represented at virtually all stages of the criminal justice system relative to their representation in the general population (Thomas, 2010). If we take the prison population in England and Wales as an example, the Ministry of Justice (2009) reported that in 2008, 27% of the prison population identified themselves as being from BME groups, that is just over a quarter of the male prison population and 29% of the female prison population. Furthermore, statistics released by the Ministry

of Justice (2011) show that in 2009, Black offenders received the longest average custodial sentence length at 20.8 months, compared to Asian and White offenders at 19.9 and 14.8 months respectively.

Black offenders are also more likely to be sentenced at the Crown Court (Hood, 1992; Thomas, 2010). Data obtained from the Crown Court (CREST data) revealed that BME members are two and a half times more likely to be sentenced at the Crown Court relative to their representation in the population - 21.7% of those sentenced at the Crown Court compared to their representation in the population (8.7%) (Thomas, 2010). Black offenders are four times more likely to be sentenced in the Crown Court - 10.7% of those sentenced at the Crown Court compared to their representation in the population (2.8%) (ibid). The fact that more BME offenders are sentenced at the Crown Court could be viewed as one example of indirect discrimination (Spohn, 2009).

Additionally, it is evident from Table 3.2 there are some differences in the proportions of offenders receiving each disposal by their ethnic group. Although we cannot take these figures at face value [like the previous figures for gender], there are some substantial differences in those receiving the lesser and more severe disposals. For example in 2009, a higher proportion of White offenders received condition discharges and community penalties compared to the Black and Asian offenders, whereas a much higher proportion of Black and Asian offenders were sentenced to immediate custody than Whites.

Sentencing is just one stage of the criminal justice system: prior to this stage those appearing in court are already from a selected group constituted as a result of reporting, investigating and filtering in the pre-trial stages (Ashworth, 2010b). For example, Ashworth (2010b) highlights the use of stop and searches carried out by the police. A consistent finding is that people from ethnic minorities - particularly Black people - are stopped and searched by the police more often than the White population

	White	Black	Asian
Absolute discharge	0.5%	0.4%	0.3%
Conditional discharge	12.7%	8.7%	7.8%
Fine	16.3%	18.2%	18.2%
Community order	34.1%	30.1%	27.7%
Suspended sentence order	9.6%	8.3%	10.7%
Immediate custody	23.4%	29.1%	32%
Otherwise dealt with	3.3%	5.2%	3.3%

Source Ministry of Justice (2010) Table 1(vi)

Table 3.2: Proportions of persons (aged 21 and over) sentenced for indictable offences in 2009 by ethnicity and disposal

(Bowling and Philips, 2003). This targeting of BME people in this manner is likely to play a part in the over-representation of this group within the criminal justice system (Phillips and Bowling, 2002).

It was found BME members are also more likely to be refused bail and therefore are more likely to be remanded in custody prior to appearing in court for sentencing (Phillips and Brown, 1998). In Phillips and Brown's (1998) study, they found 26% of White offenders were refused bail compared with 34% of Black offenders and 35% of Asian offenders. Evidence suggests that bail decisions are affected by criteria such as homelessness (Walker, 1989), unemployment and socio-economic factors that disproportionately affect Black offenders (Phillips and Brown, 1998). Studies have shown that Black offenders are more likely to be remanded in custody than their White counterparts and as a result of this, they are then more likely to be given a custodial sentence if found guilty (Hood, 1992). This can be viewed as one form of indirect discrimination against Black offenders (Hudson, 1993; Spohn, 2009).

At the sentencing stage BME offenders are more likely to plead not guilty (43%) than White offenders (35%), resulting in more of these cases going to trial (Thomas, 2010). BME offenders are three and a half times more likely to face a jury verdict (29% of all trials by jury are for BME offenders), and Black offenders are five times more likely to face trial by jury (14% of all trials by a jury). There is a general

belief that juries are fairer that judges (Ashworth and Redmayne, 2010) (see Section 3.5.4). Analysis by Thomas (2010) would suggest this may be the case as the overall conviction rate for BME offenders is 65% compared to 63% for White offenders.

The most prominent UK study into ethnic disparity in sentencing was carried out by Hood (1992). This study collected data from five Crown Court centres in the West Midlands of England in 1989, consisting of 2884 males (half of the sample were White and other half non-white) and 443 females. This work highlighted severe disparity and discrimination relating to offenders' ethnicity. Namely, variations in custody showed a 5% greater probability of Black offenders being sent to prison, which varied by court, and variations in the use of alternatives to custody showed that Black offenders are more likely to receive more severe disposals, and for those given custody there were also differences relating to sentence length, with Black offenders receiving longer sentences especially those over three years.

This situation appears similar if not worse in the U.S, where studies have shown severe racial disparity in sentencing, particularly for the Black and Hispanic offenders (Johnson, 2006; Engen and Gainey, 2000; Bushway and Piehl, 2001). For example, Albonetti (1997) found there to be substantial differences in the mean length of imprisonment for drug offences varying by Black, White and Hispanic offenders. Black offenders received the longest sentences, averaging 101.97 months, compared to White and Hispanic offenders, which averaged at 72.45 and 79.58 months respectively: suggesting that sentencing is influenced by an offender's ethnicity.

These results appear consistent throughout the literature from the US, highlighting a major problem with ethnic discrimination in the criminal justice systems in the US, as well as England and Wales.

3.5.3 Additional characteristics of the offender

There is a lack of recent empirical research which investigates the effects of additional characteristics of the offender - other than gender and ethnicity - on sentencing practice. Studies which have been conducted in this area, tend to be outdated and limited in the methodologies employed to conduct the research. I will return to the issue of methodologies in the following chapter but for now I will briefly summarise some of the existing research that has been conducted into disparity.

Additional offender characteristics may be considered as part of the offender's personal mitigation. Research conducted by Jacobson and Hough (2007) held focus groups with Crown Court judges and found that often personal mitigation was the deciding factor for borderline cases. This research also found evidence that being employed is treated as a mitigating factor, with judges reluctant to jeopardise employment prospects. However, as I have previously pointed out, this potentially discriminates against offenders that are not employed prior to sentencing.

Other studies that have looked at the effects of employment status on sentencing generally date back to the 1980's and 1990's. Nonetheless, Flood-Page and Mackie (1998), Moxon (1988) and Crow and Simon (1989) all found that employment status was strongly associated with the use of fines. Flood-Page and Mackie's (1998) results showed that for first-time offenders, 10% of employed offenders received a fine compared with just over 2% of unemployed offenders and among those with previous convictions, 8% of employed offenders were fined but none of the unemployed offenders were fined. However, they do acknowledge that this does take into account the offenders means to pay. These studies also found evidence to suggest that as a result of judges' reluctance to fine unemployed offenders, their sentences were more likely to be up-tariffed to a community order or probation. Therefore, unemployed offenders are being sentenced more severely due to financial situations.

In the study by Hood (1992), it found that unemployment was treated as an

aggravating factor but only for Black offenders, this did not apply to the White or Asian offenders in the sample. The interaction between Black and unemployed led to harsher penalties.

Other studies have looked at the effects of living accommodation and offending, such as Grimshaw (2002) and Kirkwood and Richley (2008). Their research suggests a cyclical link between committing crime and becoming homeless, and once homeless committing crime. This link is perpetuated by drug use: many will take drugs to help cope with their experiences and their addiction will propel them to offend to feed their addiction (Arnull et al., 2007). However, there does not appear to be any research which has looked at the relationship between accommodation status and sentencing.

Treating offenders differently goes against the principle of equality (Ashworth, 2010b). Certainly fines are means tested but for any other sentence, Ashworth (2010b) explains, offenders should not be treated differently on account of their financial circumstances. As far as I am aware, there appears to be a gap in the research which investigates the relationship between the financial circumstances of the offender and sentencing, although perhaps employment status is indicative of the offenders financial status.

3.5.4 Characteristics and attitudes of the judge

Judges are the people making the decisions on sentencing. They hold a lot of power and although, as I have explained, guidelines have been brought in to minimise their discretion, they still have the overall say in the sentencing of offenders.

Hogarth (1971) maintains that sentences can be better explained by knowing more about the judge rather than the particulars about the case. Research by Johnson (2006) would endorse this claim. In looking at sentencing disparity in the U.S., Johnson found that the effect of the judge's race was noteworthy: ethnic minority judges were less punitive and therefore less likely to incarcerate offenders. Where

they did incarcerate offenders, the sentences were 6% shorter than those given by White judges (Johnson, 2006). The age of the judge was also significant: older judges were less likely to incarcerate, and when they did they also gave shorter sentences (ibid). However, the influence of the judge's gender was found to be minimal (ibid).

Judges are human and therefore it is important to recognise that they are open to being shaped by influences from their personal backgrounds and experiences (Hogarth, 1971) just like everyone else. They are not immune to stereotypes and common conceptions of certain groups within society. As I mentioned previously, Johnson (2005) rightly argues that the social context of the court has an influence on courtroom decision-making, that is because, judges and other courtroom actors do not make decisions in a social vacuum. These decisions are influenced by social, political and organisational environments (ibid). Judges may also have different views about the effectiveness of the various sentences and may favour one type of disposal over another (ibid). Thus, sentencing practices are likely to vary across courts, as well as, within courts.

This argument finds itself in the realms of legal theory. I will not go into this in great detail here but the point needs mentioning. The philosophy of law and legal theory holds that judges apply legal reasons to the facts of a case in a rational, mechanical, and deliberative manner (Danziger et al., 2011). Judges are supposed to be fair and consistent in their decision-making because they base decisions solely on legal rules and reasons (Leiter, 2005). This is known as Legal Formalism. However, Leiter explains the twentieth century movement of Legal Realism was the most indigenous jurisprudential movement in the U.S.. Legal realism argues that the rational application of legal reasons does not sufficiently explain the decisions of judges and that psychological, political, and social factors influence judicial rulings (Danziger et al., 2011). Therefore extraneous factors, such as gender, ethnicity or socio-economic status, may sway decisions, out with the legal factors of the case.

Research by Danziger et al. (2011) tested this theory. Their study examined the decisions made by a sample of judges in Israel on granting parole to prisoners. The study sought to assess whether the judges decision-making was consistent throughout the day when making repeat decisions. They sought to investigate whether judges were more favourable in their decision-making following a break in which they had time to rest and eat etc. Their study split the day into three sections to access functionality first thing in the morning and following two breaks throughout the day.

The results showed the likelihood of a favourable decision was much greater at the start of the working day and directly following breaks than later in the sequence of cases. The probability of a favourable ruling declines from approximately 0.65 at the start of the session to nearly 0 and jumps back up to around 0.65 following a break (Danziger et al., 2011). They then accounted for the legal characteristics of the case (seriousness of the original crime, months served, previous incarcerations and rehabilitation programme), characteristics of the prisoners (sex, nationality), and the proportion of favourable rulings to that point in the day, using logistic regression with the ruling decision, and the judge-specific fixed effects. While controlling for all these covariates, the pattern of favourable decisions still held (Danziger et al., 2011).

Ashworth (2010b) explains that it is fairly well established that a major source of disparity in sentencing is the different penal philosophies among judges. Anderson and Spohn (2010) also comment on the varying sentencing philosophies of the judges. A judge with a retributive philosophy is more likely to ignore the extra-legal characteristics of the offender, focusing primarily on the seriousness of the offence. A utilitarian judge may give a harsher sentence to an offender who is deemed less likely to be rehabilitated based on their perceived risk of reoffending. This point is summed up in this quote by Hogarth, who explains that judges

"differ widely in their purposes, their views as to the effectiveness of dif-
ferent kinds of sentences, the criteria applied in deciding between different

kinds of sentences, the ways in which conflict between the offender's needs and community protection is resolved, the amount of information support they have for their views, and the kinds of situation in which they experience difficulty in sentencing" (Hogarth, 1971, p. 91)

Therefore, although the Criminal Justice Act 2003 states that sentencers must have regard for all five of the sentencing principles (see Subsection 2.2.2), there is no way of adequately ensuring this. Therefore, this is one way in which disparity is likely to occur when there is a lack of guidance as to how this can be achieved as well as different underlying philosophies of the judiciary.

Judges may sentence offenders more severely, not because of the characteristics of the offender but rather because they are predisposed to using one form of disposal (Johnson, 2005) or believe one type of disposal is more effective than another. This again highlights judges do not make decisions purely on legal facts and rules. Discretion may leave room for disparity which may result in certain groups being discriminated against.

In England and Wales the judiciary has also come under a number of criticisms, largely because of the use of lay magistrates (Flood-Page and Mackie, 1998). Research conducted in the mid-1990s compared the sentencing patterns of stipendiaries (now District Judges) and lay magistrates. Flood-Page and Mackie (1998) found that stipendiary magistrates were more likely to sentence an offender to custody than lay magistrates who made more use of community sentences (ibid).

However, Flood-Page and Mackie (1998) also point out the workloads of stipendiary magistrates are often very different from lay judges, and stipendiary judges are also more likely to be involved in more complex cases lasting several days, making it more difficult to compare the two. Using logistic regression and type of magistrate, sex of the offender, nature and number of offences, previous convictions or similar convictions and whether they were subject to any other court orders at the time of

the offence, Flood-Page and Mackie (1998) found the type of magistrate had an independent effect on the outcome and confirmed lay magistrates used custody less. Although, the offence type, whether the offender was already subject to a court order and total number of offences that they were being sentenced for had a bigger effect on the outcome (ibid).

Additionally Flood-Page and Mackie tested their hypothesis that the presence of stipendiaries in courts beside lay magistrates may have an influence on the decision-making of sentences. Stipendiary magistrates occasionally contribute to the training of lay magistrates and in some areas sit with them on the bench (Flood-Page and Mackie, 1998; Morgan and Russell, 2000). Flood-Page and Mackie (1998) found evidence of an increase in severity of sentencing by lay magistrates in courts with stipendiaries. Morgan and Russell (2000) also tested the theory and found similar results.

Morgan and Russell (2000) carried out similar work to Flood-Page and Mackie in which they observed 930 occasions where the magistrates passed sentence. They found that stipendiaries were more likely to use higher sentencing tariffs (25% of cases resulting in immediate custody compared to just 12% for lay magistrates, and 12% of cases resulting in probation compared to 8% for lay judges) and less likely to use lower tariff sentences (8% (stipendaries) and 28% (lay judges) for conditional discharges and 12% (stipendaries) and 39% (lay judges) for fines respectively). Whereas stipendiaries used probation orders to a greater extent than the lay judges, the reverse is the case with community penalties and compensation orders - 10% for lay magistrates and 7% for stipendiaries.

The argument surrounding lay and profession judges also extends to the social characteristics of these judges. Lay-judges are supposed to represent a more democratic and socially representative judiciary but studies have shown this not to be the case. Cavadino and Dignan (2007) report lay judges are still overwhelmingly middle-aged and middle class, very few are from ethnic minorities and they tend to be more

'middle-minded', with a higher proportion of them inclined to vote Conservative than the general population. Although women now make up around half of lay judges this was not always the case. Crown Court judges tend to be slightly older than magistrates, are predominantly male, and have even fewer members from ethnic minorities (Cavadino and Dignan, 2007). The fact they are previously barristers is said to narrow the backgrounds that judges come from. Morgan and Russell (2000) suggest that in reality the social composition of both lay and stipendiary judges are likely to be very similar, and both very removed from the spheres of those appearing before them.

There is a general perception among defendants that trials at the Crown Court are fairer than trials at magistrates' courts (Ashworth and Redmayne, 2010) and many defendants opt to go to trial at the Crown Court. In doing so, offenders run the risk of being sentenced more severely due to the increased sentencing powers of the Crown Court. However, Ashworth and Redmayne (2010) explain there are a number of reasons why an offender would opt for a Crown Court trial: juries are fresh to cases and are not 'case hardened' in the way magistrates may be; previous convictions are not known to juries which may benefit the offender; and finally, juries tend to be more representative of the general population and therefore are perhaps more likely to understand the predicament of the offender compared to the judge.

This section has shown that it is not only offender characteristics that can lead to disparate sentencing but also the characteristics of the judge. As Johnson (2005) adequately phrased it - sentencing does not happen in a "social vacuum" and therefore a number of extraneous factors can affect the sentencing decisions. These factors need to be taken into consideration when looking at disparity in sentencing, as has been done in a number of studies (mainly from the US). However, information on the characteristics of judges are not recorded in any official data in England and Wales. This makes it extremely difficult to control for this.

3.5.5 Geographic location of the court

Kautt (2002) compared American sentencing to real estate, suggesting that sentencing depends significantly on the geographic location of the court. Ulmer and Johnson also comment that "if the sentence one receives and the grounds for that sentence depend on location, then the notions of equal justice that underlie most Western legal systems may be undermined" (Ulmer and Johnson, 2004, p. 137). Therefore regardless of where about in the country an offender is sentenced they should be treated with parity.

Research has shown there to be differences in the sentencing practices between courts. For example, Halliday (2001) published findings which revealed an 'unnamed court' sentenced half of its domestic burglary offenders to immediate custody, where as another 'unnamed court' court only sentenced 13% of the domestic burglary offenders to the immediate custody. Similarly, for the offence of Actual Bodily Harm, courts ranged from sentencing 2% to 53% of offenders convicted of Actual Bodily Harm to immediate custody. This indicates major sentencing disparity.

Hood (1992) also found sentencing variation between courts and judges, which was initially masked by only looking at the overall average effect of race. For example Birmingham Crown Court sentenced a much lower proportion of Black offenders to custody than would have been expected, whereas the Dudley courts sentenced a much higher proportion of Black offenders to custody (Hood, 1992). This disparity had initially been over-looked.

In the US, Ulmer and Johnson (2004) found that sentencing varied across the county courts: courts were more likely to sentence offenders to jail depending on the jail capacity. In areas where there was more jail capacity, more offenders were sentenced to jail after controlling for the legal characteristics of the offence (Ulmer and Johnson, 2004).

In England and Wales, the Ministry of Justice published results of a study which

investigated the geographic variation in custodial sentencing (Mason et al., 2007). Using custody rates and average custodial sentence lengths (ACSLs) for the 42 Criminal Justice Areas (CJAs) in England and Wales, they compared all recorded crime offences sentenced in the magistrates' and Crown Courts between 2003 and 2006. The study found large variations in custody rates, average custodial sentence lengths and the use of Life and IPP sentences across the 42 CJAs. In particular the custody rates in the magistrates' courts and the use of Life and IPP sentences in the Crown court had the most variation (Mason et al., 2007).

The results showed that the average custody rates in the magistrates' courts were fairly stable between 2003 and 2006, although the actual range had narrowed from 15.7% in 2003 to 9.8% in 2006 (ibid). In 2006 the average custody rate was 11% and half of the 42 (21) CJAs were within 1.7 percentage points of the average. Of those CJAs with custody rates in the top five in 2006, three (Essex, Bedfordshire and London - Metropolitan and City of London police force areas were combined) were consistently in the top five for all years. Similarly those CJAs in the bottom five for 2006, two (Dyfed-Powys and Lincolnshire) were also in the bottom five for 2003, 2004, and 2005 (ibid).

They report ACSL also remained fairly stable between 2003 and 2006. In 2006 the ACSL at magistrates' court was 3.0 months with a total range from 2.3 months to 3.6 months, and half of the CJAs were within 0.2 months of that average. Of those CJAs in the top five for 2006, two (Northumbria and Cumbria) were consistently in the top five for 2003, 2004 and 2005. Likewise in the bottom five for ACSL in 2006, two CJAs (Devon and Cornwall and Essex) were in the bottom five for all years.

Turning to the Crown Court, the custody rates were also fairly stable between 2003 and 2006, apart from the small fall in custody rates between 2005 and 2006. The authors attribute this fall to the introduction of the suspended sentence order, which was introduced in 2005. Although, the same pattern of custody rates did

not occur in the magistrates' courts in which suspended sentence orders were also available, so it may be there is an alternative explanation for this fall. In 2006 the average custody rate was 56% with a total range from 45% to 68%. Half of the 42 CJAs were within 3.2 percentage points of the average.

ACSL has also remained relatively stable over the period of study between 2003 and 2006. The ACSL in 2006 was 24 months with a range of 19.5 months to 28.3 months with half (21) of the CJAs being within 1.8 months of the average. The range of ACSL has narrowed from 15.2 months in 2003 to 8.8 months in 2006. Of those CJAs in the top five for ACSL in 2006, two (Sussex and London) were consistently in the top five. In the bottom five in 2006, only one (South Wales) CJA was also in the bottom five during 2003, 2004 and 2005 (ibid).

Mason et al. (2007) explain the two measures - custody rates and ACSL - were influenced by a variety of factors including legislative changes. For example the introduction of Indeterminate Sentences for Public Protection (IPPs) in April 2005 resulted in a slight fall in ACSL in the Crown Court. In 2006 this average ranged from 1.3% to 4.6%, and 21 of the 42 CJAs were within 2.5 percentage points of the average.

Mason et al. (2007) point out that although their analysis found large variation in custody rates, ACSLs and the use of Life and IPP sentences, the sentencing practice of individual CJAs was relatively consistent. The majority of variation in sentencing was down to a relatively small number of CJAs.

In their analysis, they found there to be a link between the seriousness of the offence and sentencing practice in the Crown Court. Although at best 38% and 20% of the observed variation in custody rates and ACSL respectively can be attributed to the difference in the seriousness of offences between CJAs (ibid). Therefore this only partially explains variation in sentencing, leaving the rest unaccounted for. To add to this, there was a lack of a clear relationship between offence seriousness and sentencing at the magistrates' court and so implies that the mix of offences sentenced

in magistrates' courts was not a significant factor in explaining sentencing variation across the CJAs (ibid).

Furthermore, they found committal practices did not explain geographical variations in sentencing. Neither could the differences in local crime rates and the changes in recorded crimes and sentencing in the magistrates' and Crown Courts, suggesting these two are not linked with sentencing in CJAs (Mason et al., 2007).

Mason et al. (2007) concluded that variation could not be explained solely in terms of the characteristics of the cases or offenders coming before the courts. These factors do play a part in the variation but cannot fully account for the geographic variation. Variations may be down to 'local justice'[17] or the 'human factor' (Tarling and Weatheritt, 1979), which maintains that consistency within individual courts is more important than maintaining consistency at a national level. Finally Mason et al. (2007) acknowledge the relationship between judges and other CJAs agencies who inform the sentencing process, were identified as possibly influencing sentencing practice.

3.6   Concluding remarks

The literature on sentencing disparity draws from a number of different themes such as its association with consistency and inconsistency, discrimination, equality and inequality. All these notions then resonate with injustice. There are also a number of beliefs about how justice can be achieved: one is to argue justice can be achieved through parity and therefore disparity leads to injustice; whereas an alternative approach would suggest disparate sentencing is necessary to achieve justice in an unequal world. These two approaches are somewhat debated in the literature.

However, there may be a way around this if we were to consider equivalent severity

---

[17]Local justice refers to the fact that not all disposals will be available in every area especially when disposals are being piloted (Mason et al., 2007)

or parity in terms of equivalent sentence severity, as opposed to the same sentence or the same treatment. This would allow for an objective measure in which to assess sentencing disparity.

To date, there have been a substantial number of studies which have sought to investigate sentencing disparity, more so in the US than in England and Wales, and mainly concentrating on gender and ethnicity. There have been some studies which have also considered additional offender characteristics, such as employment status and accommodation status, characteristics of the judiciary in terms of the gender, age and ethnicity, and geographic location of the court. Nevertheless, the overwhelming picture suggests these extra-legal factors provide the basis for disparity to occur.

In the following chapter I will explain how these studies which investigate sentencing disparity generally measure the dependent variable - sentence severity, and the problems associated with these measures. I then go on to suggest an alternative approach which will ultimately provide a more robust measure of sentence severity and with that more reliable research into sentencing disparity.

Chapter 4

Previous work on measuring sentence severity and assessing disparity

4.1 Introduction

The previous chapter reviewed the literature on disparity in sentencing. It highlighted that much of the previous research has identified four main factors; offender's gender and ethnicity, characteristics of the judge and location of the courts contribute to disparity in sentencing. At the end of the previous chapter, I suggested there are some weaknesses in this research due to the lack of a clear and explicit definition of disparity, and the methodologies employed to measure sentence severity, which tend to be based on hypothetical scenarios or custodial sentences. This chapter will firstly explain why we measure sentence severity, then discuss the dimensionality of sentence severity and will finally go on to explain the methods which have been adopted to measure sentence severity and disparity in this thesis.

4.1.1 Why measure sentence severity?

Sentence severity can essentially be thought of as a latent concept because it is not directly measurable. Sentence disparity is intertwined with sentence severity: only by measuring sentence severity can we begin to assess disparity. Disparity is said to occur when 'two like-situated offenders' are sentenced for the same offence, and where one offender receives a more severe sentence than the other. For example, after controlling for the offence type and legal characteristics of the case, one offender receives a custodial sentence of 6 months and the other offender receives a custodial sentence of 9 months. However. it becomes more difficult to assess sentencing disparity when offenders receive different types of sentences (von Hirsch, 1993). The different punitive elements (i.e. monetary, restrictions on time or liberty) of sentences handed down by the courts, make it very difficult to compare the different types of sentences. Therefore it can also be thought of as a problem in alignment - how do we determine the community order equivalence of a 30 day custodial sentence?

To facilitate investigation into disparity it is necessary to quantitatively measure sentence severity. In 1978 Leslie Sebba wrote;

> "In view of the proliferation of sophisticated techniques of quantification currently applied in criminological research, it is rather surprising to observe that in one vital area the development of such techniques seems to have been neglected. That is the problem of the severity of penalties meted out by the criminal courts" (Sebba, 1978, p. 247).

Some thirty years later we appear to be not much further ahead.

The practice of measuring sentence severity has generally existed within two spheres; the data driven approach and the opinion/perception approach. These spheres have remained distinctly separate in the methods used to measure sentence severity. The data driven approach uses existing data or data that has been collected specifically for the research by survey or from administrative records. This

data is then analysed using a variety of statistical techniques and from this, disparity in sentencing can be investigated. This approach differs from research conducted in the opinion/perception approach, which asks the public or experts their perceptions and opinions of sentence severity. This method uses a vignette approach, whereby members of the public or experts (i.e. judges, probation officers, police officers etc.) are asked to rate the severity of sentences and apply corresponding scores to each of them. In doing so this creates a series of scores which make up a severity scale. I will refer back to these approaches throughout the course of the chapter as I discuss the various methods used to explore sentence severity and disparity.

4.2   Dimensionality of severity

Sentence severity cannot be directly observed or easily measured. However, the principle of proportionality permits that punishments/sentences can be ordered on a scale so that their relative severity reflects the seriousness ranking of the crimes involved (von Hirsch, 1993). Indeed Sweeten (2012) argues that determinations of seriousness are intertwined with the notion of unidimensionality. In this vein, sentence severity can then be thought of as a unidimensional concept of the punitiveness of a criminal sentence given in a court of law. In fact, a great deal of literature in this area also considers offence seriousness and sentence severity to be a unidimensional concept whereby offences and/or sentences can be arranged on a continuum from the least serious/severe to the most serious/severe sentence (for example Sellin and Wolfgang (1964); Francis et al. (2001); Sebba (1978); Sweeten (2012)).

Another type of scale that has been used previously is a qualitative or descriptive scale. This can be seen as a multidimensional representation of sentence severity but the qualitative factors, such as mobility in the community and financial obligations, are incorporated into a single unidimensional scale. Schiff's (1997) research did just

that by incorporating a number of the punitive elements that make up a sentence and from that calculated an overall severity score. More will be said about this study in section 4.3.3.

There are some that are critical of this unidimensional approach to sentencing (for example Marinos (2005); Braithwaite (1982); Garland (1990)), deeming it as "...defective and unsuccessful because of the lack of consideration of the multidimensional qualities of the offence and the offender" (Marinos, 2005, p. 44). Braithwaite (1982) argues that the sole focus on severity (as well as just-desert and the idea of proportionality) ignores the complexities associated with offending and the offender. Indeed many different factors need to be taken into consideration when sentencing offenders, including the "purposes and functions of individual punishments"(Marinos, 2005, p. 44). Marinos (2005) also states that in order to embrace the multi-dimensional nature of sentencing it is also important to take into consideration the nature of the offence and the age of the offender, as well as the broader relationships among the purposes of the punishment and the individual sanctions. This suggests that the different punishments seek to serve different purposes, such as denunciation, rehabilitation, incapacitation etc (ibid). This also links back to the point made by Tonry (1996) and Kolber (2009) (see Section 3.3.1) that sentences need to take into consideration the effects that different sentences can have on the individual offender. These views emphasise the different theoretical approaches to sentencing and their conflicting ideologies.

However, I would argue that a scale based on sentence severity does not reject the complex nature of sentencing, instead we can think of severity as just one-dimension or aspect of sentencing. Similarly, although referring to offending, Sellin and Wolfgang (1964) argued that offending is multidimensional but insisted that it is possible to use a "...homogeneous dimension which can allow for quantitative measurement of this phenomenon" (Sellin and Wolfgang, 1964, p. 338). Likewise, it is possible to use the

severity as a homogeneous dimension of sentencing to allow for the measurement of disparity.

Furthermore, the sentencing framework in England and Wales adopts a hybrid approach to sentencing, proportionality remains the guiding principle in sentencing offenders. It therefore seems sensible to use this existing principle to model sentence severity and to provide a measure in which to then assess sentencing disparity

4.3 Methods to measure sentence severity

This section describes the methods used to measure sentence severity.

4.3.1 In-out decision to incarcerate

Within the data driven approach, one of the methods that is used to measure sentence severity is the 'in-out' decision to incarcerate. This method constructs a binary variable, coded 1 if the offender was sentenced to immediate custody and 0 if they received any other disposal e.g. a fine or community penalty. This is thus a very simple and crude measure of sentence severity, and implicitly assumes that any length of custodial sentences is more severe than any non-custodial sentences. To assess disparity, the method then models the probability that an offender will receive a custodial sentence (Ulmer and Johnson, 2004). This can be used to assess for example, whether being male or female increases or decreases the probability of receiving a custodial sentence.

This method has been popular, with studies investigating sentencing disparity carried out by, for example; Wheeler et al. (1982); Ulmer and Johnson (2004); Merrall et al. (2010); Holleran and Spohn (2004). Wheeler et al. (1982) explain that in interviews conducted with judges, they explicitly state the hardest decision they make is whether to incarcerate or not, therefore it is imperative that research investigates

the decision to incarcerate. These studies also explain there are different variables which impact on the decision to imprison and in giving a custodial sentence, the ultimate length of that sentence. Ulmer and Johnson (2004) state it also important to model the in-out decision and length of the custodial sentence separately to test for different hypotheses. A number of studies model the in-out decision, as well as the length of custodial sentence in their research. I will say more about this in the following section.

The methodology that focuses on the 'in-out' decision to incarcerate creates a very simplified measure of sentence severity and implies that all custodial sentences are more severe than any non-custodial sentences. It dismisses the possibility that two different sentences, for example a short custodial sentence and a community order, can have the same level of sentence severity or parity.

The 'in-out' decision method also omits the magnitude of the sentence. It is not possible to determine whether a sentence is, for example twice as severe than another. It assumes all custodial sentences are more severe than any non-custodial sentence. Consequently, this method of measuring sentence severity would appear to be a rather crude and erroneous measure.

4.3.2   Custodial sentence length

The second and most popular method that is generally used to measure sentence severity also falls within the data driven sphere. Studies which use real sentencing data will tend to only include the principal offence[18] . These studies then use the custodial sentence length to determine sentence severity: longer sentences are seen as more severe then shorter custodial sentences. This method is used in a number of studies, such as those conducted by; Anderson and Spohn (2010), Britt (2009), Albonetti (1997), Bushway and Piehl (2001), Helms and Jacobs (2002), Ulmer and

---

[18]Where the offender has been found guilty of one or more offences, the principal offence is the most serious offence which receives the heaviest/most severe sentence (Ministry of Justice, 2010)

Johnson (2004), Mueller-Johnson and Dhami (2009), Wheeler et al. (1982), Miethe and Moore (1986), Mustard (2001), and Pina-Sánchez and Linacre (2013).

Studies vary in how sentence length is determined. Studies conducted in England and Wales, such as the study by Pina-Sánchez and Linacre (2013), used the length of custodial sentences recorded by the judges' in the Crown Court Survey. However, this differs to those studies conducted in the U.S. especially where the state in which the research was conducted use a sentencing grid, such as in Minnesota, where the sentencing grid will provided a minimum and maximum presumptive sentence range for cases that share the typical criminal history and offence severity characteristics (Minnesota Sentencing Guidelines Commission, 2013). For example, in looking at sentencing disparity and departures from the pre-defined guidelines, Helms and Jacobs (2002) used the awarded minimum and maximum sentence lengths transformed into months, summed the two and then averaged them. They then added 1 (to avoid taking the log of 0 in cases where the offender received a non-custodial) and took the natural logarithm. In contrast, Britt (2009) used the minimum number of months the offender was sentenced to custody, and again took the logarithm of that sentence length. Taking the natural logarithm of sentence length reduces the positive skew, as the majority of offenders receive relatively short sentences (Britt, 2009); (Albonetti, 1997). Others have used the midpoints of sentence length (Miethe and Moore, 1986) and the average sentence length (Mustard, 2001) to model sentencing disparity.

However, as with any method, there are drawbacks of using this custodial sentence length approach. This method focuses on only one type of disposal - immediate custody. It therefore dismisses a large proportion of sentences which do not result in immediate custody. In doing so, Leslie Sebba comments it "ignores the disparities in terms of severity found among non-custodial penalties"(Sebba, 1978, p. 250). This method inevitably leads to bias in the data, as it excludes important sentencing information (Merrall et al., 2010) by omitting the multitude of sentences which do

not result in a custodial sentence.

By omitting all non-custodial sentences, this method does not facilitate the idea that certain sentences can be equivalent in terms of their severity, whereby a non-custodial sentence i.e. a community order, could be as severe as certain custodial sentences. Therefore it rejects the idea that two different disposals can have the same level of severity.

A number of studies have opted to combine the 'in-out method' with the length of custodial sentence to measure severity, and ultimately sentencing disparity. Studies such as those conducted by Bushway and Piehl (2001) and Albonetti (1997) combine the decision to incarcerate with sentence length, while others have opted to model them separately, for example Wheeler et al. (1982), Ulmer and Johnson (2004), and Merrall et al. (2010). There are arguments for and against modelling the probability of custody and custodial sentence length together, which will be discussed in Section 4.4.

4.3.3 Severity scales

Sebba (1978) argues that in order to conduct any comprehensive analysis on sentencing practice, it is imperative that the various disposals are placed on a single scale. This is so all disposals can be considered "simultaneously, and in a meaningful way" (Sebba, 1978, p. 249). Consequently, a third way of measuring sentence severity does just that. It involves constructing a single scale to allow all sentences to be incorporated, and so the focus is not on just custody which the previous two methods rely on but the entire spectrum of sentences meted out by the courts in England and Wales. Generally these scales have fallen into the opinion driven approach of measuring sentence severity due to the methods used to construct these scales.

The diagram in Figure 4.1 represents the unidimensional severity scale which is based on the principle of proportionality, and more specifically resonates with von

Sentence severity scale
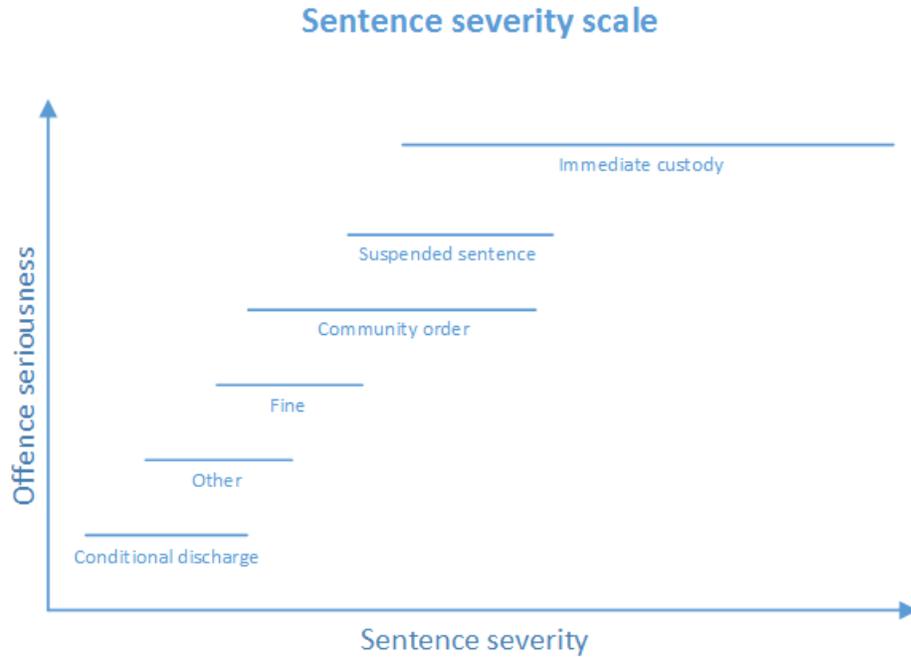
Figure 4.1: Sentence severity scale

Hirsch's (1993) notion of ordinal proportionality. This requires that sentences are determined by the seriousness of the offence, and therefore offenders that have committed the same gravity of offence should be treated with like severity, but not necessarily the same sentence (ibid). Sentences must however be comparable (Duff, 2001).This ties in with parity, which I discussed in the previous chapter. Sentences must be ranked in order of their severity based on the ranked seriousness of the corresponding offence. Thus, as the seriousness of the offence increases (y-axis) it passes through the various thresholds for each type of disposal, whether that is a fine, a community order or the immediate custody threshold. The distribution of disposals are superimposed onto a single scale of sentence severity (x-axis). The spacing on the scale then represents the relative severity of each sentence in conjunction with the other sentences in the scale (ibid).

Morris and Tonry (1990) refer to this type of scale as a "continuum of sanctions", and by arranging sentences in this way, it is possible to determine how sentences

relate to each other and in relation to imprisonment. When sentences are ordered by their severity, like in this example, sentences can overlap in terms of their comparative severity. When von Hirsch (1993) refers to parity, he refers to sentences that have the same level of onerousness and not necessarily the same sentence. When two sentences are of comparative severity, von Hirsch (1993) refers to this as the sentences having an equivalent penal bite. In von Hirsch et al.'s (1989) earlier paper, they address the ways in which non-custodial sentences can be arranged on a penalty scale and how much substitution is acceptable. The first model he discusses is the 'No Substitution model', whereby sentences are arranged on a continuum but the non-custodial and custodial sentences do not overlap. The second is the 'Full Substitution model' which operates on a penalty unit basis. Therefore all sentences are given a score, which can then be substituted with a number of non-custodial sentences which equal the same penalty unit amount. However, von Hirsch et al. (1989) suggest this type of model is complex and overly complicated and therefore may not be useful. The third model is the 'Partial Substitution model'. Here there are a number of bands and within some of these bands it would be possible to substitute penalties of equivalent severity. This is the most useful of the three. Therefore, I argue that different types of sentences can be used interchangeably as long as they carry the equivalent 'penal bite'. The idea of interchangeability, meaning that sentences can overlap in terms of their severity can be seen in the diagram. For example, an onerous community order can have the same severity as a short-term custodial sentence: both penalties carry the same severity-equivalence. I will return to this point in Section 10.3 where I examine the implications of such a scale.

The 1960's, 1970's and 1980's witnessed a large number of studies which focused on scaling offence seriousness and sentence severity. The work by Sellin and Wolfgang in 1964 appears to have started off this trend which continued for some years. However, since the mid-1980's there have been very few studies which have looked at creating

similar scales, suggesting a drop in popularity. The late 1980's/early 1990's saw a brief revival in scaling research in the U.S. due to the introduction of intermediate sanctions, defined as punishments with middle-range severity, employable for crimes of medium and upper-medium seriousness (von Hirsch et al., 1989). However, the methods and techniques did not change or move forward. The remainder of this section will outline the various types of severity scales which have been created in order to measure sentence severity.

Research conducted by Tiffany et al. (1975), used a pre-existing sentence severity scale based on a version created by the US Administrative office for their reporting purposes[19] . Arbitrary scores were assigned to a range of 16 sentences, including suspended sentences, fines and custodial sentences (for example a suspended sentence was assigned the value of 0, a fine was assigned the value of 1, 1-6 months imprisonment was assigned a value of three, and imprisonment of over 120 months received a value of 50) to create a single severity scale. The scale was then used as the dependent variable in which to model sentencing disparity.

Similarly, Malila (2012), in looking at disparity in sentencing in Botswana, assigned arbitrary scores to different types of punishment to represent their severity, relative to the other punishments. Malila then identified the severity of multiple punishments by adopting a unit scoring approach, similar to that employed in research by Schiff (1997). I will return to Schiff's (1997) research later in this chapter.

Studies like those conducted by Tiffany et al. (1975), and Malila (2012), have received criticism for assigning arbitrary scores to sentences (McDavid and Stipak, 1981); (Klepper et al., 1983). Klepper et al. make the case that although these scales are generally reasonable, questions arise as to whether "findings are simply an artefact of an artificial scale" (1983: 58). Consequently, using a different scale could

---

[19]Tiffany et al. (1975) explain that the scale implemented by the U.S. Administrative Office was developed by James A. McCafferty in a paper entitled 'Weighting', and was presented at the 96th Congress of Correction in 1966, Baltimore, Maryland. Attempts have been made to source this paper but they have been unsuccessful.

give different results. There are also issues with how to interpret these scores. Without the relevant context, the scores are not easy to interpret and are not intuitive.

In the past severity scales have also been constructed on the basis of perceived sentence severity. Here, groups of people in society, such as members of the judiciary, the public, students, or police officers, were asked to score sentences in respect to their perceived severity. Often research of this kind used hypothetical offences, offenders and cases, and asked respondents to decide on an appropriate sentence. The outcome was a series of scores which were ranked in order of severity within a single scale.

Studies such as those conducted by Erickson and Gibbs (1979), Sebba and Nathan (1984), Allen and Anson (1985), and Harlow et al. (1995), used a psychometric scaling technique called magnitude estimation to create a sentence severity scale. Magnitude estimation employs a survey method of perceived severity, where respondents are required to rank a sample of sentences based on their perceived severity. Respondents are presented with a list of sentences, which are randomly ordered and where one penalty - for example a 1 year custodial sentence - is anchored, and given a value of 100, or some such value. The respondents are then asked to assign values to the additional sentences relative to the score of 100 (i.e. 1 year custodial sentence). The respondents are asked not to look back at their previous scores so that each score will be measured against the 100 score. The individual scores for each penalty are usually then averaged to produce a final scale with various scores for each sentence.

In contrast, Harlow et al. (1995) used both line production (LP) and number estimation (NE) which are also types of psychometric scaling. LP involves respondents drawing lines to replicate the severity of a series of sentences relative to one another. NE is similar to magnitude estimation in which respondents are asked to score sentences relative to the previous sentence. They acknowledge the scale values are not interpretable on their own because the values only have meaning in comparison to each other but the ratio of numbers corresponding to different sentences are mean-

ingful. They give this example; 1-year probation with a scale value of 54.29 can be considered about half as severe as a 1-year intensive supervision programme, with a scale of 111.79. This gives the scores far more context than previous studies have managed.

Harlow et al. (1995) admit magnitude scaling is time-consuming for both the researcher and respondents: the actual preparation of the surveys and scale validation once the data has been collected is also time-consuming. However, they reason that it allows freedom in determining the adequate range of responses (Harlow et al., 1995). In other words, respondents are not limited by having to choose from pre-chosen categories, which is likely to restrict their perceptions of sentence severity.

Other studies have used paired comparisons to create severity scales. In this manner, two stimuli - sentences - are compared to establish which sentence is perceived to be more severe. In the case of Buchner's (1979) study, she based her assessment of sentence severity on Thurstone's Law of Comparative Judgements using paired-comparisons. This bases the difference in sentences on a 'psychological continuum' (Thurstone, 1994, p. 266) and creates a single severity scale in which all sentences are ordered.

Unlike Buchner's (1979) study which used hypothetical cases, Francis et al. (2001) used real data from the courts. Although, they were modelling offence seriousness rather than sentence severity, the same principles apply. In their study they used data in the form of pairs of real offences sentenced at the same time for the same offender and in doing so, this controls for the various factors which come into the judges decision-making process. Judges then identified the offences they perceived as being more serious within each pair. From this information a preference scale can be constructed, which ranks offences by their seriousness (ibid). Although, this methodology is more robust than the studies which use hypothetical cases, it still relies on a *judgement* of which offence is more serious.

There are problems associated with approaches which focus on people's perceptions of sentence severity. McDavid and Stipak (1981) make the point that it is highly subjective: opinions relating to severity are likely to vary depending on those surveyed. This is evident in the work by Sebba and Nathan (1984) who found that prisoners gave the highest mean scores relating to the various sentences forming the severity scale, followed by students, probation officers and then finally police officers. In contrast, Erickson and Gibbs (1979) found police officers scored sentences more severely than general citizens. This emphasises the subjective nature of studies which base their scales on the perception of sentence severity. Francis et al. (2001) also criticise these types of studies, referring to them as volatile because different societal groups have different views. Also members of the public are more susceptible to the influence of the media and the moral panics which they can create. This will likely affect their perceptions and opinions of sentences (Francis et al., 2001, p. 729-730).

Francis et al. recognise the merit in "tapping into a real situation... using the representativeness of the community whose task is to assess 'seriousness' "(Francis et al., 2001, p. 730) referring to the judiciary and the magistracy who are actually deciding on the seriousness of the various offences and punishing offenders accordingly. Although they were looking specifically at offence seriousness as opposed to sentence severity, they recognise the benefit of using real sentencing data as opposed to perceived or hypothetical cases which may not capture or illustrate the actual truth.

These types of studies are also costly to perform. Francis et al. (2001) comment that these types of studies are unwieldy because they involve expensive fieldwork in the form of interviewing respondents, a point also echoed by McDavid and Stipak (1981). The data to create this type of analysis is not readily available and therefore needs to be collected, which is also time consuming and labour intensive (ibid).

Furthermore, a major drawback of most of these studies is the relatively small

sample sizes that they use. The studies are frequently conducted with a small number of respondents, for example 44 members of the public (Harlow et al., 1995); 51 judges (Buchner, 1979); 81 students (Allen and Anson, 1985); 84 respondents (Sebba and Nathan, 1984). These samples are therefore not generalisable to any one population and the results are likely to vary were a different sample from the same population used in the analysis.

Another type of scale that has been used is a qualitative or descriptive scale. This can be seen as a multi-dimensional scale but where the qualitative factors are incorporated into a single unidimensional scale. Schiff (1997) created the Criminal Penalty Severity Scale (CPSS), and employed what he calls a 'penalty unit approach': whereby each component of a sanction is said to characterise a 'unique and essential dimension of criminal punishment'. This model is based on the Delaware Model, which uses a number of accountability levels to help determine the restrictiveness of a particular sanction. Schiff's (1997) model uses five levels relating to 'restriction to a prescribed space', 'direct supervision of compulsory activities', 'categorical supervision of normal activities', 'resource limitations' and 'length of supervision'. Each penalty is scored in reference to these five categories and each level is also assigned a subjective weight. No score has an absolute value or cardinal meaning but is said to give a sense of relative severity for one penalty vis-a-vis another (Schiff, 1997). The CPSS then computes a numerical value for each individual penalty. Values are created for each separate component of the sanction by multiplying the level of accountability by its weight. The total score is then divided by 10 to make the numbers more manageable to create a score for each penalty. Again, this method is computationally intensive and does not use real sentencing data.

An alternative method that has been employed by McDavid and Stipak (1981) is canonical correlation. Their study used real sentencing data obtained from the Pennsylvania Commission on Crime and Delinquency records for 1977. Their study

used 2321 cases/offenders and seven categories for type of offence and seven categories of various sentences. Canonical correlation extracts a linear combination called a 'canonical variate' from each set of variables (offence and sentence). The method extracts pairs of canonical variates so that the correlation (canonical correlation) between variates from each set of factors is maximised (ibid). The highest correlation is the best. Scores for each category of sentence and offence are then produced. In other words canonical correlation measures the linear relationship between two variables (Borga, 2001). Borga (2001) explains that canonical correlation finds the two bases in which the correlation matrix between offence and sentence is diagonal and the correlations on the diagonal are maximised.

McDavid and Stipak (1981) point out that an advantage of this method is that other variables can be built into the model: it is possible to account for variables which are likely to impact on the severity of the sentence, for example the offenders plea and previous convictions. This will increase the correlation between offence and sentence as these additional variables are being included in the model which explains the relationship of the data better.

The one factor that all of these methods have in common is they produce some form of sentence severity measure. This measure then acts as a dependent variable in which to investigate sentencing disparity. The next section will discuss the methods that have previously been used to measure sentencing disparity.

4.4 Methods to measure sentencing disparity

Assuming that a sentence severity score can be produced, how can this score be used to assess disparity in sentencing? In the existing literature, generally five quantitative methods have been used. This section will discuss each of these methods in turn. It will then discuss an alternative way of assessing sentencing disparity using a

qualitative approach using simulated case studies.

4.4.1 Ordinary least squares regression

Regression analysis is a statistical tool for investigating the relationship between variables (Sykes, 1993). A number of studies, including those by Croyle (1983), Tiffany et al. (1975) and Engen and Gainey (2000) have used regression analysis in the form of ordinary least squares (OLS) to model disparity in sentencing. For example, sentence length is often used as the dependent variable in a regression analysis. A number of independent variables, such as gender and ethnicity are then fitted to the model to assess how these variables affect sentence length. We could then use the results of this type of analysis to predict sentences for offenders with a given set characteristics (Tiffany et al., 1975).

This method is rather simplistic in its nature and therefore is not used in many sentencing disparity studies. However, it is described here to document the evolution of methodology.

4.4.2 Logistic regression

Logistic regression models the relationship between a dichotomous dependent variable (e.g. the 'in-out' decision) and one or more explanatory variables, to estimate the probability of an event occurring. As well as using OLS, Steffensmeier et al. (1993) use logistic regression to model the effects of gender on the 'in-out' decision to incarcerate. This then estimates the probability of being incarcerated for males compared to females in the sample.

Steffensmeier et al. (1993) use both OLS and logistic regression to model the effects of gender on the 'in-out' decision to incarcerate. Although the two methods produce similar results they rely mainly on OLS to report the findings. This is because they feel OLS is more familiar to most readers and it is more straightforward to interpret

than logistic regression (ibid).

4.4.3   Tobit analysis

Studies that have opted to jointly model the decision to incarcerate in combination with the length of the custodial sentence, such as Bushway and Piehl (2001) and Albonetti (1997), have combined these two measures using tobit analysis. Albonetti insists the same set of independent variables are thought to affect the decision to incarcerate and the length of sentence, and therefore, models them simultaneously. In Albonetti's (1997) sample, 91% of offenders received a custodial sentence; it was therefore necessary to take account of those (9%) who, as a result of not getting a custodial sentence received a zero length. This is referred to as left-censoring because the non-custodial sentences, with length of zero, occurs outside of the range of custodial sentence length. Tobit analysis therefore allows for left-censoring to generate unbiased results.

As mentioned previously, Wheeler et al. (1982) and Ulmer and Johnson (2004) however chose not to combine the decision to incarcerate with sentence length. Ulmer and Johnson (2004) state three reasons for this: firstly, their analysis indicated that different independent variables affect the decision to incarcerate and the length of the custodial sentence. Wheeler et al. (1982) also report this. Therefore they believe combining the two would bias their results. Secondly, some of Ulmer and Johnson's hypotheses look to establish the link between prison capacity and the decision to incarcerate. Consequently, combining the two measures would not allow them to test this hypothesis. The third reason they give for modelling both methods separately is that to their knowledge, tobit analysis cannot be used in the multilevel modelling which they used to carry out their analysis. I will return to the method of multilevel modelling later in this section.

4.4.4 Quantile Regression Modelling

Britt (2009) criticises these statistical models just discussed, because they only estimate constant effects of the independent variables on the conditional mean of the dependent variable. This assumes that the effect of the independent variable is constant across all sentence length. Another method that has been identified to measure disparity is quantile regression modelling. This method is similar to regression analysis but looks at the effects of certain characteristics on sections of the sentence length distribution. A quantile is a percentile in the distribution of a variable (Britt, 2009, p. 348).

Consequently, quantile regression analysis in this instance looks at various quantiles within sentence length to identify if and where disparity exists and to what extent. Research conducted by Britt found that the biggest discrepancy between OLS and quantile regression modelling in his research was the effect of being convicted through a jury trial (compared to pleading guilty). Britt reports that the OLS result showed sentence length to be almost 16 months longer than those who plead guilty. Using quantile regression analysis, the results ranged from about 1 month (at the 0.10 quantile) to more than 31 months (at the 0.9 quantile). Concluding that quantile regression modelling is far more accurate than traditional regression analysis. Additionally, like the tobit model, quantile regression analysis allows for left-censoring of the dependent variable to ensure unbiased estimates.

4.4.5 Multilevel modelling

The most common method used in the study of sentencing disparity is multilevel modelling. Studies, such as those conducted by Weidner et al. (2005), Ulmer and Johnson (2004), Johnson (2006) and Pina-Sánchez and Linacre (2013), have employed this method. Once more these studies have used either the decision to incarcerate or sentence length as the dependent variable, and a number of independent variables,

such as gender, and race, to assess where disparity occurs.

Multilevel models are a type of regression model, used to analyse data with a hierarchical or nested structure (Bryk and Raudenbush, 1992), such as individual offenders nested in courts (Pina-Sánchez and Linacre, 2013) or individual offenders nested in courts, nested within counties (Johnson, 2006). They differ from regression models in that the equations defining the multilevel models contain more than one error term (one for each level)(Luke, 2004). Independent variables can then be added to the model, corresponding to the relevant level. For example, in the study by Johnson (2006) age of the offender is an individual characteristic (level 1), size of the court is court level characteristic (level 2) and population demographic is a county level characteristic (level 3).

This method has numerous advantages over traditional analytical strategies such as ordinary least squares (OLS). Weidner et al. (2005) explain that unlike single-level regression models, multilevel models can account for the lack of independence across levels of nested data. Where data are nested, there are likely to be dependencies between individual responses, and therefore not taking this into account can lead to biased parameter estimates. Single-level multivariate techniques are also likely to underestimate standard errors and in doing so exaggerate the significance of parameter estimates (ibid).

Secondly, multilevel modelling allows one to partition the variance for each level in the model, for example, individual and county levels, which will allow the researcher to assess the degree of variation amongst these levels individually (Ulmer and Johnson, 2004). The researcher can then assess the variation that exists when explanatory variables are added to the model and how these help to explain sentencing disparity (ibid).

Another advantage of using a multilevel model is that it does not assume the impact of explanatory variables is the same across all counties (Weidner et al., 2005).

By using a random slope model as opposed to a random intercept model, it may reveal that certain individual explanatory variables, such as race, do not have the same impact across all counties. Weidner et al. (2005) explain that multilevel models are far more accurate at estimating model parameters.

4.4.6   Simulated case studies

Simulated case studies are a qualitative approach used to assess sentencing disparity. This approach focuses on groups of judges to identify if their opinions vary in terms of the types and severity of sentences they impose. In this approach, judges are given a number of identical case studies with hypothetical information relating to a number of offences, the circumstances surrounding these offences and information on the suspects/offenders. In some instances, judges are asked to indicate whether they find the suspect(s) in each of the cases guilty, such as in the study by Austin and Williams (1977). Judges are then asked to impose a sentence for each offender. Palys and Divorski (1986) argue that this approach has the ability to hold cases constant and obtain multiple judgements on identical cases. In doing so, any sentencing disparity will be apparent. However, they also acknowledge that these are not 'real' cases with 'real' consequences, and therefore we must be cautious about the external validity of the results (Palys and Divorski, 1986).

Studies that have used this method have found there to be disparity in the verdict (for example, Austin and Williams (1977)), choice of disposal (for example, Austin and Williams (1977); Palys and Divorski (1986); Maguire (2010)) and magnitude of the penalty (ibid). The type of offence tends to be related to the degree of disparity (Austin and Williams, 1977): judges were more likely to be consistent on the type of disposal where the offence was more serious, and less likely to agree for less serious offences (Maguire, 2010). Although, even where there was agreement on the type of disposal, there was disparity in the magnitude of the sentences, such as the amount

of the fine or length of the community sentence (ibid).

Although these studies identify sentencing disparity, they tend to rely on small sample sizes and may not reflect the population of judges. For example, 15 judges were included in Irish study by Maguire (2010), and 47 judges were included in US study by Austin and Williams (1977). Although, the Canadian study by Palys and Divorski (1986) had a much bigger sample (n ≈ 200). However, these studies are all based on hypothetical sentencing vignettes, and this may not be a true representation of what happens in real life situations. Furthermore, this method makes it hard to detect systematic disparity associated with the personal characteristics of the offender, which may influence sentencing decisions.

4.5   Concluding remarks

In summary, sentence severity can be thought of as a unidimensional concept that can be modelled on the existing sentencing framework of England and Wales. Research in this area has very much been conducted within two distinct spheres. The data driven sphere, uses real sentencing data and the 'in-out' decision to incarcerate, or custodial sentence length. On the other hand, research conducted in the opinion/perception sphere, has tended to use hypothetical cases, where certain groups within society are asked to rate their perceived severity of certain sentences. The outcome of this is a severity scale where sentences are ranked by their perceived severity. All of these methods create an output of sentence severity that can be modelled as a dependent variable in which to then assess sentencing disparity. Again there are a number of possible ways of doing this, such as; regression analysis, tobit regression analysis, quantile regression analysis and multilevel modelling. Alternatively there are also more qualitative method using simulated case studies. However, these have very small sample sizes and are based on hypothetical sentencing vignettes.

Conducting a thorough review of the existing methodological literature in this area has lead me to develop an alternative method to measure sentence severity and assess sentencing disparity. This research will bridge the gap between the methods used within the spheres. In doing so, it will use real sentencing data relating to offences and sentences. I will use a method that has not been previously used in this type of research, called Goodman Row Column Association analysis. The method will be developed and extended in this thesis to allow for legal covariates which affect sentencing decisions. The estimates produced from this method will then be used to build a scale measuring sentence severity. In subsequent chapters this unidimensional scale will then be used as the dependent variable in a multilevel model with levels at the offence level (level 1) and the offender level (level 2). A number of independent/explanatory variables will then be added to this model to identify which covariates impact on and lead to disparity in sentencing. This will lead to a far more technically robust investigation of sentencing disparity than has been conducted to date.

The following chapter will discuss the data and sample used to carry out this research before going on to discuss the methods involved in the process.

Chapter 5

Data and preliminary analysis

5.1 Introduction

The crux of this doctoral research is to investigate disparity in sentencing. However, this requires developing a better measure of sentence severity than has previously been used. The first part of this research therefore sets out to construct a scale of sentence severity; this uses real sentencing data and includes a wide variety of offences and sentences. In doing so, this facilitates the second part of the research which allows the assessment of disparity by using this new measure.

This chapter will firstly discuss the sources of data used in this research and will go on to describe the sample. The final part of the chapter will see the data being arranged into a two-way contingency table which provides the conceptual basis for developing the scale. The two-way contingency table forms the basis for conducting a new and innovative statistical analysis which will be discussed in Chapter 6.

5.2 Sources of data

The research uses data derived from the Offender Assessment System (OASys) merged with data from the Police National Computer (PNC). Applications for data were

made to the OASys Data Evaluation and Assessment Team (O-DEAT) and the PNC Information Access Panel (PIAP) within the Justice Statistics Analytical Services Unit (JSAS) of the Ministry of Justice.

The OASys data is the major dataset used in this research. A sample of the data in the form of pre-sentence reports (as described in section 2.5.1) was collated and later matched with PNC data by O-DEAT following sentencing, specifically for this research. OASys data provides information on the offence, socio-demographic characteristics of the offender, needs of the offender, and their predicted risk of future offending. The PNC data then provides details of the offence, such as: the Home Office (HO) offence code, HO disposal code, duration or amount of disposal, plea, and the previous conviction history of the offender.

5.2.1 Offender Assessment System

The Offender Assessment System was jointly developed and piloted by the Prison Service and Probation Service between 1999 and 2001 (Holden, 2007). A paper-based system was initially used prior to an electronic version becoming available, and by 2003 the Prison Service and Probation Service had both developed their own electronic versions due to different IT structures (ibid). It was not until early 2006 that the systems were amalgamated to allow the exchange of OASys assessments (Howard, 2011). Around the same time the National Probation Directorate issued a new court report framework which standardised the provision of pre-sentence reports and with the implementation of OASys, it was then possible to produce a standard delivery report from the OASys assessment (Holden, 2007).

As discussed briefly in section 2.4, the Offender Assessment System is a tool used to assess the risk and needs of offenders. It is used by the Probation Service generally at the pre-sentence stage - in the form of pre-sentence reports. As well as at the start of most community and custodial sentences, and at regular intervals throughout the

offender's sentence (Howard, 2006). The OASys assessment is generally reserved for offenders that have committed relatively serious offences and are therefore likely to receive a community or custodial sentence.

There are five main parts to the Offender Assessment System;

- Risk of reconviction and offending-related factors

- Risk of serious harm, risks to individuals, and other risks

- OASys summary sheet

- Supervision and sentence planning

- Self-assessment

The first part consists of 12 sections, which includes information on the offender and analysis of the offence, the factors associated with risk, as well as information relating to the health and other considerations such as the suitability of electronic monitoring. These sections are illustrated in the diagram shown in Figure 5.1. The arrows pointing to the offender at the centre of the diagram represents the fact that these factors are said to contribute to the offender's potential future behaviour which are assessed by the actuarial risk assessment.

The second part of OASys looks at the risk of serious harm, risks to the individual and risks to others. This section covers significant events and behaviours, which are also likely to impact on future risk. It also assesses the offender's mental state, i.e. their vulnerability or whether they are coping. It also assess if there are other groups that may be at risk of their offending, for example children.

The summary sheet draws together key information from the assessment and includes a scoring schedule to allow the scores to be tallied. The OASys General reoffending Predictor (OGP) score is out of 100 and estimates the likelihood of the offender reoffending: a higher score indicates they are more likely to reoffend. The supervision

Figure 5.1: OASys risk factors

and sentence planning section includes pre-sentence report plans and review plans. The self-assessment aspect of OASys summarises the attitudes and behaviours that the offender needs to work on and consequently improve throughout their sentence. There is also a confidential section which is not disclosed to the offender.

Missing data

There is a substantial amount of missing data in this sample. Although this does not cause problems for the initial part of this research - constructing the severity scale - it does impact later on in the analysis. As a result, not all the factors shown in

Figure 5.1 can be included in the analysis aimed at identifying sentencing disparity. Furthermore, to allow for a complete case analysis, additional cases had to be removed from the sample. I will go on to say more about this in Chapter 9.

5.2.2 Police National Computer

The Police National Computer (PNC) was introduced in 1974 and was initially used as a database for stolen vehicles. Since then it has grown in size and usage, housing several separate databases as well as links to many others. The PNC is an electronic database containing information on arrests, court summons and court disposals for each offender in England and Wales, as well as Scotland. The police authority initially records the data, which is then updated with information from the courts following sentencing via Crest and Libra; the court systems of the Crown and magistrates' court respectively.

The operational database also records information on local intelligence, co-defendants, known associates, previous addresses, as well as offender's distinguishing scars, marks and descriptions. This includes links to fingerprints and DNA. In general this extra intelligence information is not available to researchers. Additionally, the PNC does not hold information on the social characteristics and circumstances of the offender. Therefore this additional data had to be sought from another source i.e. the Offender Assessment System.

Data quality issues in the PNC

Following sentencing, information from the court systems (Crest and Libra) is uploaded to the PNC. Within the PNC database, up to four sentence disposals for each offence can be recorded, with each disposal consisting of four fields: type of disposal, duration of the disposal, amount of the disposal (fines) and the Home Office disposal code. Custodial sentences and fines appear relatively straight-forward to record. How-

ever, the complex nature of some of the relatively new sentences make these difficult to accurately record in the existing PNC database categories.

The complexity of new disposals such as the community order, make recording them in the PNC very difficult. A community order can have a number of requirements attached to this disposal; these requirements dictate the severity of the sentence. To use an example to demonstrate; an unpaid work requirement can vary from 40-80 hours within the low range; 80-150 for medium range; and 150-300 hours in the high range category under the guidelines (Sentencing Guidelines Counci, 2004). Ideally the PNC would record all of this information, as well as the length of the full order. However, there is currently no way of doing this. Therefore, we only know the length of the overall order.

There are also instances from the data, which suggest offenders are receiving multiple community orders and suspended sentence orders. We know this is not the case because they can only receive one community order or suspended sentence. It is the combination of requirements that reflect the severity of that sentence. It is not clear why community orders and suspended sentence orders have been recorded in this way. It could be interpreted that each subsequent community order or suspended sentence order that is recorded is actually referring to the number of requirements attached to the order but this is merely speculation.

Furthermore, in recording a suspended sentence order, the PNC would ideally record the length of the custodial sentence and the length of the suspension period. However, this is not the case. Only after conducting exploratory analysis on the data did it become clear that the PNC only records the length of the custodial sentence and not the length of the suspension period.

The PNC does not record information relating to the conditions placed on offenders' that are given a conditional discharge. Again, only the length of the conditional discharge is recorded in the PNC.

Finally, it is not possible to differentiate between consecutive and concurrent custodial sentences in instances where offenders are being sentenced for multiple offences. My understanding is that magistrates courts cannot grant consecutive custodial sentences but this does not correspond with the data. This suggests either that, magistrates are giving consecutive custodial sentences or that the Home Office disposal codes are being recorded inaccurately by the PNC operatives. The latter explanation would seem more likely considering the other problems identified with recording sentences in the PNC.

As a result of all these issues, the decision was taken to use only the first disposal recorded in the PNC; based on the assumption that the first disposal is the most severe sentence and likely to be the most accurate.

5.3   The dataset

The original dataset was provided in IBM SPSS for Windows Version 19.0, which was then stored on an external hard drive and held in a secure office. All the cleaning of the data and data manipulation was carried out in SPSS before it was imported into R for data analysis. The following section will discuss the steps taken before any analysis for this research could be undertaken.

5.3.1   Data manipulation

The original dataset contained information relating to 39,520 offenders and 83,423 offences. The data was in a long format: each line in the dataset related to one offence. In total there was 527 variables, and very few of these variables had data labels. Some of the variable names were self-explanatory and therefore did not require additional data labels, for example *psrcase* related to the pre-sentence report case. Other variable names, such as *S1Q3* was less clear initially and would have benefited

from having a data label. However, it was possible to match these variables up with a paper copy of the OASys assessment questionnaire. Additionally, these variables also required value labels which were again obtained from the OASys questionnaire and entered manually into the SPSS dataset.

A substantial amount of time was spent cleaning the data due to the different ways in which certain variables were manually entered into the OASys system. For example, the type of court the offender was sentenced in was recorded, in most instances, as *CRN* for Crown Court and *MAG* for magistrates' court. However, there were a number of variations of this, such as *crn*, *Crn*, *CRO*, and *mag* which therefore required recoding.

In the original dataset the type of offence and the length of the sentence were in two separate categories. These two categories had to be manipulated into one category, for example an Immediate custody sentence of 3 months 6 months. As explained in Section 5.2.2, for community orders, suspended sentences and conditional discharges, the requirements were not recorded and therefore these sentences were also categorised by sentence length.

Research conducted in the data driven sphere will, as a rule, only include the principal offence in subsequent analysis. This will inevitably create an inaccurate depiction of sentencing in practice. To avoid this, and to create a more realistic picture of current sentencing practice, all offences are included in this research. Data was aggregated by the pre-sentence report case number to identify the number of offences the offender was being sentenced for. In this sample 12,753 offenders (52%) were sentenced for only one offence, 24% were sentenced for two offences, and the remaining 24% were sentenced for three or more offences. The average number of offences that offenders are sentenced for is 2.19 (SD=2.47). To take into account the cases where the offender was sentenced for multiple offences, an additional variable was created *Numoff*. This variable was dichotomised: those sentenced for two or less

offences in the first category and those sentence for three or more offences in a second category. It is then possible to investigate the principle of totality.

There is one slight caveat that needs to be mentioned, which concerns custodial sentences in the data: where an offender is sentenced for multiple custodial sentences there is no indication of whether that sentence ran concurrently or consecutively. The new sentencing guidelines which came into operation after this data was collected explain that concurrent sentences may be longer for multiple offences than for single offences to reflect the overall seriousness of the combined offences. On the other hand, sentences that are run consecutively may then be shorter where there are multiple offences, so that the overall time the offender spends in prison is proportionate to the seriousness of the combined offences. This said, we would still expect to see offenders that are being sentenced for multiple offences sentenced more severely than an offender sentenced for a single offence, once the legal characteristics of the offence are taken into consideration.

For any single offence an offender can receive a combination of disposals. For example, an offender can receive a community order with a number of requirements, a fine, and a compensation order. The most severe disposal would usually be recorded as the first disposal. In this example, the community order is the most severe of the three but all three disposals add to the overall severity of the sentence. Every offence will have at least one disposal as a proportion of the 49,590 offences, 60% of the sample received at least two disposals, 42% received at least three disposals and 24% received four disposals. However, the decision was made to only use the first disposal due to problems with the way disposals are recorded in the PNC which has already been discussed in Section 5.2.2.

An additional variable was also created at this stage which indicated whether the offender had previously appeared in court because the dataset did not contain information relating to the offender's conviction history. Information on whether the

115

offender had previously appeared before the court was included and therefore this was used as an indicator of prior offending. *Prior* therefore became a dichotomised variable indicating whether the offender had a prior court appearance to be used in the subsequent analysis.

Offenders excluded from the sample

As explained in 5.2.2, data quality issues relating to the way offences are recorded in the PNC meant that information relating to 4,448 offenders also had to be removed from the sample. An additional 4,355 offenders were also excluded from the sample where the offence plea was missing.

Following the discussion in section 2.7.1, cases resulting in an absolute discharge were also removed from the sample for two reasons: a conviction resulting in an absolute discharge does not count as a conviction for most future purposes, and secondly there were very few cases which resulted in this type of disposal (N=45 offenders).

Under the age of 21, offenders are treated differently within the Criminal Justice System in England and Wales. The decision was taken to only include adult offenders in this research and so all offenders under the age of 21 were removed from the sample (N=6134 offenders).

The final sample size for the first analysis therefore comprised of 24,538 offenders and 49,590 offences.

5.3.2 Demographics of the sample

This section will document the socio-demographics of the offenders in the sample (Table 8.1 also displays these characteristics).

Sex

The majority (88%) of offenders in the sample are male with only 12% being female. There are slightly more male offenders in this sample compared to figures reported in the sentencing statistics for England and Wales 2009 (Ministry of Justice, 2010), where only 77% of the sample are male. The higher number of males in this sample is likely to reflect the more serious nature of offences recorded in OASys data. As noted earlier, Cavadino and Dignan (2007) and Schwartz and Steffensmeier (2007) have suggested that women tend to play subsidiary roles in offences which implies that they are convicted of less serious crimes.

Age

Figure 5.2 shows the range and frequency of ages in the adult sample. The age range of the sample is 21 to 90, where each bar refers to one year of age. The average age of the sample is 34.9 years (SD = 11.29). The sample peaks at 21 and gradually declines with age with a few small unexplained spikes. It is unclear why there is such a strong peak at age 21. Official statistics break age down into age groups; 10-17, 18-20 and 21 and over: there is no information on the individual age years of the offender. It is therefore not possible to compare the age demographics of this sample with the general offending population.

Prior court appearances

This variable relates to the offender's prior court appearances which acts as an indicator variable for their offending history. Just under 5,500 offenders in the sample (22%) had previously not appeared in court, and therefore were first time defendants. The other 78% of the sample had previously appeared in court.
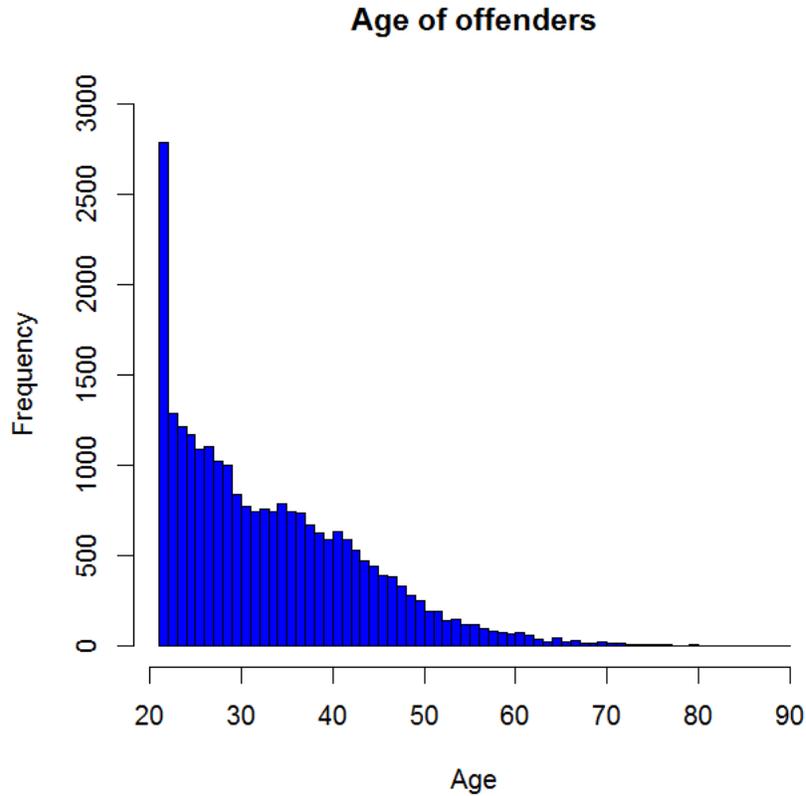
Age of offenders

Figure 5.2: Age of offenders in sample (N=24,538)

Number of offences

This variable records the aggregate number of offences for each offender in the sample, which will allow the principle of totality to be explored. Approximately 76% of the sample were sentenced for one or two offences, the remaining 24% were sentenced for three or more offences.

Court type

58% of the sample (14,178 offenders) were eventually sentenced at the magistrates' court, compared with 42% who were sentenced at the Crown court. This is very different from the figures reported by the Ministry of Justice (2010). The Ministry of Justice report that in 2009 the majority (93%) of offences were sentenced at the

magistrates' court and the remaining 7% were sentenced at the Crown Court (Ministry of Justice, 2010). This discrepancy is likely due to using OASys data. OASys data is collected by the Probation Service in the form of pre-sentence reports and these are only prepared where offences are more serious in nature passing the community penalty threshold. The offences within this sample are more serious in nature than those in the general population.

Plea

Offenders mostly plead guilty to their offences - 84% and 7% plead not guilty. The plea for the remaining 9% of the sample was unknown. The Crown Prosecution Service (2012) report that in 2009-10 68% of cases at the magistrates' court gave a guilty plea, and 74% of cases at the Crown Court gave a guilty plea.

5.4 Primary variables of interest

It is the offence and sentence categories which are of interest initially, as this thesis looks to see if it is possible to incorporate these into a single scale measuring severity. The following sections then discuss the offence and sentence categories and how the data can be arranged to facilitate model building.

5.4.1 Offence categories

PNC Home Office offence codes were grouped into major offence codes, reducing the number of categories of offences in the sample from 477 to 99 categories. For example, the Home Office offence code 3.01 relates to making threats to kill, code 3.03 relates to assisting an offender by impeding his apprehension or prosecution in a case of murder, and code 3.04 is intentionally encouraging or assisting commission of murder: these codes are then summarised by major offence code 3.00 Threat, Conspiracy or

Incitement to Murder.

Some major offence categories were excluded from the sample where there was very small numbers of offenders or no offenders having committed these types of offences. For example, no offenders in the sample were sentenced for major offence code 6 - endangering railway passengers - or major offence code 7 - endangering life at sea.

As can be seen from Table 5.1, offences such as malicious wounding, obscene publications, misuse of drugs, common and other types of assault, fraud and shoplifting occur more than any other types of offences. Offences such as murder, frauds by agents, trustees, company directors etc., and health and safety at work are much less common.

Table 5.1: Offences in the sample

Major offence code	Type of offence	Frequency	Percent
1.00	Murder	5	< 0.1
2.00	Attempted murder	8	< 0.1
3.00	Threats, Conspiracy or Incitement to murder	112	0.2
4.00	Manslaughter	91	0.2
5.00	Wounding and other acts endangering life	384	0.8
8.00	Malicious wounding and other like offences	4,867	9.8
11.00	Cruelty to or neglect of children	263	0.5
13.00	Abduction of a child	18	< 0.1
16.00	Buggery or attempted buggery[*]	9	< 0.1
17.00	Indecent assault on a male	214	0.4
18.00	Indecency between males[*]	20	< 0.1
19.00	Rape of male or female	311	0.6
20.00	Indecent assault on a female	1,166	2.4
21.00	Unlawful sexual intercourse with girl under 13	137	0.3

Table 5.1-*continued from previous page*

Major offence code	Type of offence	Frequency	Percent
22.00	Unlawful sexual intercourse with girl under 16	438	0.9
23.00	Incest with girl under 13 years old	115	0.2
24.00	Procuration of males and females	20	< 0.1
28.00	Burglary in a dwelling	1,175	2.4
29.00	Aggravated burglary in a dwelling (including attempts)	39	0.1
30.00	Burglary other than a dwelling	608	1.2
33.00	Going equipped for stealing	120	0.2
34.00	Robbery and assaults with intent to rob	709	1.4
35.00	Blackmail	36	0.1
36.00	Kidnapping	109	0.2
37.00	Aggravated taking of a vehicle	98	0.2
38.00	Concealing	319	0.6
39.00	Stealing from another person	315	0.6
40.00	Stealing in a dwelling other than from automatic machines and meters	112	0.1
41.00	Stealing by an employee	464	0.9

Table 5.1-*continued from previous page*

Major offence code	Type of offence	Frequency	Percent
42.00	Unlawfully taking away or opening a mail bag	28	0.1
43.00	Abstracting electricity	46	0.1
44.00	Stealing pedal cycles	29	0.1
45.00	Stealing from vehicles	305	0.6
46.00	Stealing from shops and stalls (shoplifting)	2,581	5.2
47.00	Stealing from automatic machines	30	0.1
48.00	Theft of motor vehicle	93	0.2
49.00	Other stealing and unauthorised takings	930	1.9
51.00	Frauds by agents, trustees, company directors etc.	8	< 0.1
52.00	False accounting	101	0.2
53.00	Other frauds	2,903	5.9
54.00	Receiving/handling stolen goods	562	1.1
55.00	Bankruptcy	36	0.1
56.00	Arson	153	0.3
57.00	Other criminal damage endangering life	10	< 0.1

Table 5.1-*continued from previous page*

Major offence code	Type of offence	Frequency	Percent
58.00	Other criminal damage	384	0.8
59.00	Threat or possession with intent to commit criminal damage	58	0.1
60.00	Forgery of prescription	10	< 0.1
61.00	Other Forgery	447	0.9
65.00	Public Order Act 1986 Sec 2 Violent Disorder	45	0.1
66.00	Other offences (against the State and Public Order)	1,413	2.8
67.00	Perjury	16	< 0.1
69.00	Intending to facilitate prison escape	12	< 0.1
71.00	Child prostitution	11	< 0.1
73.00	Abuse of trust - Sexual offences	31	0.1
74.00	Gross indecency with a child	217	0.4
78.00	Immigration Act including people trafficking	30	0.1
79.00	Attempting to pervert the course of justice	256	0.5
80.00	Absconding from lawful custody	20	< 0.1
81.00	Firearms offences	315	0.6

Table 5.1-*continued from previous page*

Major offence code	Type of offence	Frequency	Percent
82.00	Offences against laws relating to Customs, Excise and Inland Revenue	34	0.1
83.00	Bail offences	1,419	2.9
84.00	Trade Descriptions and similar offences	256	0.5
85.00	Health and Safety at work	6	< 0.1
86.00	Obscene publications	4,274	8.6
88.00	Sexual Offences Act 2003 including grooming and other sexual acts	263	0.5
92.00	Misuse of Drugs Act	4,357	8.8
93.00	Permitting premises to be uses for unlawful (drug-related) purposes	70	0.1
99.00	Other indictable offences	121	0.2
104.00	Assault on a constable	1,183	2.4
105.00	Common and other types of assault	6,416	12.9
108.00	Cruelty to animals, ill treatment, neglect etc.	170	0.3
111.00	Offences in relation to dogs Dangerous Dogs Act 1991	10	< 0.1
112.00	Preventing child from receiving information	36	0.1
115.00	Firearms Act 1968	30	0.1

Table 5.1-*continued from previous page*

Major offence code	Type of offence	Frequency	Percent
125.00	Offences against Public Order	976	2.00
126.00	Interference with a motor vehicle	72	0.1
130.00	Stealing and unauthorised takings	174	0.4
131.00	Aggravated vehicle taking	98	0.2
141.00	Drunk and disorderly	103	0.2
143.00	Licensing Act offences	5	< 0.1
149.00	Criminal Damage	1,899	3.8
151.00	Benefit offences	195	0.4
162.00	Offences relating to Police regulations	5	< 0.1
164.00	Public health - By-laws	7	< 0.1
169.00	Transport offences	7	< 0.1
185.00	Picklock related offences	10	< 0.1
193.00	Misuse of drug related offences	77	0.2
195.00	Other offences excluding motoring offences	755	1.5
196.00	Improper use of public electronic communications network	127	0.3

Continued on next page

Table 5.1-*continued from previous page*

Major offence code	Type of offence	Frequency	Percent
802.00	Dangerous Driving	431	0.9
803.00	Driving or attempting to drive after consuming alcohol or drugs	1,131	2.3
804.00	Careless driving	44	0.1
805.00	Accident offence	124	0.3
807.00	Driving licence offences	925	1.9
809.00	Vehicle insurance offences	365	0.7
813.00	Vehicle Test Offences	9	< 0.1
814.00	Fraud, Forgery associated with vehicle or driving records	10	< 0.1
816.00	Speed limit offences	8	< 0.1
818.00	Road traffic offence, neglect, directions etc.	19	< 0.1
	Total	49,590	100.00

[*]These offences have now been re-classified as male rape

5.4.2 Sentence Categories

This research uses only the primary disposal. There are six different types of disposal, which are displayed in Table 5.2, along with the number of cases and percentage of the sample receiving each type of disposal.

Table 5.2: Disposals in the sample

Type of disposal	Frequency	Percent
Immediate custody	19,844	40.0
Suspended sentence	10,252	20.7
Community order	15,059	30.4
Conditional discharge	938	1.9
Fine	1,133	2.3
Other	2,364	4.8
Total	49,590	100.00

Sentencing statistics produced by the Ministry of Justice (2010) in 2009 reported that only 7% of offences resulted in immediate custody, 3% were suspended sentences, 14% were community sentences, 67% were fines and 8% were 'other' sentences. As can be seen from Table 5.2, these figures are very different to those reported by the MOJ and is the result of using data obtained from pre-sentence reports for reasons explained previously in Section 5.3.1. Therefore the sentences used in this research are slightly more severe in nature than would be observed in the general offending population. That being said, the data does still include non-custodial sentences which have previously been omitted from any empirical research in this area.

From the six disposal categories, sentences were then arranged into 35 distinct categories, by choosing sensible cut points based on the frequencies in each category. Table 5.3 displays the sentence categories in the sample.

Offences resulting in a custodial sentence of less than one week are generally for offences relating to breaching bail conditions and shoplifting. Fines are generally given for offences such as being drunk and disorderly, drug offences and public order offences. Other sentences will generally include compensation orders and confiscation

Table 5.3: Sentences in the sample

Disposal	Sentence	Frequency	Percent
Immediate custody	≤ 1 week	230	0.5
	> 1 week ≤ 2 weeks	313	0.6
	> 2 week ≤ 1 month	1,242	2.5
	> 1 month ≤ 2 months	1,361	2.7
	> 2 months ≤ 3 months	1,578	3.2
	> 3 months ≤ 4 months	1,387	2.8
	> 4 months ≤ 6 months	1,889	3.8
	> 6 months ≤ 9 months	1,490	3.0
	> 9 months ≤ 12 months	2,088	4.2
	> 12 months ≤ 18 months	2,137	4.3
	> 18 months ≤ 24 months	1,854	3.7
	> 2 years ≤ 3 years	1,800	3.6
	> 3 years ≤ 4 years	1,089	2.2
	> 4 years ≤ 5 years	530	1.1
	> 5 years ≤ 6 years	318	0.6
	> 6 years ≤ 7 years	153	0.3
	> 7 years ≤ 8 years	133	0.3
	> 8 years ≤ 10 years	86	0.2
	> 10 years including life	166	0.3
Suspended sentence	≤ 1 month	788	1.6
	> 1 month ≤ 2 months	1,121	2.3
	> 2 months ≤ 3 months	1,602	3.2
	> 3 months ≤ 4 months	1,625	3.3
	> 4 months ≤ 5 months	469	0.9
	> 5 months ≤ 6 months	897	1.8
	> 6 months ≤ 9 months	1,913	3.9
	> 9 months ≤ 12 months	1,837	3.7
Community order	≤ 12 months	12,286	24.8
	> 12 months ≤ 2 years	1,665	3.4
	> 2 years ≤ 3 years	1,108	2.2
Fine	Fine	1,133	2.3
Conditional discharge	≤ 12 months	608	1.2
	> 12 months ≤ 2 years	286	0.6
	> 2 years ≤ 3 years	44	0.1
Other	Other	2,364	4.8
	Total	49,590	100.00

orders, and these are usually given for offences such as common assault and motor vehicle offences.

5.5   Concluding remarks

The data for this research comes from OASys data which is then merged with PNC data following sentencing at the Crown or magistrates' court. Due to the nature of OASys data, this sample contains more serious offences/severe sentences than might be expected. However, this is no different to the data used by Pina-Sánchez and Linacre (2013) for example, who use Crown Court survey data.

Unlike the other studies that have used real sentencing data, this research does not use only the principle offence. Instead this research includes all offences and their corresponding sentences to more accurately depict sentencing practice. It is in-fact these two variables which are of initial interest and it is the offences and sentences which will be used to create the severity scale. The following chapter will explain the methodology employed to create this scale.

Chapter 6

Measuring sentence severity using log-linear models

6.1 Introduction

The way in which sentence severity is measured is of particular importance, as the way in which this is done may affect the findings (Sweeten, 2012). The first aim of this empirical work is to create a scale measuring sentence severity which includes the full spectrum of sentences meted out by the courts in England and Wales. As discussed in Chapter 4, there are a variety of methods that have been utilised in the past to create the measure of severity, each with their own merits and shortcomings. Fundamentally, these past measures introduce bias into the sample, and therefore a new, more robust and accurate measure is required.

This work bridges the two approaches of measuring severity: it uses real sentencing data, as opposed to using hypothetical scenarios, and constructs a scale using statistical modelling rather than psychometric scaling. The scale will then have two uses. Firstly, by producing a scale which incorporates both custodial and non-custodial sentences, this will enable use to identify sentences which are of equivalent severity.

The second aim is then to use this scale as the dependent variable within a multilevel model, and a number of legal and extra-legal characteristics will be tested to examine their effects on sentence severity to access disparity in sentencing.

In this chapter the aim is to use a statistical modelling approach to model the offence by sentence categories in a two-way contingency table. I start by considering the Goodman RC association model, which has been used by many authors in which to build scores for the levels of the categorical variables making up a two-way table (e.g. Clogg, 1982). This model provides a set of scores for the row variable (which we will take to be 'offence') and the column variable (taken to be 'sentence'). Such scores are estimated by modelling the association between the offence and sentence categories. I will explain the steps which lead to the Goodman RC model by first describing two simpler models - the independence model and the saturated model. I then discuss how it is possible to extend this model to take account of a number of legal variables which we know impact on the sentences which offenders receive. Controlling for the effects of these legal variables will provide better estimates of the sentence severity score.

6.2   Contingency tables

A two-way contingency table is a statistical tool for summarising and displaying results for categorical variables, where each variable corresponds to one dimension of the table. I can thus organise the offence and sentence data into a O by S contingency table, where the offences are the rows and sentences are the columns. There are $O = 99$ offence categories (rows) and $S = 35$ sentence categories (columns), resulting in a 99×35 table with 3465 cells.

This part of the research seeks to model the association between offences and sentences to create a scale which measures sentence severity. Firstly, I need to establish

this association statistically. Of course, in this application, we would expect a strong association between the offence and sentence variables

6.2.1 The classical test for independence: the Pearson Chi-squared test

If the offence and sentence categories are independent they are mutually exclusive, and therefore, there is no association. I can test this by conducting a Pearson's Chi-Square test. This test is well-known and is described in many introductory statistical textbooks (e.g. Fielding and Gilbert, 2006, p. 270). Testing the null hypothesis of no association gives a value of X^2 of 82408.59 - this is compared to a chi-squared distribution with $(O - 1) \times (S - 1)$, or 3332 degrees of freedom. This is highly significant with a p-value of p<0.0001, suggesting that there is strong association between offence and sentence categories as would be expected.

6.3 Log-linear models

An alternative method of testing for independence is to take a log-linear modelling approach. Let C_{os} be the observed cells of the contingency table. I assume that the cell counts are Poisson distributed with means μ_{os}

$$C_{os} \sim Poisson(\mu_{os})$$

Then the log-linear model of independence can be written as

$$\log(\mu_{os}) = \theta + \alpha_o + \beta_s$$

For identifiability, it is usual to define $\alpha_1 = \beta_1 = 0$, setting the first level of each variable to be the reference category. An equivalent way of writing the model is to

reparameterise by removing the θ parameter,

$$\log(\mu_{os}) = \alpha_o + \beta_s$$

where the α parameters have absorbed the θ parameter.

The deviance, (or minus twice the log-likelihood ratio of the model compared to the saturated model) of this model has a chi-squared distribution with $(O-1) \times (S-1)$ degrees of freedom if there no evidence of association in the table. When applying this model to the data, I obtain a deviance of 50,360.41 on 3332 degrees of freedom, which is again highly significant (p<0.0001)

A second model is the saturated model. This models the association between the row and column variables by adding in a set of interaction parameters, one for each cell. This model totally reproduces the observed cell counts and has a deviance of 0 on 0 degrees of freedom.

$$\log(\mu_{os}) = \alpha_o + \beta_s + \gamma_{os}$$

The Goodman model and other association models aim to find a model which explains the interaction structure between offence and sentence with a relatively small number of parameters - a model between the independence model and the saturated model.

6.3.1 Linear by linear and score association models

The simplest way of modelling an association in a two way table is to give each row and column a fixed score. For example, if we knew a seriousness score for offences, and a severity score for sentences, then these scores could be used directly in the model. If these scores are called $score_o$ and $score_s$ then the product of these scores can be formed, and make an additional explanatory variable in the model. The score

association model is then:

$$\log(\mu_{os}) = \alpha_o + \beta_s + \lambda.score_s \times score_o$$

A special form of this model is the linear by linear score association model, where the scores for the offences are simply the number $1, 2, \ldots, O$ for each row, and scores for the sentences are similarly $1, 2, \ldots, S$ for each column. This model really assumes that the categories are in some natural order, and that therefore it makes sense to assign a rank order, For example, this model might be considered if the columns were categories of a likert response and the rows of the table were social class categories.

However, of course, we do not know either a seriousness score for offences or a severity score for sentences, and therefore these need to be estimated from the model. I replace $score_o$ and $score_s$ by γ_o and δ_s with the Greek letter notation indicating that the scores now need to be estimated from the model. This motivates the Goodman RC association model.

6.4 Goodman RC association model

The simplest form of the model can be written as

$$\log(\mu_{os}) = \alpha_o + \beta_s + \lambda\gamma_o\delta_s$$

This model is known as a log-multiplicative model, and is also known as the Goodman RC(1) model. The (1) means that there is just one multiplicative term in the model. The λ parameter is just a scaling parameter, and can be absorbed into either the γ_o or the δ_s parameters. This gives the model

$$\log(\mu_{os}) = \alpha_o + \beta_s + \gamma_o\delta_s$$

The table displayed in Table 6.1 represents a simplified contingency table with three levels of offence seriousness - high, medium and low - and three levels of sentence severity - high, medium and low. If the offence categories were ranked in order of seriousness and the sentence categories were ranked by seriousness, then we would expect to see a high number of counts in the diagonal cells. This is because, for example, offences with high levels of seriousness would receive sentences with high levels of severity. Similarly low level offences would receive low severity sentences etc. I am thus interested in the association (represented by high counts) between the offence and corresponding sentence categories which frequent the diagonal cells in the table.

Table 6.1: Association within a contingency table

		Sentence		
		Low	Medium	High
Offence	High	Low counts	Low counts	*High counts*
	Medium	Low counts	*High counts*	Low counts
	Low	*High counts*	Low counts	Low counts

In section 6.3.1, I explained that we do not know either the rank order of offences by seriousness or the order of sentences by their severity, and therefore this ruled out using the score association model. Consequently, the Goodman RC association model does not assume the correct ordering of categories: instead the model reveals the ordering of the categories through estimating the two unknown parameters (Powers and Xie, 2000) γ_o and δ_s. This requires an iterative procedure at the model building stage and this requires starting with a random set of starting values, whereby one set of parameters are treated as being known to update the other set of estimates until they both stabilise. This presents us with a series of scores (ibid), which then becomes the latent variable/sentence severity scale.

6.5 Extending the Goodman RC association model

While the simple Goodman RC model can be used to estimate severity scores, the estimates will not be optimal. In Chapter 2 it was explained that judges are required to consider a number of aggravating and mitigating factors before reaching a final decision. I therefore want to control for some of these variables when estimating the model scores.

In Section 2.3.1 I explained the sentencing guidelines in practice as well as the use of aggravating and mitigating factors which judges must consider when passing a sentence. This includes a number of statutory factors such as, previous convictions, whether the offence was committed when the offender was on bail, whether the offence was aggravated in relation to race, religion, disability or sexual orientation of the victim, whether the offence involved a terrorist connection, and also the plea, which is considered a mitigating factor if the offender pleads guilty to the offence. Therefore, within the extended Goodman RC model, I wanted to be able to control for these statutory types of covariates.

Some of these factors were recorded in the OASys data such as the offence plea, whereas others such as whether the offence was committed when the offender was on bail and whether the offence was aggravated in relation to race, religion, disability or sexual orientation of the victim were recorded but contained a lot of missing cases which raised questions about their validity. There was no information recorded on the previous convictions of the offenders. As a result of this, the majority of these factors could not be used in the subsequent analysis.

Two additional variables were created for the subsequent analysis (see Section 5.3.1): the first was an indicator of prior offending and the second identified cases where the offender was sentenced for multiple offences. The first variable used information relating to whether the offender had previously appeared in court to indicate prior offending which would act as a proxy for previous convictions. The second vari-

able indicated whether the offender was sentenced for multiple offences to reflect the totality principle. It would therefore be expected that offenders that are sentenced for multiple offences would be sentenced more severely to reflect the their offending behaviour and ensure the principle of proportionality was adhered to.

We can think of each offence in the table as having a distribution of sentences over the sentence categories, but that distribution will be affected by the combination of legal variable, such as plea, prior convictions etc. An offence which is a debut offence for offenders will attract lower sentences compared to another offence of equal seriousness that is committed later in a criminal career. I therefore develop the model to include such effects. It would have been possible to include more legal variables (aggravating and mitigating factors) in the model but this would have made the model overly complicated.

In this research, I consider three categorised legal variables: number of offences ($\leq 2 / > 3$), prior court appearance (prior/no prior) and plea (guilty/not guilty/unknown). There are 12 possible combinations of levels of these three variables (e.g. ≤ 2 offences-prior-guilty is one combination) and each combination now has its own offence by sentence table. An example of the raw data prior to creating the table can be seen in Appendix C and a sample of the code used to create the table and the extended Goodman RC model can be seen in Appendix F.

Conceptually, the whole data now consists of a set of 12 separate *layers* of 35 by 99 tables, as illustrated in Figure 6.1 - or a three-way table of 12 x 35 x 99 with 41,580 cells in total. By creating additional cells, the data in the contingency table would be too sparse.
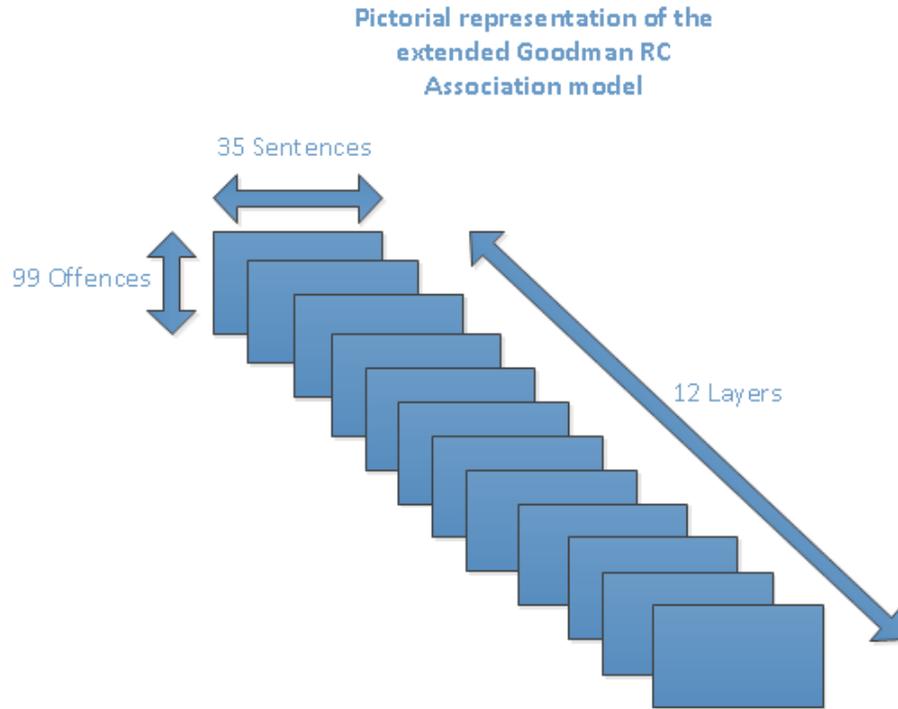
Figure 6.1: Diagram representing the Extended Goodman RC Association model

6.5.1 Existing extensions of two-way contingency table models with layers

Other authors have extended the Goodman RC association model into three dimensions to model data in the form row x column x layers. Probably the best known extension is by Xie (1992) which is known as the log-multiplicative layer effect model. This fits a full interaction between layers and rows, and between layers and columns, and assumes a common association between the rows and the columns of the table for each layer, but multiplied by a different constant for each layer. Such models have been used by Erikson and Goldthorpe (1992) to model social mobility tables and how they change over time. This model is inappropriate for this work as it does not decompose the row column interaction term into row and column scores.

A different approach was suggested by Goodman and Hout (1982), who proposed a 'modified regression-type approach' for modelling the association between rows and

columns in the presence of layers. Essentially, this approach decomposes the interaction into a common interaction term present for all layers, and a second common term which is multiplied by a layer-specific constant. This method has been used to examine how the association between job satisfaction and self-employment varies across occupational groups. This model is again inappropriate for my purposes, as again, there are no row and columns scores estimated in the model.

In addition, neither of these approaches allow for covariates to be placed on the estimated association parameters. I therefore proceed to develop a new model.

6.5.2 A new model for incorporating sentence covariates into the Goodman model

None of the previous work described above supplies a suitable form of model to allow the sentence scores δ_s to be themselves modelled by legal covariates. I therefore am proposing a new extension to the Goodman model, which is described below.

Let $C_{os\ell}$ be the three way table of counts with $C_{os\ell} \sim Poisson(\mu_{os\ell})$

Then I propose

$$\log(\mu_{os\ell}) = \alpha_o + \beta_s + \tau_\ell + \gamma_o \delta_{s\ell}$$

where

$$\delta_{s\ell} = \delta_s + \sum_p \theta_p X_{ps}$$

The δ_s in the standard Goodman model is replaced by $\delta_{s\ell}$ with the sentence scores now additionally depending on the layer. The $\delta_{s\ell}$ is decomposed into δ_s and the additive effects of the covariates. The X_{ps} will represent dummy variables for plea, prior court appearance, and number of offences and θ_p are a set of unknown regression coefficients. The adjusted sentence severity scores are given by the δ_s which will

now control for the legal variables. I can determine which of the legal variables are significant by fitting a series of extended Goodman models, each with a different subset of covariates.

However, only one scale or set of scores is actually produced. The final sentence severity scale or set of scores correct for the effects of the legal factors. Aggravating factors will increase the sentence severity, whereas the mitigating factors will decrease the sentence severity. Therefore by controlling for the legal covariates (aggravating and mitigating factors) we are able to control for the imbalance these factors have on sentence severity. These scores are then standardised using linear interpolation which is explained in Section 6.8. Custodial sentence days are anchored and therefore linear interpolation standardises the non-custodial scores in relation to these days in custody.

In the second part of this PhD research, the new scale is used as the dependent variable in a multilevel model and the same legal variables (plea, prior and number of offences) are included as independent/explanatory variables. This allows us to assess the impact that these covariates have on sentence severity. For example, we will be able to assess if there is a reduction in sentence severity where an offender pleads guilty compared to an offender that pleads not guilty, controlling for a number of other legal and extra-legal factors such as, offence type, sex, ethnicity and so on.

6.6   Software

For this research, the R software environment was used to fit the Goodman RC association model and the extended Goodman RC models. An example of the layout of the dataset prior to fitting the Goodman RC and extended Goodman RC models can be see in Appendix C.

Due to the large amount of data and the memory capacity of my own machine,

the model fitting was undertaken using Lancaster University's High End Computing Cluster. The High End Computing Cluster is a service that supports researchers and research students who require high performance computing. The service supports the use of a number of third-party software, including R.

The 'gnm' package by Turner and Firth (2012), provides the facilities for fitting generalised non-linear models, including both the Goodman RC model and the extended Goodman RC model. The gnm model is fitted using the Poisson family log link function. The 'Mult' function forms part of the model formula and provides the log-multiplicative term in the model formula. Thus the following command:

$$RC = gnm(Freq \sim SENT + OFF + Mult(SENT, OFF), family = Poisson)$$

fits the basic Goodman RC model, and

$$ExtRC = gnm(Freq \sim SENT + OFF + Mult((SENT + Plea + Numoff + Prior), OFF) + Plea + Numoff + Prior, family = Poisson)$$

fits the most complex extended Goodman RC model. See Appendix F for the sample code and commentary used to create the extended Goodman RC model.

6.7 Model fitting

A number of models were fitted and these can be seen in Table 6.2 along with the deviance (minus twice the log-likelihood ratio of the model compared to the saturated model) for that model, the difference in deviance between the nested model, the difference in degrees of freedom between the two models, and finally the associated p-value. A p-value of less than 0.001 suggested that the more complicated model (the model with the additional parameters) provided the best model fit for the data.

Table 6.2 shows the first model to be fitted was the basic Goodman RC model with no legal covariates (Model 1). Maximum likelihood estimation was used to assess model fit and provide the model deviance. The legal covariates were then added to the standard Goodman RC model - one variable at a time and then combinations of the variables were added - to produce the extended Goodman RC model. To test whether the additional covariates improve model fit, it is standard practice to use the reduction in deviance and the degrees of freedom for each of the nested models. This in turn provides a p-value which suggests whether there is sufficient evidence to reject or not reject the additional parameters in the model.

In total, seven models were fitted: the basic Goodman RC model and six extended Goodman RC models. The models with only one legal covariate (Models 2, 3 and 4) were compared to the basic Goodman RC model (Model 1). The models with two legal covariates (Models 5 and 6) were then compared to Model 2 with just plea as a covariate. The final model with all three legal covariates (Model 7) was then compared to model 5. As can be seen from table 6.2 the model with the lowest deviance score and biggest difference in deviance on 2 degrees of freedom was Model 7, which contained all three legal variables. Model 7 is therefore the optimal model. The results of this analysis will be discussed in detail in Chapter 7.

6.8   Linear interpolation

The extended Goodman RC analysis then provides us with a series of 35 sentence severity scores. Although the scores (sentence severity scale) were estimated using sophisticated statistical techniques, and provide us with a measure of sentence severity in their own right, the scores do not have any real meaning and therefore appear arbitrary. For example, a score of -5.17 (immediate custody >2 weeks ≤1 month) is less severe than a score of -4.38 (immediate custody >1 month ≤2 months), and a

Table 6.2: Models fitted to obtain the optimal model

Model	Covariates	Deviance	Difference in deviance	Difference in degrees of freedom	p-value
1	None: Basic Goodman RC	242,468			
2	RC, Plea	241,664	804	2	<0.0001
3	RC, Number of offences	241,410	1,058	1	<0.0001
4	RC, Prior court appearance	239,739	2,745	1	<0.0001
5	RC, Plea, Number of offences	240,601	1,063	2	<0.0001
6	RC, Plea, Prior court appearance	238,866	2,798	2	<0.0001
7	**RC, Plea, Number of offences, Prior court appearance**	**237,437**	**3,164**	**2**	<0.0001

Note: all models are fitted to the full 99×35×12 table

score of -4.48 (conditional discharge $\leq$12 months) is somewhere in-between. However, the scale has no anchor points.

However, I also know the modal[20] number of days in custody for each of the custodial sentence categories. For example, *immediate custody >2 weeks ≤1 month* is approximately 28 days and *immediate custody >1 month ≤2 months* is approximately 54 days. It would then be useful if it were possible to anchor the extended Goodman RC scores to estimate the equivalent number of days in custody for the non-custodial sentences.

Having said that, it is possible to use a technique called Linear Interpolation, which estimates a point (e.g. a non-custodial sentence category) between two known points (e.g. two custodial sentence categories). The custodial days for each of the custodial sentence categories act as an anchor to then standardise the non-custodial sentences in relation to the known days. As I know the modal number of days the custodial sentences equate to in each category, I can use these figures to estimate the equivalent days in custody (Y) for the non-custodial sentence categories using the following formula,

$$Y = Y_0 + (Y_1 - Y_0)\left(\frac{X - X_0}{X_1 - X_0}\right)$$

In this equation Y_1 (immediate custody >1 month ≤2 months) and Y_0 (immediate custody >2 weeks ≤1 month) relate to the modal number of days in custody for each of the two custodial sentences. The X's relate to the raw scores obtained from the extended Goodman RC Association analysis (see Table 7.2), where X is the score for the non-custodial sentence (conditional discharge ≤12 months), and X_1 (immediate custody >1 month ≤2 months) and X_0 (immediate custody >2 weeks ≤1 month)

---

[20]The modal or most common number of days in custody was used instead of the average number of days in custody so that the categories were not effected by any outliers (especially for the immediate custody category of 10 years and over including life) although there was not much difference in the number of days for each sentence category when the average was used. Alternatively the average number of days in custody could have been used.

are the raw scores for the two custodial sentences. We can therefore estimate the equivalent number of days in custody for the non-custodial sentence Y (conditional discharge ≤12 months) by inserting the appropriate numbers into the formula. The diagram in Figure 6.2 illustrates the figures that are entered into the formula to produce the equivalent days in custody for the example given above. This calculation was done by hand for each of the non-custodial sentence categories.
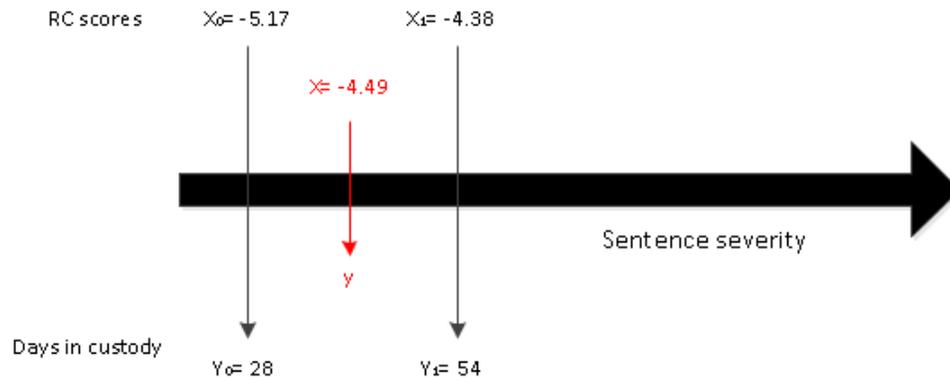
Figure 6.2: Diagram representing linear interpolation

The results of the linear interpolation are provided in Chapter 7. It is important to emphasis that linear interpolation only yields an approximate measure. However, it provides a series of sentence severity scores (or sentence severity scale) which can be interpreted in a meaningful way. These sentence severity scores - equivalent days in custody - then provide us with a new and improved measure of sentence severity, which can then be used to investigate sentencing disparity. It also provides us with a scale which allows us to assess the interchangeability of sentences, which Sebba (1978), von Hirsch (1993), and Morris and Tonry (1990) view as of primary importance.

6.9   How can the sentence severity scale be used?

In Section 6.1 I explained that the aim of creating this type of sentence severity scale (extended Goodman RC model) was two-fold. Firstly, by incorporating the

non-custodial and custodial sentences into a single scale, it is possible to identify sentences that overlap or are equivalent in terms of their level of sentence severity (see Section 7.3.2). This could potentially be an aid to judges by identifying alternative non-custodial sentences which are equivalent in terms of their severity to custodial sentences, and thus encourage/facilitate the use of non-custodial sentences. Secondly, the scale is a tool in which we can measure sentencing disparity across the spectrum of different disposals, replacing the previous methods discussed in Sections 4.3.1 and 4.3.2.

This research uses the new sentence severity scale to investigate sentencing disparity. In this instance, the sentence severity scale becomes the dependent variable in the multilevel model (see Section 8.3). A number of legal (for example; offence plea and type of court the offender was sentenced in) and extra-legal factors (For example, the offender's sex and ethnicity) are then fitted as independent variables (see Section 8.4) in the model to try and explain the variability in sentence severity scores.

6.10   Concluding remarks

In this chapter, I have shown that it is possible to use Goodman RC association analysis to model the association between two categorical variables - offence and sentence. Furthermore, it was possible to extend this model and control for legal variables, which sentencing legislation states, should impact on sentence severity.

Linear interpolation then allowed us to estimate the equivalent number of days in custody that the non-custodial sentence categories equate to using the extended Goodman RC association estimates for the custodial sentences. The outcome of this is a scale, which is meaningful, and also facilitates the investigation into sentencing disparity.

The following chapter provides the results of the Goodman RC association anal-

ysis, extended Goodman RC analysis and the linear interpolation.

Chapter 7

Results of log-linear modelling

7.1 Introduction

To recall, the main focus of this doctoral research is to assess disparity in the sentencing of offenders in the magistrates' and Crown Courts of England and Wales. In order to do this effectively, it is essential to use a measure of sentence severity that does not introduce bias in the sample. In Chapter 4, I discussed the methods typically used to measure sentence severity; the 'in-out' decision to incarcerate, length of custodial sentences and scales based on the perception of severity for a range of sentences. Additionally, I expressed my concern regarding the robustness and objectivity of these methods. This then prompted me to find an alternative, more effective and accurate way to measure sentence severity: one that incorporates the full range of sentences meted out by the courts in England and Wales, and models real sentencing data. Goodman RC association analysis facilitated this.

Goodman RC association analysis models the association between the offence and sentence categories, incorporating the full range of sentences issued by the courts. It then estimates a series of scores, which serves as a scale, measuring sentence severity. Therefore, the full range of sentences are aligned on a single scale. We can then

extend this model by adding a range of legal covariates. By controlling for the legal factors (offence plea, number of offences and prior offending), the aim is to more closely model/replicate and explain what is happening in the courts and to get a better understanding of the 'real' situation in sentencing.

Chapter 6 discussed the Goodman RC association and extended Goodman RC association models, and how these are used to estimate sentence severity scores. I then explained how linear interpolation was used to transform the scores (δ_s) from the extended Goodman RC analysis into the equivalent number days in immediate custody.

This chapter (Chapter 7) will now briefly discuss the results from the Goodman RC association and extended Goodman RC analysis, as well as, the results from the linear interpolation. This then provides a scale of sentence severity that is meaningful in its own right, and is then used to assess sentencing disparity.

7.2 Basic Goodman RC model - no legal covariates

Basic Goodman RC association analysis models the association between the offence and sentence categories, with no additional covariates. The results from this analysis are displayed in Table 7.1, where the first column in the table relates to the sentence categories, and the second column displays the raw scores from the Goodman RC analysis.

7.2.1 Sentence severity scale - alignment of sentences

The results displayed in Table 7.1 clearly demonstrate that the range or distribution of sentence categories can be arranged on a single scale modelling one dimension of sentence severity. The scale ranges from the least severe sentence - Immediate custody ≤ 1 week - with a raw score of -5.616, up to the most severe sentence - Immediate

Table 7.1: Raw scores from the basic Goodman RC model ranked in order from least to most severe

Sentence category	Raw score (δ_s)
Immediate custody ≤1 week	−5.616
Immediate custody >1 week ≤ 2 weeks	−5.432
Fine	−5.147
Suspended sentence ≤1 month	−4.915
Immediate custody >2 week ≤1 month	−4.860
Conditional discharge ≤12 months	−4.469
Immediate custody >1 month ≤2 months	−4.305
Conditional discharge >12 months ≤2 years	−4.255
Conditional discharge >2 years ≤3 years	−4.205
Community order ≤12 months	−4.139
Suspended sentence >1 month ≤2 months	−4.128
Other	−4.062
Community order >12 months ≤2 years	−3.958
Suspended sentence >2 months ≤3 months	−3.815
Immediate custody >2 months ≤3 months	−3.692
Suspended sentence >3 months ≤4 months	−3.486
Immediate custody >3 months ≤4 months	−3.338
Suspended sentence >4 months ≤5 months	−3.181
Suspended sentence >5 months ≤6 months	−1.461
Immediate custody >4 months ≤6 months	−1.386
Suspended sentence >6 months ≤9 months	0.416
Suspended sentence >9 months ≤12 months	0.831
Immediate custody >6 months ≤9 months	0.966
Immediate custody >9 months ≤12 months	2.325
Immediate custody >12 months ≤18 months	3.342
Community order >2 years ≤3 years	4.013
Immediate custody >18 months ≤24 months	5.321
Immediate custody >2 years ≤3 years	6.883
Immediate custody >3 years ≤4 years	8.723
Immediate custody >4 years ≤5 years	10.452
Immediate custody >5 years ≤6 years	11.258
Immediate custody >6 years ≤7 years	11.812
Immediate custody >7 years ≤8 years	12.235
Immediate custody >8 years ≤10 years	14.411
Immediate custody >10 years including life	17.098

custody >10 years including life - with a raw score of 17.098, incorporating and aligning the full range of sentences, each with their different punitive elements.

The raw scores provide the magnitude of severity for each sentence, which enables

the categories to be ranked in order of sentence severity. To inspect whether the basic Goodman model has achieved this, we can consider the raw scores for just the immediate custody categories. Now it is generally accepted that as the length of a custodial sentence increases, so too does the severity of that sentence. The estimates for the immediate custody categories displayed in Table 7.1, corroborate with this notion. Therefore, we can then be fairly confident that the other (non-custodial) sentence category estimates are in the correct rank order.

We would perhaps not expect to see the immediate custody categories appearing so low on the scale. However, these sentences (e.g. immediate custody ≤ 1 week the least severe sentence) are generally given for the least serious offences, such as failing to adhere to bail conditions. Additionally, these sentences will generally be awarded to offenders who are also being sentenced for another offence which will also likely result in a term of immediate custody.

Conditional discharges all seem to receive fairly low sentence severity scores. The conditions placed on the discharge would have given us a more accurate measure of the severity of these sentences, however, this information was not available.

The 'fine' category has a relatively low sentence severity score. Although we need to keep in mind that fines are also dependent on the financial situation of the offender and are means tested. This may then have a different effect on sentence severity.

The 'other' category includes sentences such as confiscation orders and compensation orders. This sentence category received a score of -4.062, and this is more severe than expected. Although, again these are usually given with a combination of sentences (e.g. community orders).

Community orders generally fall between the shorter suspended sentence categories on the scale of sentence severity. The community order category relating to greater than two years and less than or equal to three years is estimated as a lot more severe than the other community order categories. A community order of greater than two

years and less than or equal to three years receives a raw sentence severity score of 4.013 which is ranked between immediate custody greater than 12 months and less than or equal to 18 months (δ_s=3.342), and immediate custody greater than 18 months and less than or equal to 24 months (δ_s=5.321). This would seem reasonable as a community order greater than two years and less than or equal to three years would be awarded for a relatively serious offence, which may be on the border of getting a custodial sentence.

It seems reasonable that the range of suspended sentence scores are slightly less (on the scale of sentence severity) than the same range of custodial sentences. The legislation suggests that only once the length of the custodial sentence has been decided should the judge then decide to suspend it. So for example a suspended sentence order of less than one month is shown as slightly less severe than a custodial sentence of between two weeks and one month. This is replicated with the other suspended sentence and immediate custody sentence categories. There is one anomaly, where the suspended sentence category relating to greater than one month and less than or equal to two months receives a raw score of -4.007 which is more severe than the equivalent custodial sentence of similar length (greater then one month and less than or equal to two months) with a raw score of -4.379. This would suggest that this custodial sentence is less severe than this suspended sentence, which we would not expect to be the case. This result could be caused by the inaccurate recording of data when it was uploaded onto the PNC.

The actual raw scores range from -5.616 to 17.008 but they have no real meaning other than they provide the rank order for the sentence categories, based on the association between the offence and sentence categories. This is because, as explained in Section 6.4, these are estimated from a random set of starting values to facilitate the iterative procedure as there is no anchoring to the raw scale. A different set of starting values will give the same model fit - deviance - but different positioning of the

scale. Consequently, these raw scores cannot be directly compared with, for example, the estimates from the extended Goodman RC model.

I will now go on to explain the results of the extended Goodman RC model which includes the legal covariates, which is the optimal and preferred model.

7.3 Extended Goodman RC model - controlling for legal covariates

In section 6.5, I discussed extending the Goodman RC association model to include additional legal covariates (offender's plea, total number of offences they are being sentenced for, and prior court appearance, acting as an indicator of prior offending) to more accurately reflect the seriousness of an offence and in turn the severity of a given sentence. This then controls for the imbalance that legal factors have on sentencing and therefore we are able to provide a far more accurate representation of the sentence severity by controlling for some of the factors that are taken into the sentencing decision-making process.

Controlling for the legal covariates in the model correct for the imbalance that these factors have on sentencing - by either decreasing or increasing the severity of sentences. Furthermore, linear interpolation then standardises these scores relative to the severity of the custodial sentences. The end result is a single sentence severity scale which orders the sentences by their magnitude of sentence severity and allows for the imbalance that legal covariates create.

Recall that in Section 6.5.2 I proposed the Extended Goodman model to be:

$$\log(\mu_{os\ell}) = \alpha_o + \beta_s + \tau_\ell + \gamma_o \delta_{s\ell}$$

where

$$\delta_{s\ell} = \delta_s + \sum_p \theta_p X_{ps}$$

The Extended Goodman RC model therefore produces a number of parameter estimates in addition to the sentence severity scores which are later used to create the sentence severity scale. All the parameter estimates generated by fitting the Extended Goodman model are displayed and explained in the following sections.

7.3.1   Sentence severity scores

The raw sentence scores (δ_s) and final scores (obtained through linear interpolation) from this model are displayed in Table 7.2. The final scores highlighted in bold indicate the equivalent number of days in custody which were calculated using the linear interpolation equation (see Section 6.8) and are therefore estimates of the number of days in custody.

Again the raw scores provide the magnitude of severity, which enables the categories to be ranked in order of sentence severity. These estimates are different to the estimates obtained from the Goodman RC model: firstly because a different set of starting values are used at the model building stage, and secondly because there are now also a number of covariates added to the model. The two models are not nested, and therefore they are not comparable.

Although, the extended Goodman RC model produces a series of sentence severity estimates that allow the sentences to be ranked in order of severity, out-with this scale, the scores do not have any real meaning because they have not been anchored and therefore they appear arbitrary. As this model is now the optimal and preferred model, linear interpolation was used, as explained in section 6.8, to produce a final set of sentence severity scores. This then uses the modal number of days for each of the immediate custody categories to estimate the equivalent (approximate) number

Table 7.2: Scores from the extended Goodman RC model with legal covariates

Sentence category	Raw score (δ_s)	Final score (Days)
Immediate custody >1 week ≤2 weeks	−6.663	14
Immediate custody ≤1 week	−6.585	7
Fine	−5.256	**27**
Suspended sentence ≤1 month	−5.173	**28**
Immediate custody>2 week ≤1 month	−5.168	28
Conditional discharge ≤12 months	−4.478	**50**
Immediate custody>1 month ≤2 months	−4.379	54
Conditional discharge >12 months ≤2 years	−4.115	**63**
Community order ≤12 months	−4.028	**65**
Suspended sentence>1 month ≤2 months	−4.007	**66**
Community order >12 months ≤2 years	−3.881	**72**
Suspended sentence >2 months ≤3 months	−3.611	**79**
Immediate custody >2 months ≤3 months	−3.390	86
Other	−3.371	**88**
Suspended sentence >3 months ≤4 months	−3.265	**97**
Conditional discharge >2 years ≤3 years	−3.205	**103**
Immediate custody >3 months ≤4 months	−3.050	117
Suspended sentence >4 months ≤5 months	−2.717	**126**
Suspended sentence >5 months ≤6 months	−1.352	**163**
Immediate custody >4 months ≤6 months	−1.060	171
Suspended sentence >9 months ≤12 months	−0.286	**216**
Suspended sentence >6 months ≤9 months	−0.264	**217**
Immediate custody >6 months ≤9 months	0.347	252
Immediate custody >9 months ≤12 months	0.761	351
Immediate custody >12 months ≤18 months	1.023	494
Immediate custody >18 months ≤24 months	5.321	705
Immediate custody >2 years ≤3 years	2.008	998
Community order >2 years ≤3 years	2.821	**1335**
Immediate custody >3 years ≤4 years	2.886	1362
Immediate custody >4 years ≤5 years	4.085	1770
Immediate custody >5 years ≤6 years	4.449	2156
Immediate custody >6 years ≤7 years	4.835	2495
Immediate custody >7 years ≤8 years	4.914	2888
Immediate custody >8 years ≤10 years	6.416	3514
Immediate custody >10 years including life	8.008	4816

of days in custody for the non-custodial sentence categories.

The actual rank order of the sentence categories from the extended Goodman RC model are similar to the rank order that we have seen for the Goodman RC model,

albeit with some slight differences. For example, the immediate custody category of less than or equal to one week is estimated as now marginally more severe than immediate custody of greater than a week and less than or equal to two weeks category, which is why the final scores appear out of sequence in Table 7.2 with 14 appearing before 7 in the final score category based on the modal days in custody. There may be many reasons why this might have happened, such as anomalies in the data and the way the offences or sentences were recorded. Another reason for this may relate to the fact those receiving a sentence of less than one week tend to be for bail offences. In this situation, multiple offences will be recorded which will increase the severity of a sentence, and the offender is also more likely to plead not guilty for the bail offence which are now taken into consideration with the extended Goodman model. This will increase the severity of the sentences and hence will appear more severe than the immediate custody sentence of 14 days.

The suspended sentence greater than six months and less than or equal to nine months category ($\delta_s=-0.264$) is estimated as more severe than the nine to twelve months suspended sentence category ($\delta_s=-0.286$). Although the difference is quite small, this again is likely to be the result of anomalies in the data and the way that these types of offences have been recorded in the PNC. As I explained in Section 2.7.4, the severity of suspended sentences is based on the length of the custodial sentence as well as the length of time that is suspended. Therefore by only using the length of the custodial part of the sentence as a category as I have had to do here leaves it open to these discrepancies. More accurate recording of sentencing data would vastly improve these estimates.

Other sentence categories also appear slightly less or more severe in comparison to the other sentences than would perhaps been anticipated. I would have expected the conditional discharge sentences to be much lower on the severity scale than they appear, for example the conditional discharge of between two and three years which

following the linear interpolation equates to over three months in prison. However again only the length of this type of order was recorded in the PNC data and without the finer gradations of the sentence in relation to the conditions attached to the sentence, it is not possible to more accurately model these types of sentence. In a similar vein, I would have expected some of the custodial sentences to have appeared more severe, such as the immediate custody of two to three years. This could be the way these sentences were categorised and again finer gradation of custodial sentences would have prevented this.

Sentence severity scores with a different subset of covariates

The raw scores displayed in Table 7.2 are based on one combination of the legal covariates. Consequently, it would have been possible to choose another layer or combination of legal covariates using the θ_p parameter estimates obtained from the extended Goodman RC analysis. The θ_p in table 7.3 displays the additional coefficients.

Table 7.3: Theta parameters

Coefficient	θ
Plea: Not guilty	1.028
Plea: Unknown	-0.240
Numoff2: >3	0.895
Prior: Prior	-1.306

We can therefore use the θ_p parameters to change the legal characteristics or baseline characteristics by adding the relevant coefficient estimates to the raw sentence severity scores estimated in the Extended Goodman RC model. For example we can calculate scores based on the following characteristics - the offender plead not guilty (1.028), was sentenced for three or more offences (0.895) and had no prior court appearances (no change - same as previous baseline). Table 7.4 shows the new raw scores (δ_s) for each sentence and we can see that each sentence has increased in severity by 1.923 (1.028 + 0.895) indicating more severe sentence severity scores.

However, following linear interpolation (see Section 6.8) to anchor the scores, we can see that there is actually no change in the estimated equivalent days in custody for each of the sentence categories also shown in Table 7.3 and that is because only one scale is actually created. Extending the Goodman RC model allows us to control for the legal factors and therefore control for the imbalance or affects that the legal factor have on sentence severity.

The sentence severity scale now provides a quantitative measure of sentence severity which places sentences in rank order, based on how they are used in practice in England and Wales. By controlling for the legal covariates, the aim was to more accurately model the complex nature of sentencing whilst controlling for the imbalance or effects that legal factors have on sentence severity. Whilst it should be kept in mind that this provides an approximation of sentencing, this scale facilitates the second part of this research which looks into sentencing disparity.

Interchangeability of sentences

The results (final scores) shown in Table 7.2 (and Table 7.4) can also be displayed graphically in the form of a dotplot shown in Figure 7.1. The dots represent the equivalent number of days in custody for each of the sentence severity scores, estimated by linear interpolation (for the non-custodial sentence, as described in section 6.8). The X-axis provides the sentence severity scale and the various sentences are arranged on this scale to represent their equivalent severity score. This scale displays the levels of severity from left to right, with the least severe sentences on the left to the most severe sentences on the right-hand side of the scale. This graphical representation of the data replicates the diagram of the sentence severity scale (Figure 4.1) produced in Chapter 4, where the original idea was conveyed.

The 35 sentence categories are grouped by type of disposal, represented by different coloured dots on the graph. The dark blue dots represent the different community

Table 7.4: Alternative scores from the extended Goodman RC model using different theta parameters

Sentence category	Raw score (δ_s)	Final score (Days)
Immediate custody >1 week ≤2 weeks	−4.740	14
Immediate custody ≤1 week	−4.662	7
Fine	−3.333	**27**
Suspended sentence ≤1 month	−3.250	**28**
Immediate custody>2 week ≤1 month	−3.245	28
Conditional discharge ≤12 months	−2.555	**50**
Immediate custody>1 month ≤2 months	−2.456	54
Conditional discharge >12 months ≤2 years	−2.192	**63**
Community order ≤12 months	−2.105	**65**
Suspended sentence>1 month ≤2 months	−2.084	**66**
Community order >12 months ≤2 years	−1.958	**72**
Suspended sentence >2 months ≤3 months	−1.688	**79**
Immediate custody >2 months ≤3 months	−1.467	86
Other	−1.448	**88**
Suspended sentence >3 months ≤4 months	−1.342	**97**
Conditional discharge >2 years ≤3 years	−1.282	**103**
Immediate custody >3 months ≤4 months	−1.127	117
Suspended sentence >4 months ≤5 months	−0.794	**126**
Suspended sentence >5 months ≤6 months	0.571	**163**
Immediate custody >4 months ≤6 months	0.863	171
Suspended sentence >9 months ≤12 months	1.637	**216**
Suspended sentence >6 months ≤9 months	1.659	**217**
Immediate custody >6 months ≤9 months	2.270	252
Immediate custody >9 months ≤12 months	2.684	351
Immediate custody >12 months ≤18 months	2.946	494
Immediate custody >18 months ≤24 months	7.244	705
Immediate custody >2 years ≤3 years	3.931	998
Community order >2 years ≤3 years	4.744	**1335**
Immediate custody >3 years ≤4 years	4.809	1362
Immediate custody >4 years ≤5 years	6.008	1770
Immediate custody >5 years ≤6 years	6.372	2156
Immediate custody >6 years ≤7 years	6.758	2495
Immediate custody >7 years ≤8 years	6.837	2888
Immediate custody >8 years ≤10 years	8.339	3514
Immediate custody >10 years including life	9.931	4816

order categories, light grey dots represent the conditional discharge categories, orange represents the fine category, immediate custody categories are displayed in red, the green dot represents the other category , and finally the suspended sentence category

estimates are displayed in lilac.
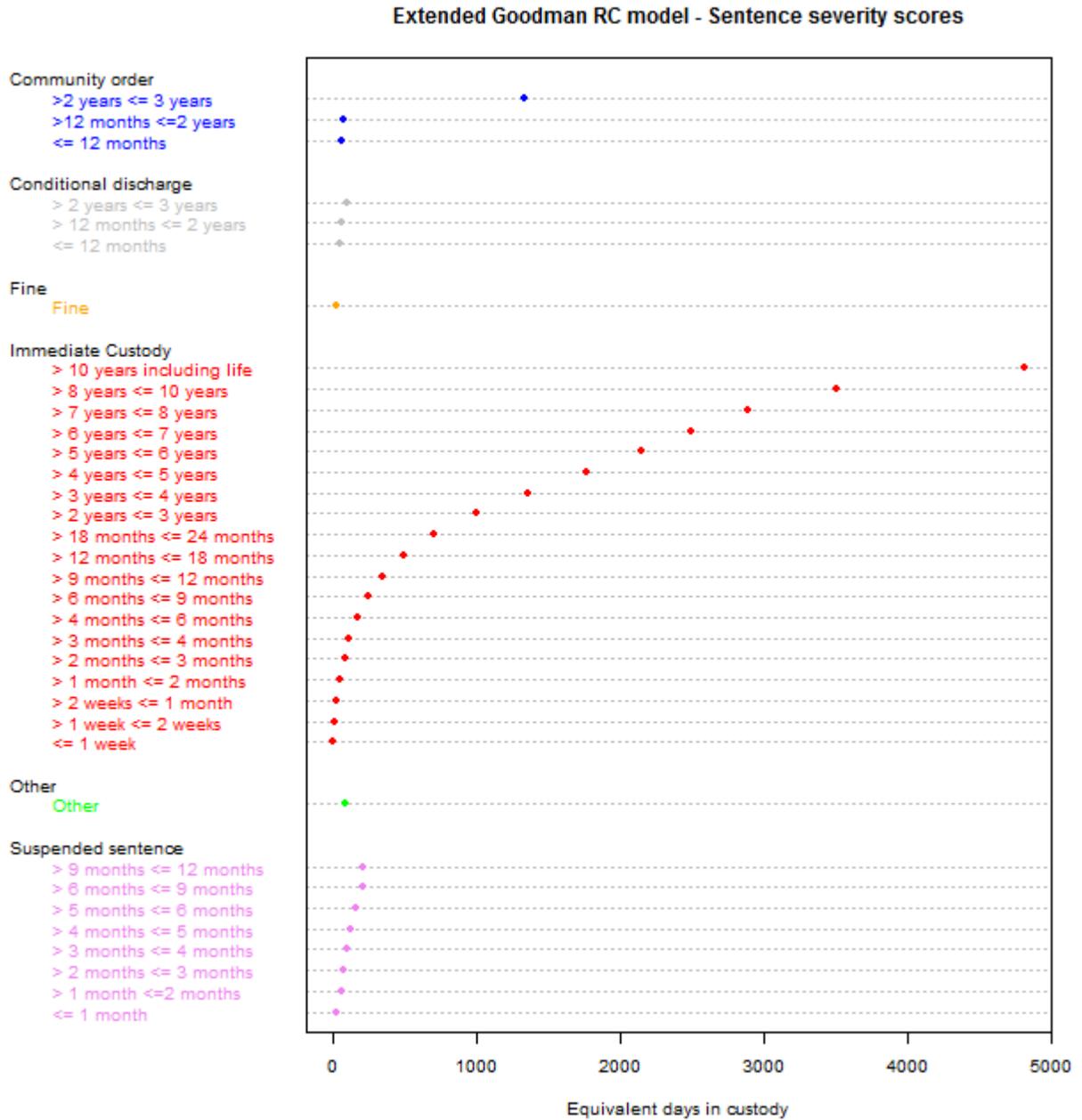
Figure 7.1: Dotplot - Extended Goodman RC model - Sentence severity scores

In creating a single severity scale which incorporates the full range of sentences issued by the courts in England and Wales, it is possible to observe the equivalent sentence severity score for each sentence, where offences of similar seriousness are given sentences of equivalent severity. I have previously argued that sentences can

and do overlap in terms of their severity, even though the type of sentence, or the punitive elements of the sentences are very different - such as monetary or liberty. This is referred to in the sentencing literature as the interchangeability of sentences. The dotplot in Figure 7.1 effectively illustrates this point by demonstrating that sentences are interchangeable, because sentences can and do overlap when comparing the severity of these sentences. By using the methodology (Extended Goodman RC association analysis) employed to create this sentence severity scale and incorporating the full range of sentences exercised by the courts, and then converting these scores into equivalent days in custody using the method of linear interpolation, it is evident that, some non-custodial sentences are essentially on par with the shorter custodial sentences. This notion of interchangeability is discussed further in Chapter 10 but first some examples of interchangeable sentences are provided to illustrate this point.

One example of interchangeability from this scale shows that, a community order of less than 12 months (δ_s = -4.028) is roughly equivalent to a suspended sentence order of greater than one month and less than or equal to 2 months (δ_s = -4.007). Following linear interpolation these two sentences are then approximately 65 and 66 days in custody respectively. Likewise the conditional discharge of less than or equal to a year category (δ_s = -4.115) equates to 63 days in custody following the linear interpolation. Another example indicates, the sentence category for greater than three and less than or equal to four years immediate custody (δ_s = 2.886) is relatively close to a community order of greater than two years and less than or equal to three years (δ_s = 2.821). In days the community order equates 1335 days in custody which is similar to the modal number of days for the immediate custody of greater than three years and less than four years category (1362 days).

The final scores obtained from the extended Goodman RC analysis and linear interpolation become the sentence severity scale, which was used to assess sentencing disparity.

7.3.2 Offence seriousness scores

In addition to the sentence severity scores, the extended Goodman model also estimates a number of γ_o parameters which relate to the offences (major offences). These offence scores can be seen in Table 7.5 ranging from the most to the least serious.

Table 7.5: Offences scores

Offence type	Raw score (γ_o)
Rape of male or female	1.73
Attempted murder	1.53
Buggery or attempted buggery	1.37
Murder	1.15
Incest with girl under 13 years old	1.11
Aggravated burglary in a dwelling (including attempts)	1.11
Unlawful sexual intercourse with girl under 13	1.08
Indecent assault on a male	1.05
Kidnapping	0.99
Gross indecency with a child	0.96
Unlawful sexual intercourse with girl under 16	0.94
Indecent assault on a female	0.92
Indecency between males	0.85
Blackmail	0.84
Robbery and assaults with intent to rob	0.83
Manslaughter	0.81
Wounding and other acts endangering life	0.77
Child prostitution	0.77
Abuse of trust - Sexual offences	0.69

Continued on next page

Table 7.5-*continued from previous page*

Offence type	Raw score (γ_o)
Obscene publications	0.64
Other criminal damage endangering life	0.59
Public Order Act 1986 Sec 2 Violent Disorder	0.58
False accounting	0.56
Immigration Act including people trafficking	0.55
Abduction of a child	0.53
Offences against laws relating to Customs, Excise and	0.53
Burglary in a dwelling	0.50
Sexual Offences Act 2003 including grooming and other	0.50
Procuration of males and females	0.50
Concealing	0.46
Other Forgery	0.46
Arson	0.44
Frauds by agents, trustees, company directors etc.	0.40
Stealing by an employee	0.39
Firearms offences	0.38
Bankruptcy	0.37
Cruelty to or neglect of children	0.37
Threats, Conspiracy or Incitement to murder	0.37
Theft of motor vehicle	0.35
Other frauds	0.33
Aggravated taking of a vehicle	0.31
Misuse of Drugs Act	0.30
Dangerous Driving	0.27

Table 7.5-*continued from previous page*

Offence type	Raw score (γ_o)
Unlawfully taking away or opening a mail bag	0.26
Trade Descriptions and similar offences	0.26
Other indictable offences	0.25
Intending to facilitate prison escape	0.25
Other stealing and unauthorised takings	0.22
Receiving/handling stolen goods	0.20
Permitting premises to be uses for unlawful (drug-related) purposes	0.19
Attempting to pervert the course of justice	0.18
Perjury	0.17
Malicious wounding and other like offences	0.16
Benefit offences	0.15
Stealing from another person	0.14
Forgery of prescription	0.13
Absconding from lawful custody	0.12
Burglary other than a dwelling	0.10
Other offences (against the State and Public Order)	0.09
Aggravated vehicle taking	0.05
Stealing from automatic machines	0.05
Threat or possession with intent to commit criminal damage	0.04
Stealing in a dwelling other than from automatic machine	0.00
Speed limit offences	-0.03
Health and Safety at work	-0.05
Firearms Act 1968	-0.12
Improper use of public electronic communications networks	-0.16

Table 7.5-*continued from previous page*

Offence type	Raw score (γ_o)
Cruelty to animals, ill treatment, neglect etc.	-0.17
Stealing and unauthorised takings	-0.20
Offences in relation to dogs Dangerous Dogs Act 1991	-0.20
Abstracting electricity	-0.21
Stealing from vehicles	-0.22
Accident offence	-0.24
Driving license offences	-0.24
Going equipped for stealing	-0.30
Vehicle Test Offences	-0.32
Road traffic offence, neglect, directions etc.	-0.32
Driving or attempting to drive after consuming alcohol or drugs	-0.34
Common and other types of assault	-0.37
Other offences excluding motoring offences	-0.37
Other criminal damage	-0.40
Vehicle insurance offences	-0.42
Assault on a constable	-0.42
Preventing child from receiving information	-0.52
Criminal Damage	-0.55
Offences against Public Order	-0.57
Careless driving	-0.58
Stealing pedal cycles	-0.59
Fraud, Forgery associated with vehicle or driving records	-0.63
Stealing from shops and stalls (shoplifting)	-0.64
Interference with a motor vehicle	-0.71

Table 7.5-*continued from previous page*

Offence type	Raw score (γ_o)
Picklock related offences	-0.81
Misuse of drug related offences	-0.93
Public health - By-laws	-0.95
Drunk and disorderly	-1.07
Bail offences	-1.14
Licensing Act offences	-1.36
Offences relating to Police regulations	-1.36
Transport offences	-1.72

The order of the major offences in Table 7.5 are determined by the γ_o parameters also obtained from the extending Goodman RC association model based on the association with the corresponding sentence categories and correcting for the imbalance of legal covariates. These parameters provide the magnitude of offence seriousness which then enable the offences to be ranked in order of offence seriousness just like the sentence severity scores shown in Table 7.2. Offences such as murder, attempted murder and rape offences are ranked amongst the most serious offences. At the other end of the scale bail offences, licensing Act offences, offences relating to Police regulations and transport offences are amongst the least serious offences.

Although there are no official Ministry of Justice offence seriousness scales in England and Wales, murder would be considered the most serious type of offence. Murder however, does not come out as the most serious offence in the scale. Rape, attempted murder and buggery (now re-classified as male rape) are all ranked slightly above murder. This could be explained by some anomalies in the data, such as the way this offence was recorded or the sentences they received. Or it could be the

result of the offence murder not being a very common type of offence with very few offences in the sample (n=5 or <1%). In Section 6.4 I explained that within a two-way contingency table of sentences by offences (see Table 6.1), the diagonal cells of a table would contain a high number of counts based on certain offences receiving certain types of sentences. Where certain offences are committed less frequency (for example murder) the counts in these cells will also be small which will reduce the accuracy of the estimates and this may cause certain offences to be out of the expected sequence.

Additionally, in all of these cases we do not know the actual culpability and harm caused or intended by the offender which determines the seriousness of all of the individual offences. For any given offence we have no information on whether the offence was rated as a category 1 offence indicating greater harm and higher culpability, a category 2 offence indicating greater harm and lower culpability or lesser harm and higher culpability, or a category 3 offence indicating lesser harm and lower culpability. Neither do we know the aggravating or mitigating circumstances of the offence which will ultimately determine the sentence passed by the judge. Therefore although we know the major offence code, we do not know exactly just how serious the offence faired in comparison to other similar types of offences. Lastly, we are unable to account for the discretion shown by the judge. It may be that some judges exhibit more discretion than others, which is likely to result in different sentences being passed.

The raw offence scores (γ_o) provide the magnitude of offence seriousness which allows the scores to be ranked in order of their seriousness. However like the sentence severity scores (δ_s) these scores also appear arbitrary and have no real meaning outwith the scale. As the focus of this research is on sentence severity, we do not attempt to standardise or anchor these scores to make them meaningful in their own right. The offence seriousness scores are not required for this research and therefore will not be discussed further.

7.3.3 Additional parameter estimates

A series of popularity parameters are also estimated in the Extended Goodman RC model for the offence (α_o) and sentence (β_s) categories. These estimates indicate the popularity or frequency of the categories compared to the reference categories, although these estimates might not be representative of offences in the general population. Table 7.6 displays the α_o estimates with Murder as the reference category and no α_o estimate. Table 7.7 displays the β_s estimates with less than one week in immediate custody as the reference category and again no β_s parameter is estimated for this category. Again these estimates do not provide any useful information for the purposes of measuring sentence severity and therefore do not feature in the remainder of this thesis.

Table 7.6: α estimates

Offence type	α_o estimates
Murder	
Attempted murder	-1.015
Threats, Conspiracy or Incitement to murder	3.892
Manslaughter	3.637
Wounding and other acts endangering life	5.131
Malicious wounding and other like offences	7.255
Cruelty to or neglect of children	4.747
Abduction of a child	2.193
Buggery or attempted buggery	-0.1849
Indecent assault on a male	4.021
Indecency between males	2.071
Rape of male or female	1.516

Table 7.6-*continued from previous page*

Offence type	α_o estimates
Indecent assault on a female	6.016
Unlawful sexual intercourse with girl under 13	3.511
Unlawful sexual intercourse with girl under 16	4.995
Incest with girl under 13 years old	3.244
Procuration of males and females	1.856
Burglary in a dwelling	6.361
Aggravated burglary in a dwelling (including attempts)	2.178
Burglary other than a dwelling	5.014
Going equipped for stealing	1.898
Robbery and assaults with intent to rob	5.66
Blackmail	2.674
Kidnapping	3.497
Aggravated taking of a vehicle	3.687
Concealing	5.039
Stealing from another person	4.476
Stealing in a dwelling other than from automatic machines and meters	3.01
Stealing by an employee	5.343
Unlawfully taking away or opening a mail bag	2.331
Abstracting electricity	1.343
Stealing pedal cycles	-0.8737
Stealing from vehicles	3.191
Stealing from shops and stalls (shoplifting)	3.386
Stealing from automatic machines	1.86
Theft of motor vehicle	3.685

Table 7.6-*continued from previous page*

Offence type	α_o estimates
Other stealing and unauthorised takings	5.76
Frauds by agents, trustees, company directors etc.	1.293
False accounting	3.92
Other frauds	7.093
Receiving/handling stolen goods	5.207
Bankruptcy	2.76
Arson	4.286
Other criminal damage endangering life	1.605
Other criminal damage	2.601
Threat or possession with intent to commit criminal damage	2.474
Forgery of prescription	0.9846
Other Forgery	5.375
Public Order Act 1986 Sec 2 Violent Disorder	3.111
Other offences (against the State and Public Order)	5.842
Perjury	1.558
Intending to facilitate prison escape	1.473
Child prostitution	1.579
Abuse of trust - Sexual offences	2.69
Gross indecency with a child	4.248
Immigration Act including people trafficking	2.706
Attempting to pervert the course of justice	4.357
Absconding from lawful custody	1.644
Firearms offences	4.952
Offences against laws relating to Customs, Excise and Inland Revenue	2.829

Table 7.6-*continued from previous page*

Offence type	α_o estimates
Bail offences	0.18
Trade Descriptions and similar offences	4.542
Health and Safety at work	-0.09262
Obscene publications	7.648
Sexual Offences Act 2003 including grooming and other sexual acts	4.864
Misuse of Drugs Act	7.456
Permitting premises to be uses for unlawful (drug-related) purposes	3.093
Other indictable offences	3.785
Assault on a constable	3.648
Common and other types of assault	5.586
Cruelty to animals, ill treatment, neglect etc.	2.79
Offences in relation to dogs Dangerous Dogs Act 1991	-0.1674
Preventing child from receiving information	-0.2949
Firearms Act 1968	1.252
Offences against Public Order	2.757
Interference with a motor vehicle	-0.528
Stealing and unauthorised takings	2.693
Aggravated vehicle taking	3.051
Drunk and disorderly	-2.06
Licensing Act offences	-6.659
Criminal Damage	3.522
Benefit offences	4.005
Offences relating to Police regulations	-6.657
Public health - By-laws	-4.064

Table 7.6-*continued from previous page*

Offence type	α_o estimates
Transport offences	-8.534
Picklock related offences	-2.994
Misuse of drug related offences	-1.609
Other offences excluding motoring offences	3.412
Improper use of public electronic communications network	2.531
Dangerous Driving	5.082
Driving or attempting to drive after consuming alcohol or drugs	3.962
Careless driving	-0.3625
Accident offence	2.208
Driving licence offences	4.197
Vehicle insurance offences	2.477
Vehicle Test Offences	-0.7597
Fraud, Forgery associated with vehicle or driving records	-2.108
Speed limit offences	0.2566
Road traffic offence, neglect, directions etc.	-0.05

7.4 Concluding remarks

Chapter 7 has reported on the results of the Goodman RC analysis, and the extended Goodman RC analysis. The extended Goodman RC model was the optimal and preferred model because it controlled for the three legal variables. Linear interpolation then made it possible to calculate the equivalent number of days in custody for each of the non-custodial sentence categories based on the modal number of days for each

of the immediate custody categories. The end result was a series of final scores which are now anchored and therefore provides meaningful estimates rather than arbitrary measures.

The final scores - displayed in Table 7.2 - are obtained from the extended Goodman RC analysis and linear interpolation and now become the scale in which to assess sentencing disparity. The scores provide an approximate measure of sentence severity that incorporates the range of sentences issued by the courts in England and Wales, modelling real sentencing data to produce an objective measure of sentence severity.

The next chapter will discuss the second stage of this PhD thesis: identifying and assessing disparity in sentencing. Chapter 8 will explain the multilevel modelling approach which will use the newly formed sentence severity scale, and a variety of explanatory variables to facilitate this research into sentencing disparity.

Table 7.7: β_s estimates

Disposal	Sentence	β_s estimate
Immediate custody	$\leq$ 1 week	
	> 1 week $\leq$ 2 weeks	0.260
	> 2 week $\leq$ 1 month	2.407
	> 1 month $\leq$ 2 months	2.761
	> 2 months $\leq$ 3 months	3.087
	> 3 months $\leq$ 4 months	2.982
	> 4 months $\leq$ 6 months	3.079
	> 6 months $\leq$ 9 months	2.399
	> 9 months $\leq$ 12 months	2.567
	> 12 months $\leq$ 18 months	2.476
	> 18 months $\leq$ 24 months	1.993
	> 2 years $\leq$ 3 years	1.815
	> 3 years $\leq$ 4 years	0.780
	> 4 years $\leq$ 5 years	0.750
	> 5 years $\leq$ 6 years	1.547
	> 6 years $\leq$ 7 years	2.604
	> 7 years $\leq$ 8 years	2.813
	> 8 years $\leq$ 10 years	4.787
	> 10 years including life	6.269
Suspended sentence	$\leq$ 1 month	1.952
	> 1 month $\leq$ 2 months	2.653
	> 2 months $\leq$ 3 months	3.076
	> 3 months $\leq$ 4 months	3.127
	> 4 months $\leq$ 5 months	1.903
	> 5 months $\leq$ 6 months	2.400
	> 6 months $\leq$ 9 months	2.867
	> 9 months $\leq$ 12 months	2.833
Community order	$\leq$ 12 months	5.043
	> 12 months $\leq$ 2 years	3.084
	> 2 years $\leq$ 3 years	0.858
Fine	Fine	2.281
Conditional discharge	$\leq$ 12 months	1.925
	> 12 months $\leq$ 2 years	1.264
	> 2 years $\leq$ 3 years	0.477
Other	Other	3.493

Chapter 8

Measuring sentencing disparity using multilevel modelling

8.1 Introduction

In Chapter 4, I explained that one of the most common methods used to measure sentencing disparity is multilevel modelling. The aim of multilevel modelling is to construct a model that expresses how the dependent variable - sentence severity - depends on, or is explained by the exploratory - independent - variables (Snijders and Bosker, 2012). In the past, studies that have used multilevel modelling, such as those conducted by Pina-Sánchez and Linacre (2013), Ulmer and Johnson (2004), and Johnson (2006), have tended to use sentence length or the probability of incarceration as the dependent variable, and numerous explanatory variables to assess sentencing disparity. However, all of these studies have not been able to include non-custodial sentences, such as fines and community orders, even though these sentences are used more often than custodial sentences (see Section 5.4.2) and are also just as important to include as custodial sentences (von Hirsch et al., 1989). This research differs in that it uses the new measure of sentence severity - the sentence severity scale or a

transformation of it - as the dependent variable, which incorporates the full range of offences meted out by the courts in England and Wales. It then models the effect of legal and extra-legal characteristics of the offence and the offender, on sentence severity to assess sentencing disparity. This chapter will explain multilevel modelling and how it is used to identify and measure sentencing disparity.

8.2 Multilevel modelling

The goal of multilevel modelling is to predict values of the dependent variable based on a function of predictor or explanatory variables at more than one level (Luke, 2004). Multilevel modelling enables me to examine how sentencing is influenced by legal factors at the offence level and extra-legal factors at the offender level. In section 4.4.1, I explained that sentencing disparity is intertwined with sentence severity: only by measuring sentence severity can we assess disparity. Consequently, I use multilevel modelling to predict the effect that the explanatory variables have on sentencing (sentence severity) by defining which are statistically significant to determine which factors are attributing towards this disparity.

Multilevel modelling is a type of regression analysis, used to analyse data with a hierarchical or nested structure (Bryk and Raudenbush, 1992). It is used in a variety of fields, and is particularly popular in educational research, specifically education attainment. In this context there are two levels; pupils (level-1) and schools (level-2), where pupils are nested within schools: multilevel modelling is then used to assess the factors which impact on test results at both these levels. Multilevel modelling has also become increasingly popular within the field of criminology, with numerous studies adopting this method, such as those mentioned in section 4.4.5. In my data, I also have a nested structure and therefore, it is appropriate to also use multilevel modelling. In this case, offences (level-1) are nested within individual offenders (level-

2) and these act as levels within the multilevel structure. If there had been court level information available, such as the specific court that the offender was sentenced in, this would have been a third-level in the multilevel model. It its absence, I will test a third-level, that is the Probation Trust area in which the court is located. I will discuss this further in sections 8.2.2 and 8.6.

Regression analysis that does not take account of the multilevel structure in the data is making the assumption that the observations are independent of one another (Luke, 2004). Where there is a nested structure, there is a good chance that this independence assumption will be violated (ibid), which can lead to biased parameter estimates (Weidner et al., 2005). We would expect that the offences committed by one specific offender are more similar in nature, such as two offences for theft and a third for handling stolen goods. Moreover the sentence for that offender will be more similar due to unmeasured variables, such as attitude towards the court.

Single-level multivariate techniques are also likely to underestimate standard errors and in doing so exaggerate the significance of parameter estimates (Weidner et al., 2005).

The ecological fallacy

The ecological fallacy occurs when conclusions are drawn about individuals based only on the analysis of grouped or aggregate data (Freedman, 2001). In this case it would be problematic to draw conclusions about sentencing disparity without taking into account the offence(s) that the offender was sentenced for. For example, an offender sentenced for assault will receive a more severe sentence than an offender sentenced for being drunk and disorderly. Without taking this level-1 factor into account, it would lead us to make unsound conclusions about the severity of the sentences offenders received and ultimately sentencing disparity.

8.2.1 Variance components model

The variance components model/random effects model containing no covariates is referred to as the null or 'empty model' Snijders and Bosker (2012). This is the simplest type of multilevel model and is usually the starting point for any multilevel modelling analysis. The variance components model is

$$Y_{ij} = \beta_0 + u_j + e_{ij}$$

where Y_{ij} is the predicted sentence severity score for the level-1 units (offence i) within the level-2 units (offender j), β_0 is the intercept, and u_j and e_{ij} are the random error terms at the offender level (level-2 units) and for offence i in group j, respectively.

The intercept (β_0) can be interpreted as the average value of the dependent variable across all offenders in the sample (Luke, 2004).

Multilevel modelling differs from regression analysis in that the equation defining the model contains more than one error term - one error term for each level in the analysis (Snijders and Bosker, 2012). We assume that u_j and e_{ij} have Normal distributions as follows;

$$u_j \sim Normal(0, \sigma_u^2)$$

$$e_{ij} \sim Normal(0, \sigma_e^2)$$

The initial variation in the data is usually calculated at this stage using the intraclass correlation coefficient.

Intraclass correlation coefficient (ICC)

The intraclass correlation coefficient (ICC) measures the proportion of variance in the dependent variable that is accounted for by the groups in the data (Hox, 2010) (i.e. the proportion of variation in sentence severity - disparity - between offenders (level 2 units)), as shown by:

$$p = \frac{\sigma_u^2}{(\sigma_u^2 + \sigma_e^2)}$$

Here σ_u^2 is the estimate for the level-2 variance and σ_e^2 is the estimate for the level-1 variance. The equation therefore states that that ICC is the proportion of between group variance compared to the total variance.

By adding covariates - independent explanatory variables - to the model, we can then begin to account for some of this sentencing variability. To do so, I fit a random intercept model.

8.2.2 Random intercept model

The random intercept model builds on the variance components model by adding P explanatory variables or covariates to the model, as shown by;

$$Y_{ij} = \beta_0 + \sum_{p=1}^{P} \beta_p x_{ijp} + u_j + e_{ij}$$

There are actually two parts to this equation: the fixed part $\left[\beta_0 + \sum_{p=1}^{P} \beta_p x_{ijp}\right]$ and the random part $[u_j + e_{ij}]$. The fixed effects part of the equation controls for the explanatory variables. Adding explanatory variables to the model aims to account

for some of the sentencing variation. For example, it is important to control for the type of offence the offender is being sentenced for, as this will/should be the most important determinant. The random part of the equation then contains the variation at the offender level (σ_u^2) and offence level (σ_e^2) that is left unexplained in the model. As the covariates - legal and non-legal factors - are added to the model, we would anticipate the variation in sentence severity to go down, providing these factors are influencing judges' decisions when sentencing.

So far I have spoken about a random intercept model that has two levels; offences at level-1 and offenders at level-2. However, it is possible to introduce a third level into the analysis to identify if there is any geographical variation[21] in sentencing outcomes. The Probation Trust area is therefore considered as the third level in the multilevel model.

Therefore the model can now be shown as,

$$Y_{ijk} = \beta_0 + \sum_{p=1}^{P} \beta_p x_{ijkp} + v_k + u_j + e_{ij}$$

The k subscript now relates to the third level in the multi level model and v_k is the additional error term where

$$v_k \sim Normal(0, \sigma_v^2)$$

Possible extensions to the model

If we take the three level model as an example, the random intercept model allows each Probation Trust area (v_k) to contribute an "effect" to sentence severity i.e. to

---

[21]Prior to obtaining the data for this research, the data was fully anonoymised by the OASys Data Evaluation and Analysis Team (O-DEAT). This meant that the information relating to the specific court the offender was sentenced in was removed from the data set.

have a different intercept but equal slopes. This is illustrated in the left-hand diagram (a) in Figure 8.1 for four Probation Trusts, where each Probation Trust's line crosses the y-axis at a different point. In this instance, covariates in the model have the same affect on sentence severity for each Probation Trust area.

Figure 8.1: Multilevel model examples
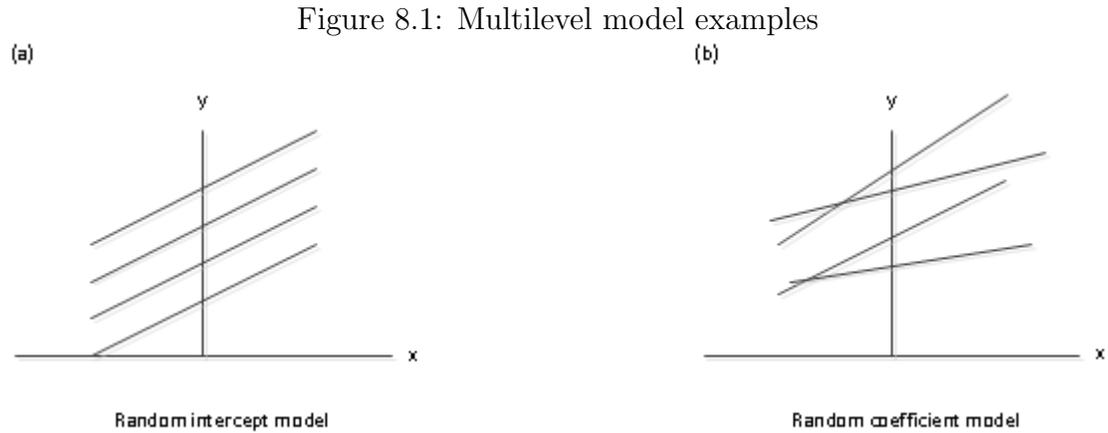
Another way of extending this model would be to introduce a random coefficient model. A random coefficient model would allow the model to have different intercepts, as well as different slopes. This is illustrated in the right-hand diagram (b) in Figure 8.1 for four Probation Trusts, where each Probation Trust's line crosses the y-axis at a different point and the slopes are no longer parallel. This indicates that the relationship between one of more of the covariates and sentence severity varies by Probation Trust. For example, this could be in relation to ethnicity, where the effects of ethnicity on sentence severity are greater in one Probation Trust compared to another Probation Trust.

In this research the random coefficient model was not fitted. In Section 8.6 I explain that the optimal model is the two-level model, and therefore fitting a random coefficient model where the intercepts and slopes for 13,495 offenders vary, would be overly complicated. Perhaps more importantly, substantively it makes little sense. On the other hand, if the three-level model had been the optimal model then fitting a random coefficient model would have been appropriate to identify if the relationship

between covariates varied by Probation Trust area.

8.3 The dependent variable of sentence severity

One of the major differences between this research and research that has been conducted in the past is the development of a new measure of sentence severity. In Chapter 7 I displayed the results from the extended Goodman RC association analysis (see Table 7.2), which produced the new sentence severity scale (raw scores) and the linear interpolation of this into equivalent days in custody (final scores). The final scores (days in custody) are treated as continuous for modelling purposes, which is used to model sentence severity and assess any sentencing disparity.

8.3.1 Log transformation of dependent variable

In Chapter 4, I explained that a number of studies, such as those conducted by Helms and Jacobs (2002) and Britt (2009), take the natural logarithm of the dependent variable - sentence length - to reduce the skew in the data. This is because more offenders receive shorter custodial sentences than mid-length or longer custodial sentences. The same skew occurs when using the new measure of sentence severity, as the majority of offenders will receive less severe/shorter sentences. Few offenders receive sentences in excess of 1000 days. Figure 8.2 displays the distribution of the equivalent days in custody for the included sample. The skew in the data is very apparent and therefore to remedy this I also take the natural logarithm of the dependent variable - final scores (equivalent days in custody). This reduces the skew in the data, which can be seen in Figure 8.3.

Once transformed, the new dependent variable is treated as a continuous variable with normal distributions assumed for the random effects. The data is treated as continuous because the dependent variable is a scale, and therefore approximates

normality (Hox, 2002). Furthermore, the bias of treating categorical variables as continuous is very small when there are a large numbers of categories (n=35) (Hox, 2002).

Figure 8.2: Distribution of the equivalent number of days in custody

Equivalent days in custody
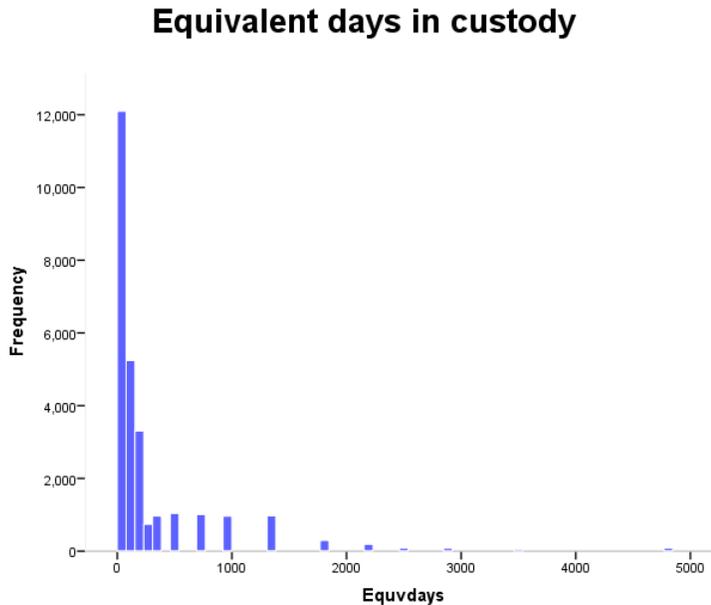
8.4 Independent variables

The covariates used in this analysis reflect the data that was available and where possible, I have tried to control for the OASys risks factors, which I discussed in Chapter 5. However, a number of variables in the dataset had large amounts of missing data, and therefore it was not possible to account for all the factors shown in Figure 5.1. It was also decided at this stage to also omit the cases where the offence plea was recorded as unknown (n=2,227) as no further information was recorded to explain why the plea was unknown. Some of the other covariates included in the analysis had some missing data in relation to the offender's ethnicity (n=4,408),

Figure 8.3: Distribution of the Log transformed equivalent number of days in custody

Log transformed equivalent days in custody
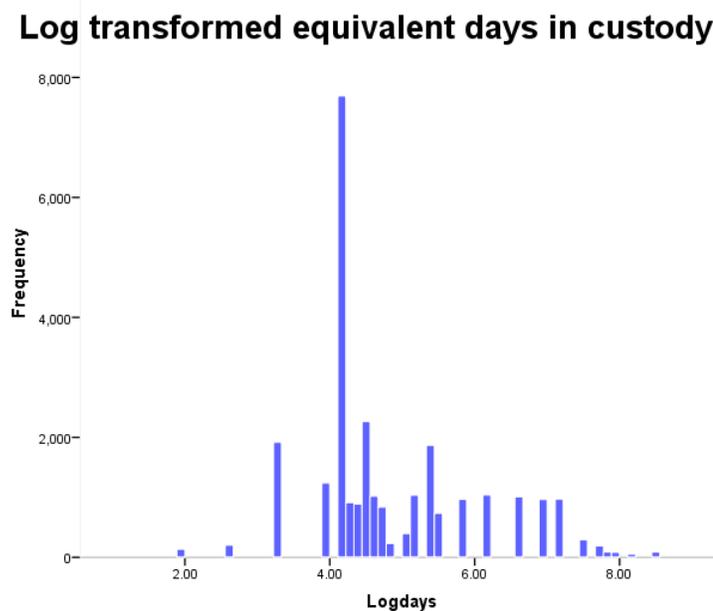
accommodation status (n=5,094), employment status (n=5,050), and income status (n=5,082). An example of the raw data used in the multilevel modelling analysis can be seen in Appendix D.

The next section will discuss how these cases were dealt with. The following section will then go on to introduce the independent variables used in the analysis.

Missing cases

For the multilevel modelling analysis a complete case analysis was undertaken: cases with missing data were removed from the sample based on the assumption they were missing completely at random (MCAR). Treating the data as MCAR assumes that the missing data is unconnected to any of the characteristics of the offenders in the sample. The data is a sub-sample of pre-sentence reports carried out between March 2008 and April 2010. It is my belief that the missing data is the result of those who carried out the original OASys assessments and who failed to complete sections of the

assessment prior to these becoming mandatory.

To test the MCAR assumption an analysis of the missing sample (excluded sample) characteristics was carried out and this was compared with the characteristics of complete case data (included sample).

Table 8.1: Full sample

Type of covariate	Covariate	levels	Full[1] Frequency (%)	Included[2] Frequency (%)	Excluded[3] Frequency (%)
Legal	Plea	Guilty	41,589* (83.9)	25,265* (93.1)	16.321* (91.4)
		Not guilty	3,397* (6.9)	1,860* (6.9)	1,537* (8.6)
		Unknown	4,604* (9.3)		
	Numoff	<2 offences	18,726 (76.3)	10,619 (76.1)	6,559 (76.2)
		>3 offences	5,812 (23.7)	3,326 (23.9)	2,054 (23.8)
	Prior	Prior	19,173 (78.1)	11,576 (83.0)	6,064 (70.4)
		No prior	5,363 (21.9)	2,369 (17.0)	2,548 (29.6)
Additional legal covariates	Court	Crown	10,360 (42.2)	5,394 (38.7)	4,162 (48.3)
		Magistrates'	14,178 (57.8)	8,551 (61.3)	4,451 (51.7)
	Weapon	No	43,052* (86.8)	23,455* (86.5)	15,579* (87.2)
		Yes	6,537* (13.2)	3,670* (13.5)	2,279* (12.8)
	Direct victim	No	333,319* (67.2)	18,116* (66.8)	12,004* (67.2)
		Yes	16,270* (32.8)	9,009* (33.2)	5,854* (32.8)

Continued on next page

Table 8.1-*continued from previous page*

Type of covariate	Covariate	levels	Full[1] Frequency (%)	Included[2] Frequency (%)	Excluded[3] Frequency (%)
Extra-legal	Sex	Male	21,549 (87.8)	12,272 (88.0)	7,535 (87.5)
		Female	2,989 (12.2)	1,673 (12.0)	1,078 (12.5)
	Ethnicity	White UK and Irish	15,419 (62.8)	11,026 (79.1)	3,115 (36.2)
		White other	845 (3.4)	541 (3.9)	224 (2.6)
		Mixed	594 (2.4)	416 (3.0)	130 (1.5)
		Asian	1,117 (4.6)	738 (5.3)	286 (3.3)
		Black	1,601 (6.5)	1,089 (7.8)	396 (4.6)
		Other	201 (0.8)	135 (1.0)	54 (0.6)
		Missing	4,761 (19.4)		4,408 (51.2)
	Age category	21-24	5,284 (21.5)	3,037 (21.8)	2,548 (29.6)
		25-29	5,388 (22.0)	3,143 (22.5)	1,806 (21.0)
		30-40	7,979 (32.5)	4,692 (33.6)	2,646(30.7)
		41-54	4,942 (20.1)	2,617 (18.8)	1,917 (22.3)
		55+	945 (3.9)	456 (3.3)	415 (4.8)

Continued on next page

Table 8.1-*continued from previous page*

Type of covariate	Covariate	levels	Full[1] Frequency (%)	Included[2] Frequency (%)	Excluded[3] Frequency (%)
	Employment status	Employed	5,697 (23.2)	3,851 (27.6)	1,306 (15.2)
		Temp/casual	478 (1.9)	359 (2.6)	80 (0.9)
		Educ/training	314 (1.3)	220 (1.6)	64 (0.7)
		Unemployed	9,504 (38.7)	7,205 (51.7)	1,486 (17.3)
		Unavailable for work	3,224 (13.1)	2,310 (16.6)	627 (7.3)
		Missing	5,321 (21.7)		5,050 (58.6)
	Income status	Wages	6,456 (26.3)	4,408 (31.6)	1,448 (16.8)
		Pension	217 (0.9)	132 (0.9)	62 (0.7)
		State benefits	10,607 (43.2)	8,058 (57.8)	1,645 (19.1)
		Other	634 (2.6)	443 (3.2)	142 (1.6)
		No income	1,271 (5.2)	904 (6.5)	234 (2.7)
		Missing	5,353 (21.8)		5,082 (59.0)
	Accommodation status	Permanent	14,698 (59.9)	10,556 (75.7)	2,872 (33.3)
		Probation/supported	592 (2.4)	452 (3.2)	86 (1.0)

Table 8.1-*continued from previous page*

Type of covariate	Covariate	levels	Full[1] Frequency (%)	Included[2] Frequency (%)	Excluded[3] Frequency (%)
		Short-term	2,855 (11.6)	2,105 (15.1)	458(5.3)
		No fixed abode	955 (4.2)	832 (6.0)	123(1.4)
		Missing	5,074 (22.5)		5,074 (58.9)
	Drug misuse	No	9,168 (37.4)	4,585 (32.9)	3,876 (45.0)
		Yes	15,370 (62.6)	9,360 (67.1)	4,737 (55.0)

[1]Offenders in Full sample = 24,538 Offences in Full sample = 49,590

[2]Offenders in Included sample = 13,945; Offences in Included sample = 27,125

[3]Offenders in Excluded sample = 8613; Offences in Excluded sample = 17,858

*Figures relate to offences

Table 8.1 displays the comparisons of the full sample (used to conduct the Goodman RC analysis), included sample (used to conduct the multilevel model analysis) and the excluded sample (missing data) for all of the covariates used in this research, grouped by legal covariates, additional legal covariates and extra-legal characteristics of the offenders. All of the frequencies and percentages in the table are aggregated to the offender level, except for plea, weapon and direct victim which relate directly to each offence and these are indicated with asterisk.

The biggest difference between the three samples relate to the offence plea which is the result of removing offenders where at least one of their offences was unknown (n=2,227). The included and excluded samples show similar proportions pleading guilty and not guilty. Similarly, the proportion of the number of offences the offenders are sentenced for are also similar across all three samples. Although the included sample have a higher proportion of offenders who have had a prior court appearance (83%) compared to the excluded sample (70%). This could potentially indicate that assessments are completed more thoroughly if offenders have a history of offending (i.e. no fields are left incomplete resulting in no missing data) than if the offender had no history of offending. Although, we have no way of validating this hypothesis. Comparing the additional legal covariates reveals no noteworthy differences between the three samples

The same proportion of males and females are present in all three samples, as are the proportions. The age groups of offenders are also similar for the full and included samples, with slightly more younger (aged 21-24) and older (aged 55+) offenders in the excluded sample.

It seems that there are only very subtle differences in the samples which indicates the missing data is MCAR and potentially an administrative oversight is responsible for the missing data.

Snijders and Bosker (2012) explains there are two basic reasons for attempting to

deal with missing data in a suitable way. The first is that in some cases, the absence of the data is related to meaningful information, and ignoring this can lead to bias in the results. Secondly, deleting incomplete cases is wasteful and leads to unnecessarily large standard errors of estimators. Although there are other more appropriate ways of dealing with missing data, such as maximum likelihood and multiple imputation (Snijders and Bosker, 2012), it was felt that there were too many missing data points (9,226) to accurately use imputation methods. It was felt that even with the reduced sample size, the sample was still sufficiently large.

8.4.1 Legal covariates

This section will now explain the legal covariates or characteristics of the offence which were added first in the multilevel model building stage (see Section 8.6). These covariates relate to the offence which are level 1 explanatory variables, although prior court appearances and number of offences are level 2 explanatory variables. These covariates are listed below: the variable name is given in parentheses and the proportions reported are based on the included sample as shown in Table 8.1.

Offence type(*Offence*)

This variable relates to the type of offence the offender was sentenced for. The 99 major offence categories were previously displayed in Table 5.1. It was anticipated that offence type would be the most important factor in explaining a substantial amount of the sentencing variation based on the principle of proportionality.

Plea(*Plea*)

Offence plea is a legal characteristic: an early guilty plea can reduce the severity of the sentence by up to 30% (Sentencing Guidelines Council, 2004). In the included sample, 93% of all the offences resulted in a guilty plea, the remaining 7% resulted

in a not guilty plea. A guilty plea is the reference category in the multilevel model analysis.

Prior court appearance (*Prior*)

This variable acts as an indicator of previous offending. It is a dichotomous variable: 83% of the sample had previously appeared in court (labelled as 'prior') and therefore have a history of offending, 17% of the sample had not previously appeared in court ('no prior'). The category for no prior court appearance is the reference category.

Number of offences (*Numoff*)

This variable records the aggregate number of offences that an offender is being sentenced for on that sentencing occasion in which they were subject to a PSR. This variable has been dichotomised into 'two or less offences' (approximately 76% of the sample) and 'three or more offences' (approximately 24% of the sample). The 'less than two offences' category is the reference category.

8.4.2 Additional legal covariates

These are the additional legal covariates which were fitted during stage two of building the multilevel model. The covariates indicating whether there was a direct victim or a weapon used are both level 1 explanatory variables and the type of court the offender was sentenced in is a level 2 explanatory variable.

Type of court (*Court*)

The type of court the offender was sentenced in is a dichotomous variable: 'magistrates' court' or 'Crown Court'. In the included sample, a bigger proportion of offences (61%) were tried in the magistrates' court, while the other 39% of offences were tried in the Crown Court. The 'Crown' court category is the reference category. In 2009,

93% of offences were sentenced at the magistrates' court Ministry of Justice (2009) which illustrates that the more serious cases are included in this sample due to using OASys data.

Direct Victim (*Dir*)

This variable indicates if the offence involved a direct victim, either through contact or targeting, which inevitably makes an offence more serious (Sentencing Guidelines Council, 2004). In the included sample, fewer than 10,000 cases (33% of offences) involved a direct victim. This is a dichotomous variable indicating whether the offence involved a direct victim - 'yes' or 'no'. No direct victim is the reference category.

Weapon (*Weap*)

This variable indicates whether a weapon was carried during the offence. An offence which involves carrying or using a weapon increases the seriousness of that offence (Sentencing Guidelines Council, 2004). In this sample, 14% of cases were identified as using a weapon ('yes'). Again this is a dichotomous variable, whereby 'no' is the reference category.

8.4.3 Extra-legal - characteristics of the offender

These covariates were added at the third stage of the model building. All of these variable relate to the offender and are therefore level 2 explanatory variables.

Sex (*Sex*)

The sample is male dominated: 88% of the included sample being male, and 12% being female. The 'female' category is the reference category.

Ethnicity (*Eth*)

The majority (79%) of the included sample are categorised as 'White British and Irish' (reference category) and a further 4% are categorised as 'White Other'. The second largest group are 'Black' offenders (8%) followed by 'Asian' offenders (5%). The 'Mixed' offender category and 'Other' offender category account for 3% and 1% of the sample respectively.

Flood-Page and Mackie (1998) do heed warning over using such broad ethnic categories as this may mask some of the findings, and therefore it would be more robust to refine the ethnic categories; for example to account for Pakistani, Indian, and Bangladeshi offenders in the Asian category. However, the numbers in these sub-categories were too small and therefore had to be grouped together.

Age

Over the age of 21, age is not considered a relevant factor in sentencing. Within the literature and previous research into disparity age has tended to be under-conceptualised (Mueller-Johnson and Dhami, 2009) and therefore it was included in this analysis. Age was organised into 5 age categories; 21-24 (22%), 25-29 (23%), 30-40 (34%), 41-54 (19%), and 55+ (3%). The reference category was the youngest age group (21-24). Age was included as a categorical measure to access if there was any significant differences between age groups. A continuous measure of age or the square root of age would only have identified if there was a linear or curve-linear relationship.

Employment (*Emp*)

The majority of the sample were 'unemployed' (52%) prior to sentencing, with a further 17% were recorded as 'unavailable to work' (for example, those who are unable to work due to incapacity or ill health). 28% were employed (either full or part-time), while just under 3% were in temporary or casual employment. The remaining 2% were

in education or training prior to sentencing. The employed group are the reference category.

Income (*Inc*)

A large proportion of the sample (58%) were in receipt of benefits prior to sentencing, less than 1% were living on pensions or maintenance payments. Just under 7% stated they had no income prior to sentencing, whilst 3% were classified as 'other'. Only 32% of the sample relied on their wages or partners wages as their source of income. This is the reference category.

Accommodation status (*Acc*)

The majority (76%) of the sample reported they lived in 'permanent independent housing' (reference category). Just over 3% lived in 'probation or supported housing' and 15% were living in short-term transient accommodation. The remaining 6% of the sample were recorded as having 'no fixed abode'.

Misuse of drugs (*Drugs*)

The UK Drug Policy Commission (UKDPC) report problem drug users are more likely to be found within the criminal justice system than within the wider population (UK Drug Policy Commission, 2008). Problem drug users also tend to have higher rates of offending and are particularly associated with acquisitive crimes, such as shoplifting and burglary (ibid). This variable is dichotomous, in which 67% of the offenders in the sample admitted to misusing drugs - 'yes', the other 33% reported they had not misused drugs - 'no' (reference category).

8.4.4 Variables not available to use in the analysis

A number of variables, such as the number of the offender's prior convictions, previous spent sentences, the court (location and size) the offender was sentenced in, and the characteristics of the judges, i.e. their gender and ethnicity was not available in the data. Had this information been available then we would have a more accurate picture of the possible factors impacting on sentencing. Additionally, a number of variables that were available form OASys could also not be included, such as, marital status, attitudinal factors, risk of recidivism scores etc., as there was a substantial amount of missing data.

At this stage it is also important to highlight that it is not possible to ever fully explain sentencing disparity by modelling the data: there will always be some variability that cannot be explained. As I have previously highlighted, sentencing does not occur in a vacuum: there may be factors impacting on sentences that are not recorded in the data, such as the offender's attitude in court or the judge's mood. We would therefore expect there to be some unexplained variation in sentencing to account for these extraneous factors.

8.5 Software

A number of software packages (e.g. R, SPSS, SAS and STATA) allow multilevel models to be fitted. There are also special software packages (e.g. HLM, MLwiN, and the MOXOR suite (Snijders and Bosker, 2012)) that are specifically aimed at applying multilevel modelling techniques and have been designed by pioneers in the field of multilevel modelling (Snijders and Bosker, 2012). Snijders and Bosker (2012) describes MLwiN as the most extensive multilevel package, written by researchers at the University of Bristol. It is a flexible package, which also allows for data manipulation, graphing, simple statistical computations and file manipulation (ibid). For these

reasons, it seemed sensible to use this package to conduct the multilevel analysis for this research.

After attempting to fit the first basic model in MLwiN, it became clear that due to the amount of data and the memory capacity of my machine, it would not be possible to continue with this software. The same problem also occurred when using the R software environment on my own machine, even with a reduced sample size. Therefore, I again used Lancaster University's High End Computing Cluster. Although this service supports a number of third-party software, including R, it does not support MLwiN.

Consequently, the multilevel modelling analysis was conducted using the R software environment and the University's High End Computing Cluster service. An example of the layout of the dataset used to conduct the multilevel modelling can be seen in Appendix D.

The 'lme4' package by Bates et al. (2014) provided the functions for fitting the multilevel models. The final model (Model 24) was fitted using the following command:

$$MODEL24 \leftarrow lmer(Logdays \sim OFFENCE + Plea + Numoff + Court + Emp + Acc$$
$$+ Inc + Sex + Eth + Age + (1 \mid SentOcc), data = DATA, REML = FALSE)$$

A sample of the code used to create the optimal multilevel model in R can be seen in Appendix F.

8.6   Model fitting

A typical approach to multilevel model building according to Luke (2004) is to work from the bottom up. By this, he means fitting a variance components model or

null model with no explanatory variables/covariates to assess the initial variation in the model and then gradually adding each covariate to the model one at a time. Luke (2004) suggests starting with the level 1 explanatory variables before going on and including level 2 and possibly level 3 explanatory variables. In doing so, it is hoped that the covariates will reduce the deviance in the model which will go some way towards explaining the observed data. This is slightly different to the technique utilised in this work, which I will now explain.

The model building for this part of the analysis was essentially conducted in four stages:

Stage 1

   Legal covariates; Offence, Plea, Numoff, and Prior were added

Stage 2

   Additional legal covariates; Court, Dir, and Weap were added

Stage 3

   Extra-legal covariates; Sex, Age, Eth, Emp, Acc, Inc, Drugs, and interactions Sex*Eth, and Sex*Age were added

Stage 4

   Third level; Probation Trust was added

The first stage sees the main legal covariates added to the model (the same three legal variables that were used in the previous analysis to estimate the severity scale, and discussed in Chapter 6) as well as the offence covariate, stage two extended this model by adding three additional legal covariates, stage three incorporated the extra-legal covariates relating to the characteristics of the offender including two interactions and finally, stage four extended the model to account for a third level. Each of these stages and the corresponding tables will be discussed in turn.

To note, once again maximum likelihood estimation was used to assess model fit and provide the model deviance (minus twice the log-likelihood ratio of the model compared to the saturated model). To test whether the additional covariates improved model fit, the reduction in deviance and the degrees of freedom for each of the nested models was used to provide a p-value, which suggested whether there was sufficient evidence at the 1% level to reject or not reject the additional parameters in the model.

The first stage of the model building process is displayed in Table 8.2. The first model fitted was the variance components or null model. This model provided an estimate of the initial variance in sentence severity between offenders, which will be discussed in the following results chapter. From here, the main legal covariates were added to the model.

Arguably the most important variable to be controlled for in this type of analysis is the type of offence that the offender was sentenced for. As can be seen from Table 8.2, by controlling for this covariate alone, the deviance was reduced by 11,636 on 98 degrees of freedom. This yielded a p-value of <0.0001, which indicated there was sufficient evidence to suggest this additional parameter should remain in the model. This model now becomes known as the base model, and the additional legal covariates are added to this model.

The deviance of Models 2-4 were compared to the base model (Model 1). The covariates relating to the offender's plea and the number of offences they were being sentenced for at the sentencing occasion were highly significant, indicating they should be included in the model. The covariate indicating prior court appearances was not significant, suggesting there was insufficient evidence to include this additional covariate. It may be a little surprising to see that this covariate was not statistically significant when the sentencing legislation states prior offending as an aggravating factor. However, the prior court appearance only acts as an indicator of prior offending and is not a direct measure.

Combinations of the legal covariates were then fitted. *Plea* was then added to the model controlling for *Numoff* (Model 3) and this became Model 5. The deviance from Model 5 was then compared to that of Model 3. As a final check to see whether having controlled for plea and number of offences, prior court appearance should now be included: *Prior* was added to create model 6 and compared with the previous model (Model 5). This additional parameter reduced the deviance by 1.86 on 1 degree of freedom. The resultant p-value (p=0.1726) was not significant and therefore this covariate was not included in the model. Model 5 provided the best model fit and will be discussed in the following results chapter (Chapter 9).

Table 8.2: Stage one multilevel model building (legal covariates): Goodness of fit results I

Model	Covariates	Deviance	Difference in deviance	Difference in degrees of freedom	p-value
0	Null	64,504.78			
1	Base (Offence type)	52869.20	11,636	98	<0.0001
2	Base, Plea	52,804.20	65	1	<0.0001
3	Base, Number of offences (Numoff)	52,770.19	99.00	1	<0.0001
4	Base, Prior court appearance (Prior)	52,867.19	2.01	1	0.1566
5	**Base, Numoff, Plea**	**52,702.01**	**68.19**	**1**	**<0.0001**
6	Base, Numoff, Plea, Prior	52,700.15	1.86	1	0.1726

Table 8.3: Stage two multilevel model building (testing for additional legal covariates): Goodness of fit results II

Model	Covariates	Deviance	Difference in deviance	Difference in degrees of freedom	p-value
5	*Base, Numoff, Plea*	*52,702.01*			
7	**Base, Numoff, Plea, Court type (Court)**	**50,310.37**	**2391.60**	**1**	**<0.0001**
8	Base, Numoff, Plea, Weapon (Weap)	52,689.35	12.65	1	0.0004
9	Base, Numoff, Plea, Direct victim (Dir)	52,685.60	16.41	1	<0.0001
10	Base, Numoff, Plea, Court, Dir	50,305.71	4.66	1	0.0308
11	Base, Numoff, Plea, Court, Weap	50,310.34	0.03	1	0.8536
12	Base, Numoff, Plea, Court, Dir, Weap	50,05.63	0.08	1	0.7711

The second stage of the analysis built on the analysis conducted in stage one. Stage two began with the best model (Model 5) from Table 8.2, hence it is repeated in Table 8.3. At this stage I attempted to control for three additional legal covariates; the type of court the offender was sentenced in, whether the offence involved the use of a weapon and whether there was a direct victim.

Model 7 controlled for the type of court the offender was sentenced in, which reduced the deviance by 2,391.6 on 1 degree of freedom (compared to Model 5). This yielded a highly significant p-value (p-value <0.0001) which suggested there was sufficient evidence to include this additional covariate in the model. Furthermore, Models 8 and 9 also yielded significant p-values (p-value <0.001), suggesting there was also evidence to keep 'Weap' and 'Dir' in the model. Combinations of the three covariates were then also fitted (Models 10-12). Model 9 controlled for court and direct victim variables, which reduced the deviance in the model by 4.66 on 1 degree of freedom (compared with model 7). The resultant p-value was 0.0308 and not significant at the 1% level, therefore 'dir' should not be included in the model. Model 11 controlled for use of a weapon as well as court, adding this variable reduced the deviance by less than 1 on 1 degree of freedom (compared with Model 7). The result was not significant (p-value 0.8536) indicating there was not enough evidence to include 'weap'. Just as a precaution Model 12 then included both these variables - direct victim and use of a weapon - and reduced model deviance by less than 1 on 1 degree of freedom (compared with Model 10). Again the non-significant p-value of 0.7711 indicated there was insufficient evidence to warrant the inclusion of these additional variables. Therefore once we control for the type of court the offender was sentenced in, the additional covariates do not improve the model enough to warrant their inclusion in the model.

On reflection this result is not a complete surprise as we would anticipate offences which involve the use of a weapon and/or involve a direct victim will automatically

increase the seriousness of the offence. In which case, the offender would likely be sent to the Crown court for sentencing, and therefore by controlling for the type of court the offender was sentenced, this is enough to adequately explain what we are witnessing in the data.

Highlighted in bold, Model 7, controlling for offence type, the number of offences the offender was sentenced for, offence plea and type of court the offender was sentenced in proves to be the optimal model fit. The results from this model are reported in the following chapter.

The third stage in the model fitting process attempted to control for the extra-legal covariates of the case. The additional seven covariates were added individually to the model and compared with the nested model (Model 7) to assess their respective deviance, as shown in the first half of Table 8.4.

In the second half of Table 8.4 combinations of the covariates were then added to the model in descending order of their deviance. For example, from the first part of Table 8.4 we can see Model 16 - offender's employment status - reduced the deviance by 44.99 on 4 degrees of freedom compared with Model 8 (p-value <0.0001) and so this was the first covariate to be included. This was followed by income, accommodation status and so on to produce the models in the second part of Table 8.4. As each additional covariate was added to the model, the deviance was compared with the previous nested model. The last model to be fitted (Model 26) attempted to control for effect of drug misuse on sentence severity. Comparing this model (Model 26) to Model 24 saw a reduction in deviance of 3.66 on 1 degree resulting in a p-value of 0.5580, indicating there was not enough evidence to suggest this covariate should be included in the model.

Consequently, from the third stage of this multilevel analysis, Model 24 is the optimal model controlling for; offence type, the number of offences the offender was sentenced for, offence plea and type of court the offender was sentenced in, the employ-

ment, accommodation and income status of the offender, the offender's sex, ethnicity and age. One small caveat needs to be mentioned in relation to the inclusion of age in the final model. At the beginning of this section it was stipulated that only covariates that were significant at the 1% level of significance would be included in the models. Table 8.4 shows that including age in the model reduces the model deviance by 12.42 on 4 degrees of freedom, and the resulting p-value was 0.0145. Therefore age was only just not significant at the 1% level. However, as will become clear in the following chapter, there are significant differences in the effects of age on sentence severity and therefore it was decided this covariate should remain in the model.

Two further models were fitted as an extension of Model 24 to test for interaction effects and can be seen at the bottom of Table 8.4. The first model (Model 24a) added an interaction term between sex and ethnicity. Comparing the goodness of fit with Model 24 reduced the deviance by 1.23 on 5 degrees of freedom. This yielded a p-value of 0.942, indicating the additional parameter did not improve the model. The second model (Model 24b) added an interaction term between sex and age. Comparing the goodness of fit with Model 24 reduced the model deviance by 2.91 on 4 degrees of freedom. The p-value of 0.574 indicated the additional parameter (interaction between sex and age) did not significantly improve model fit. Therefore Model 24 remains the optimal model. The results from Model 24 are reported in the following chapter

The last stage of this multilevel analysis extended the optimal two level model (Model 24) to three levels - offence (level 1), offender (level 2), Probation Trust (level 3) to fit Model 27. This additional model is shown in Table 8.5. Adding this extra 'random effect' reduced the deviance by 2.6 on 1 degree of freedom (compared to Model 24), which resulted in a p-value of 0.8931. There was consequently insufficient evidence to warrant the inclusion of this additional parameter.

This result indicates there is little evidence of sentencing variation at the Probation Trust level, which is a positive outcome for the criminal justice system. Of course, this

does not rule out the fact that there may be variation within, as well as, between courts within the Probation Trust which cancels itself out - some courts may sentence some offenders more severely whilst others sentence offenders less severely, and therefore this disguises any disparate sentencing. This was observed in Hood's (1992) study which looked at disparity based on the ethnicity of offenders in the West-Midlands. Although, this is not something that can be tested within this research due to court level information being unavailable.

Model 24 remains the optimal model.

Table 8.4: Stage three multilevel model building (extra-legal covariates): Goodness of fit results III

Model	Covariates	Deviance	Difference in deviance	Difference in degrees of freedom	p-value
7	*Base, Numoff, Plea, Court*	*50,310.37*			
13	Base, Numoff, Plea, Court, Sex	50,287.23	23.15	1	<0.0001
14	Base, Numoff, Plea, Court, Ethnicity (Eth)	50,291.46	18.92	5	0.0020
15	Base, Numoff, Plea, Court, Age(cat)	50,302.87	7.51	4	0.1114
16	Base, Numoff, Plea, Court, Employment (Emp)	50,265.39	44.99	4	<0.0001
17	Base, Numoff, Plea, Court, Accommodation (Acc)	50,269.58	40.79	3	<0.0001
18	Base, Numoff, Plea, Court, Income (Inc)	50,269.48	40.87	4	<0.0001
19	Base, Numoff, Plea, Court, Drug misuse (Drugs)	50,303.21	7.16	1	0.0075
16	*Base, Numoff, Plea, Court, Emp*	*50,265.39*			
20	Base, Numoff, Plea, Court, Emp, Acc	50,232.41	32.98	3	<0.0001
21	Base, Numoff, Plea, Court, Emp, Acc, Inc	50,190.00	42.40	4	<0.0001
22	Base, Numoff, Plea, Court, Emp, Acc, Inc, Sex	50,169.53	20.47	1	<0.0001
23	Base, Numoff, Plea, Court, Emp, Acc, Inc, Sex, Eth	50,152.79	16.75	5	<0.01
24	**Base, Numoff, Plea, Court, Emp, Acc, Inc, Sex, Eth, Age**	**50,140.36**	**12.42**	**4**	**0.0145**
25	Base, Numoff, Plea, Court, Emp, Acc, Inc, Sex, Eth, Drugs	50,150.78	2.01	1	0.1560
26	Base, Numoff, Plea, Court, Emp, Acc, Inc, Sex, Eth, Age, Drugs	50,136.70	3.66	1	0.5580
24a	Base, Numoff, Plea, Court, Emp, Acc, Inc, Sex, Eth, Age, Sex*Eth	50,139.13	1.23	5	0.942
24b	Base, Numoff, Plea, Court, Emp, Acc, Inc, Sex, Eth, Age, Sex*Age	50,137.45	2.91	4	0.574

Table 8.5: Stage four multilevel model building (extending the model to three levels): Goodness of fit results IIII

Model	Covariates	Deviance	Difference in deviance	Difference in degrees of freedom	p-value
24	***Base, Numoff, Plea, Court, Emp, Acc, Inc, Sex, Eth, Age***	*50,140.36*			
27	Three levels (Probation Trust): Base, Numoff, Plea, Court, Emp, Acc, Inc, Sex, Eth, Age	50,137.76	2.60	1	0.8931

8.6.1 Confidence intervals

Once the models were fitted, 95% likelihood-based confidence intervals were computed for each of the levels within each of the categorical covariates and for the random effects or variance parameters. This was to establish whether there were any statistically significant differences between the reference category and the additional levels of the covariate, and to assess the accuracy of the variance estimates. For example, as I will illustrate in Chapter 9, the results reveal there is sufficient evidence at the 5% level to suggest that offenders who pleaded guilty (reference category) were sentenced less severely than offenders who plead not guilty, when controlling for a variety of legal and non-legal characteristics. Confidence intervals were estimated using the profile likelihood and confint Bates et al. (2014) commands in R.

The lme4 package does not provide p-values associated with the parameter estimates from the model output. The authors admit to not knowing how to calculate exact values, as well as being unsure how best to approximate the degrees of freedom in a mixed model framework (Chang, L, 2010). It is argued "simple formulas for the degrees of freedom for inferences based on t or F-distributions do not apply in such cases. In fact, the pivotal quantities for such hypothesis tests do not even have t or F-distributions in such cases so trying to determine the "correct" value of degrees of freedom to apply is meaningless" (Baayen et al., 2008, p. 396). For example, the MIXED procedure in SAS offers 6 different calculations of degrees of freedom which lead to different p-values but none of them are "correct" (ibid). Therefore the authors have chosen to abstain from providing p-values until they have developed a more accurate method to do this (Chang, L, 2010).

8.6.2 Log-multiplicative effects on sentence severity

Modelling the natural logarithm of the dependent variable results in a multilevel multiple linear regression with covariates that have a linear relationship with the de-

pendent variable on a log scale. The resulting parameter estimates can be interpreted on a log scale additively or alternatively the exponent of the parameter estimate will be interpreted multiplicatively. Aitkin et al. (2009, p. 125) explain that although a "model is fitted on a log scale, we usually want to interpret the model on the original scale: if fitted values from the model are transformed by *exp(fitted(model.object))* to the original scale, these are fitted values for the median response, not for the mean". Therefore, here I took the exponent of each parameter estimate and interpreted these multiplicatively. For example, if we take a result of 1.15, this would suggest that an offender is sentenced 15% more severely than the baseline or reference category. This then allows me to assess the effect that each parameter has on sentence severity in terms of the change in percent that each parameter has compared to the reference or baseline category. If for example, a not guilty plea is shown to significantly effect sentence severity (which will be confirmed by the 95% confidence intervals for the estimate), the result may indicate that an offender who pleaded not guilty is sentenced 10% more severely than a like-situated offender who pleaded guilty. Had the exponent of this estimate not been taken, the results would be interpreted in terms of the difference in days. This is not sensible due to the vast distribution of sentences.

8.7 Concluding remarks

This chapter has discussed the statistical method of multilevel modelling and in particular the random intercept model, and how it is used to measure sentencing disparity for this doctoral research. This research differs from the previous studies conducted in the area because the scale was derived from statistically modelling real sentencing data obtained from the courts to create a sentence severity scale.

To conduct this analysis the log-transformed final scores - equivalent days in custody - obtained from the extended Goodman RC association analysis were used. The

transformation was then used as the dependent variable with a number of legal and extra-legal independent variables in the multilevel model. A random intercept model was fitted to the data using the R statistics package. The stages of the model fitting were then reported in this chapter.

The following chapter - Chapter 9 - provides the results of the multilevel modelling analysis. Chapter 10 then facilitates a discussion surrounding these results and how they fit in to the wider context of sentencing disparity.

Chapter 9

Results of multilevel modelling

9.1   Introduction

This chapter will report the results from the multilevel modelling analysis conducted in the R statistics package using lme4 as described in the previous chapter. I will first report the null or variance components model then go on to report the additional random intercept models that were explained in the previous chapter.

9.2   Variance components model

Table 9.1 displays the results from the fitted variance components Null Model. There are no covariates in this model - just the intercept parameter. Table 9.1 displays the estimated intercept coefficient and the 95% confidence intervals below this in brackets (column 2), and multiplicative effect of the coefficient (column 3).

   The intercept provides an estimate for the average sentence (severity) score for the offenders in the model. By calculating the multiplicative effect of the intercept coefficient (4.699) (as discussed in section 8.6), this provided us with the average number of days in custody within the sample, which in this case was approximately 110 days. Although this information is good to know it does not assess whether or

not there is any sentencing disparity.

Table 9.1: Null Model

Parameter	Coefficient (CI* 95%)	Multiplicative effect	Model deviance
Intercept	4.699	109.84	64, 504.78
	(4.6826, 4.7155)		

*Likelihood-based Confidence Intervals (CI 95%)

Table 9.2: Null Model - random effects

Random effects	Variance σ^2	Confidence intervals (95%)
σ_u^2 Level-two variance	0.8965	0.8839, 0.9094
σ_e^2 Level-one variance	0.5003	0.4944, 0.5062

In section 8.2.1 I explained that the variance components model usually provides a starting point for any multilevel analysis. This allows the researcher to ascertain the initial variance within the data (before any explanatory variables are added to the model) using the intraclass correlation coefficient (ICC). Table 9.2 displays the random effects from the null model, which are used to calculate the ICC. The confidence intervals for the estimate is also shown to demonstrate the accuracy of the estimates.

In this case the ICC was calculated at 0.6418 using the formula $\frac{\sigma_u^2}{(\sigma_u^2+\sigma_e^2)}$, therefore 64% of the sentencing variation occurs between offenders. This is a very high

result indicating that there is considerable variation between offenders (level-2 units). However, this is not a surprising result as we know from Chapter 5 that there are 99 different offence categories ranging from murder and manslaughter to drunk and disorderly and public order offences, and consequently 35 sentence categories ranging from conditional discharges to prison sentences of over 10 years. Therefore, without taking at least the offence type into account - not to mention the other legal factors of the case - we would expect to see this result.

The remainder of this chapter will now discuss the additional models fitted using the random intercept model, which controls for explanatory variables in an attempt to account for this sentencing variation between offenders.

9.3 Random intercept models

This section will report on the three models which were explained in Section 8.6 as being the optimal models. The first model (Model 5) contains the legal variables explicitly referred to in the sentencing legislation which should be taken into consideration when sentencing offenders in a court of law. The second model that is discussed here is Model 7, which controlled for the type of court that the offender was sentenced in, as well as the legal characteristics of the case. The third model (Model 21) builds on the previous model, taking into consideration a number of extra-legal characteristics of the offender. These three models will be discussed in turn. The results from each of the three models are displayed in the relevant tables within each section. Any result which is different and consequently statistically significant (at the 5% level) from the reference category is highlighted in italics, based on the confidence intervals obtained from the profile and confint commands in R, as also discussed in Chapter 8.

9.3.1 Model 5 - controlling for the legal characteristics of the offence

Before any investigation into sentencing disparity can be conducted it is imperative to control for the legal characteristics of the case. The principle of proportionality places great emphasis on the severity of the sentence being proportional to the seriousness of the offence. The type of offence that the offender is being sentenced for must be taken into consideration in any analysis into sentencing variation. On top of this, as discussed in Chapter 2, sentencing legislation sets out a legal framework within which judges and magistrates must work. Offender plea is one legal factor which judges should take into consideration when sentencing offenders: a discount for a guilty plea is recommended. Relevant previous convictions should also be taken into consideration as an aggravating factor thereby adding to the seriousness of the offending behaviour. Unfortunately this information is not available in this data and therefore an indicator of previous offending was used in its place, that is, prior court appearance, although this was not found to improve the accuracy of the model. In addition, where the offender is being sentenced for multiple offences, sentencing legislation states that judges should apply the principle of totality. This means that judges should sentence proportionately, taking into account all the offences so not to apply an excessive sentence, which is common in certain states of the U.S, such as a 100 year prison term. It must be made clear here that it was not possible to account for every legal factor (i.e. all aggravating and mitigating factors listed in the legislation) due to the availability of information in the data used in this research.

In Section 8.6 I pointed out that the variable relating to prior court appearance did not significantly improve the model fit, and so it was not included in the optimal model. Although this is a surprising result, as research by Ulmer and Johnson (2004) and Britt (2009) for example, found previous offending to be highly significant. In this research it is likely that the covariate used to indicate prior offending was not

an accurate measure of prior offending. Had the number of previous convictions been available in the data, this would have been included in the model.

Three covariates were included in Model 5 - offence type, offence plea, and the number of offences that the offender was being sentenced for - and the results are displayed in Table 9.3. The first column of Table 9.3 displays the covariates included in the model, column two is the parameter or category level within the covariate, column three provides the estimate, and column four displays the multiplicative effect of the estimate for each parameter.

The results provided in Table 9.3 and subsequent tables relate to the average sentence an offender is sentenced to for the offence of burglary in a dwelling and where the offender was sentenced for a one or two offences and pleaded guilty to the offence. These estimates are the reference or baseline categories. As can be seen from Table 9.3, the average sentence an offender received for an offender sentenced for burglary in a dwelling (sentenced for less than two offences and pleaded guilty to this offence) was approximately 329 days in custody.

One point to mention, in Table 9.3 and subsequent tables in this chapter, the parameters relating to the offence type have not been included, only the estimate for burglary in a dwelling. This because there are 99 offence parameters and therefore too many to display in this format.

Table 9.3: Model 5 results

Covariate	Parameter/level	Coefficient (CI 95%)	Multiplicative effect	Model deviance
Intercept		8.3013		52,702.01
		(4.6826, 4.7155)		
Offences	Burglary in a dwelling	-2.5069	328.46	
		(-4.0089, -1.0049)		
Number of offences	<2 offences (reference)			
	>3 offences	*0.1460*	*1.1572*	
		(0.1177, 0.1742)		
Plea	Guilty (reference)			
	Not guilty	*0.1778*	*1.1946*	
		(0.1356, 0.2199)		

[*]Note - parameters in italics are significantly different from the reference category based on the 95% likelihood-based CI

The results from this model (Model 5) reveal offenders sentenced for three or more offences are sentenced on average 16% (1.1572) more severely than like-situated offenders that were sentenced for two or fewer offences. This result is statistically significant as can be seen from the 95% confidence intervals displayed below the estimate in brackets and which do not include 1. It would appear reasonable and also proportionate that offenders who are sentenced for multiple offences are sentenced more severely than offenders sentenced for fewer than two offences and this would be in line with the principle of totality.

Offenders who pleaded not guilty but are then found guilty of an offence, will not benefit from the guilty plea discount, As a result, offenders can expect to be sentenced 19% (1.1946) more severely than a like-situated offender pleading guilty to their offence, when controlling for the offence type, the number of offences they are being sentenced for and their offence plea.

Table 9.4: Model 5 - random effects

Random effects	Variance	Confidence intervals
	σ^2	(95%)
σ_u^2 Level-two variance	0.6239	0.6138, 0.6342
σ_e^2 Level-one variance	0.4433	0.4381, 0.4487

Recall previously the ICC for the variance components Null Model was 64%. Table 9.4 provides the estimates for the random effects for Model 5. The ICC for Model 5 - controlling for the type of offence, the number of offences being sentenced and the offence plea - is reduced to 0.5846 or 58%. Therefore these three variables included in Model 5 account for around 6% of the sentencing variation between offenders. This sentencing variation is still incredibly high and therefore needs further investigation.

9.3.2 Model 7 - controlling for the legal characteristics of the offence and the court effect

As discussed in Chapter 2, sentencing legislation under the Powers of the Criminal Court (Sentencing) Act 2000 limits the sentencing powers of the magistrates' courts. Therefore we would expect to see a substantial difference in the sentences issued by the Crown and magistrates' courts. Model 7 controls for this, as well as offence type, number of offences the offender was being sentenced for and offence plea.

Table 9.5 provides the parameter estimate for Model 7 which shows that for the offence of burglary in a dwelling the average sentence is approximately 408 days, which is higher than that reported in Model 5. This is due to the fact we are now controlling for the type of court the offence was sentenced in. The Crown Court is now the reference category for Model 7 and subsequent models. Controlling for the type of court the offence was sentenced in indicates the increased level of offence seriousness and this will reflect the characteristics of the offence, i.e the value of the goods stolen during the burglary.

On the contrary, and as we would have expected, offenders sentenced in the magistrates' courts are sentenced less severely than an offender in the Crown Court. The model predicts that offenders sentenced in the magistrates' court rather than the Crown Court would be sentenced 51% less severely. This is a bigger reduction that would have perhaps been anticipated but this may be down to the fact that there are additional factors that we have not been able to account for due to the availability of that information.

Controlling for the type of court the offender was sentenced in, also alters the other parameter estimates in the model. As can be seen from Table 9.5, offenders sentenced for more than three offences would expect to see a 13% (1.1272) increase in the severity of their sentence, as opposed to being sentenced for one or two offences. This is lower than previously estimated in Model 5, where the estimate was 16%. There

could be various reasons for this, one reason being, as we know offences sentenced at the Crown Court are more serious than those offences sentenced at the magistrates' court, therefore once the seriousness of the offences (where there are multiple offences) reaches a certain threshold, so too does the sentence severity. A way in which judges may then deal with this is after a certain point they begin to constrain the overall severity of the sentence to ensure they are issuing a proportionate sentence within the realms of totality. Of course there may be other explanations for this result, which I have not been able to control for in this research.

Model 7 also controlled for offence plea, as did model 5. The effect of entering a not guilty plea and being found guilty still has a positive, significant effect on sentence severity. Having controlled for the additional parameter (type of court) this effect is slightly less. Offenders pleading not guilty as opposed to guilty can expect a sentence in the region of 14% more severe. This is slightly less that estimated in the previous model.

The random effects estimates are displayed in Table 9.6. The ICC for Model 7 is now 0.5658, which means there is still 56.58% variation at the offender level after controlling for offence type, number of offences being sentenced, plea and the type of court the offender was sentenced in. This combination of variables only explain 8% of the variation in sentencing.

Table 9.5: Model 7 results

Covariate	Parameter/level	Coefficient (CI 95%)	Multiplicative effect	Model deviance
Intercept		8.3450		50,310.37
		(6.9387, 9.7512)		
Offences	Burglary in a dwelling	-2.3328	408.38	
		(-3.7400, -0.9256)		
Number of offences	<2 offences (reference)			
	>3 offences	*0.1198*	*1.1272*	
		(0.0935, 0.1460)		
Plea	Guilty (reference)			
	Not guilty	*0.1345*	*1.1440*	
		(0.0946, 0.1745)		
Court	Crown (reference)			
	Magistrates'	*-0.6956*	*0.4988*	
		(-0.7226, 0.6686)		

[*]Note - parameters in italics are significantly different from the reference category based on the 95% likelihood-based CI

Table 9.6: Model 7 - random effects

Random effects	Variance σ^2	Confidence intervals (95%)
σ^2_u Level-two variance	0.5687	0.5593, 0.5783
σ^2_e Level-one variance	0.4365	0.4314, 0.4417

9.3.3 Model 24 - also controlling for the extra-legal characteristics of the offender

Model 24 extended the previous models, by controlling for numerous extra-legal variables which according to various sentencing legislation should not be taken into consideration when sentencing offenders but are included in the actuarial risk of recidivism calculation carried out by the OASys assessment. The results are shown in Table 9.7.

Table 9.7: Model 24 results

Covariate	Parameter/level	Coefficient (CI 95%)	Multiplicative effect	Model deviance
Intercept		8.1703		50,140.36
		(6.7688, 9.5719)		
Offence (See Appendix F)	Burglary in a dwelling	-2.3446	338.90	
		(-3.7466, -0.9425)		
Number of offences	<2 offences (reference)			
	>3 offences	0.1165	1.1235	
		(0.0905, 0.1428)		
Plea	Guilty (reference)			
	Not guilty	0.1357	1.1553	
		(0.0958, 0.1758)		
Court	Crown (reference)			
	Magistrates'	-0.6913	0.5005	

Continued on next page

Table 9.7-*continued from previous page*

Covariate	Parameter/level	Coefficient	Multiplicative effect
		(-0.7180, -0.6639)	
Employment status	Employed full/part-time (reference)		
	Temp/Casual work	*0.0885*	*1.0925*
		(0.0135, 0.1637)	
	Education/Training	0.0761	1.0791
		(-0.0201, 0.1725)	
	Unemployed	*0.1707*	*1.1862*
		(0.1211, 0.2214)	
	Unavailable for work	*0.1279*	*1.1364*
		(0.0696, 0.1830)	
Accommodation status	Permanent (reference)		
	Probation Housing	*-0.0823*	*0.9210*
		(-0.1462, -0.0190)	
	Transient/Short-term	-0.0051	0.9949
		(-0.0368, 0.0271)	

Continued on next page

Table 9.7-*continued from previous page*

Covariate	Parameter/level	Coefficient	Multiplicative effect
	No fixed abode	*0.1038*	*1.1094*
		(0.0546, 0.1522)	
Income status	Wages (reference)		
	Pension/Maintenance	*-0.1413*	*0.8682*
		(-0.2729, -0.0091)	
	State benefits	*-0.1102*	*0.8956*
		(-0.1591, -0.0620)	
	No income	-0.0292	0.9713
		(-0.0953, 0.0310)	
	Other	0.0426	1.07435
		(-0.0318, 0.1184)	
Sex	Female (reference)		
	Male	*0.0852*	*1.0890*
		(0.0491, 0.1208)	
Ethnicity	White UK & Irish (reference)		

Table 9.7-*continued from previous page*

Covariate	Parameter/level	Coefficient	Multiplicative effect
	White other	*-0.0772*	*0.9257*
		(-0.1362, -0.0178)	
	Mixed	-0.0075	0.9925
		(-0.0738, 0.0595)	
	Asian	-0.0203	0.9799
		(-0.0723, 0.0230)	
	Black	*0.0439*	*1.0449*
		(0.0010, 0.0869)	
	Other	*0.1214*	*1.1291*
		(0.0061, 0.2383)	
Age	21-24 (reference)		
	25-29	*0.0396*	*1.0404*
		(0.0056, 0.0731)	
	30-40	*0.0496*	*1.0508*
		(0.0180, 0.0808)	

Table 9.7-*continued from previous page*

Covariate	Parameter/level	Coefficient	Multiplicative effect
	41-54	*0.0486*	*1.0498*
		(0.0121, 0.0849)	
	55+	*0.0819*	*1.0853*
		(0.0087, 0.1565)	

[*]Note - parameters in italics are significantly different from the reference category based on the 95% likelihood based CI

Model 24 now controls for 10 variables: 4 legal covariates and 6 extra-legal covariates. Using the random intercept model equation in Section 8.2.2 and the estimates from the multilevel model shown in Table 9.7 and Appendix F (offence category estimates), we can estimate the expected sentence of an offender convicted of burglary in a dwelling (major offence code 28) in the Crown Court (reference category), at this sentencing occasion they were sentenced for one or two offences (reference category) and pleaded guilty (reference category) to their offence. At the time of sentencing the offender was in full-time employment (reference category), and therefore relying on their own income (reference category) through working, and lived in permanent accommodation (reference category). The offender was female (reference category) with White British or Irish ethnicity (reference category) and aged between 21 and 24 (reference category). The average sentence for burglary in a dwelling controlling for the legal and extra-legal covariates in Model 24 was therefore estimated at 339 days in custody.

Alternatively we can estimate the average number of days in custody for any of the offence categories using the coefficients shown in Appendix F. For example, we could estimate the average sentence for the offence of common assault to be 110 days (8.3450 (intercept) - 3.4663 (coefficient for common assault)) controlling for the same legal and extra-legal covariates in Model 24. This sentence is consistent with the sentencing guidelines for common assault set by Sentencing Council (2011), and which ranges between a fine and conditional discharge at the lowest level of seriousness (category 3) to, a medium level community order (category 2), and for the most serious offence, a community order or custodial sentence (top range of 26 weeks in custody) (category 1). In terms of disparity, the same (multiplicative) effects apply. For example, if this offender also plead not guilty they would get a sentence 16% (1.1553 multiplicative effect) higher which is approximately an additional week in custody. If the offender were also unemployed at sentencing they would also expect to receive a sentence just

under 20% (1.1862 multiplicative effect) more severe. In terms of days, that equates to approximately 10 additional days in custody.

The effects of the legal covariates after controlling for the extra-legal characteristics of the offender do not appear to change much from those estimated in Model 7.

As can be seen from Table 9.7, a not guilty plea compared to a guilty plea will see an increase in the severity of the sentence by approximately 16%. We already know that offenders will receive a discount on their sentence for pleading guilty, and therefore this result is commensurate with the sentencing legislation.

Offenders charged with multiple offences (three or more) are sentenced just over 12% more severely than offenders who are sentenced for one or two offences. This is similar to the estimate from the previous model. These results indicate that the principle of totality is being adhered to by the courts.

These results would suggest that once the relevant legal factors are taken into consideration there is agreement between what the guidelines suggest and what is happening in the courts. In some way these findings support the work of Pina-Sánchez and Linacre (2013) who found there to be 'consistency' in the sentencing of assault cases. Although as I have highlighted previously, the work by Pina-Sánchez and Linacre (2013) fails to investigate sentencing disparity by not considering the extra-legal characteristics of the case.

That being said, Model 24 also controlled for numerous extra-legal characteristics of the offender. I will discuss these results in turn.

Employment status

Employment status appears to be one of the biggest contributors in explaining sentencing disparity. The result of the multilevel modelling provides significant evidence to suggest that when controlling for the additional covariates in the model, offenders who are unemployed prior to sentencing are predicted to get a sentence that is almost

19% more severe than a like situated offender that was in full-time employment prior to sentencing. This result is highly significant. This is consistent with the research conducted by Flood-Page and Mackie (1998), Moxon (1988) and Crow and Simon (1989): all three studies reported that unemployed offenders were less likely to be fined and more likely to receive a more severe sentence.

Additionally, those who were unavailable for work prior to sentencing were also sentenced approximately 14% more severely than those who were employed prior to sentencing, and this result is significant. Being unavailable for work may be viewed similarly to being unemployed, and therefore it is the lack of being in paid employed that is the deciding factor. Although interestingly, offenders in causal or temporary employment prior to sentencing were also sentenced more severely than those in employment - 9% more severely (the result is also significant). Even though these offenders may have some work, or go through periods of employment, this does not appear to be seen as any kind of mitigation. Previous research tends to model employment as a binary outcome - employed/unemployed - and does not differentiate between the types of employment status, and so it is not possible to compare results.

Finally, offenders who were in education (whether full-time or part-time) prior to sentencing appear to be sentenced more severely than those in employment but this difference is not statistically significant.

Accommodation status

Living in permanent accommodation is the baseline or reference category for this covariate. The results of the multilevel modelling suggest that offenders recorded as having no fixed abode are sentenced just under 11% more severely than the reference category.

In spite of this, offenders that reported as living in probation or supported housing are sentenced 8% less severely than those in permanent accommodation, whilst

controlling for the other variables in the model.

There appears to be no significant difference in the sentence severity for those living in transient or short-term accommodation as opposed to permanent accommodation.

Income status

Income status is generally linked with employment but there are some interesting and conflicting results. Namely, those offenders who were reliant on state benefit prior to sentencing are sentenced 10% less severely than offenders who are reliant on their own or partner's wage. This seems to contradict the previous results which suggest unemployed offenders are sentenced more severely than those in employment. Those offenders who were in receipt of a pension or maintenance were also sentenced less severely, some 13% less severely, and this was also a significant result. There were some differences between those claiming they had no income, as well as those who said they had other means of income, but these categories were not significantly different from the baseline category.

Sex

The results of this analysis suggested that whilst controlling for offence type and the additional legal and extra-legal covariates included in Model 24, men are sentenced approximately 9% more severely than like-situated women offenders. This finding is also consistent with previous research, which has generally found that higher proportions of men are sentenced to immediate custody (Albonetti, 1997), and where women offenders do receive a custodial sentence, they are usually shorter in length (Albonetti, 1997); (Stacey and Spohn, 2006). Using the new scale, we can now be confident in concluding that men are sentenced more severely than women across the full distribution of sentences, controlling for offence type and the legal characteristics of the case.

Ethnicity

In this instance, the baseline or reference category refers to the 'White British and Irish' ethnic group. Comparing sentence severity with the 'White other' group reveals this group are sentenced 7% (0.9257) less severely than the reference group. Consistent with previous research (e.g. Hood, 1992), black offenders were sentenced 5% more severely than the 'White British and Irish' offender group. This result is statistically significant (p=<0.001). The group classified as 'Other' shows some significant differences in sentence severity. This group are sentenced just under 13% more severely than the 'White British and Irish' group. There was very little significant difference in sentence severity between the reference category and the Asian and Mixed ethnic categories.

Age

The results from the multilevel modelling suggest that age plays a significant role in sentencing disparity: as offenders get older they are sentenced more severely. offenders aged 25-29 can be expected to be sentenced 4% more severely, offenders in the age categories 30-40 and 41-54 can be expected to be sentenced 5% more severely and those over the age of 55 are sentenced just under 9% more severely.

ICC

The random effects estimates for Model 24 are displayed in Table 9.8, and are used to calculate the ICC.

Table 9.8: Model 24 - random effects

Random effects	Variance	Confidence intervals
	σ^2	2.5%, 97.5%

σ_u^2 Level-two variance	0.5646	0.5551, 0.5742
σ_e^2 Level-one variance	0.4362	0.4311, 0.4414

The ICC for Model 24 is 0.5641 or 56%, which means that even after controlling for the legal characteristics of the case and a number of extra-legal characteristics of the offender, there is still considerable variation in sentencing. The covariates included in this model only account for almost 9% of the sentencing variation between offenders compared with the Null Model. There is a considerable amount of sentencing variation that cannot be explained by the covariates included in this model and further research into this needs to be conducted to adequately address this disparity.

9.3.4 Non-significant covariates

A number of aggravating factors were included in the analysis but did not appear to have a significant effect in modelling sentence severity. At the offence level, the use of a weapon did not appear to significantly affect sentence severity. It is possible this factor is already accounted for in the type of offence in which the offender is being sentenced for and therefore not necessary to have an additional variable for this but it is a surprising result.

Similarly, having a direct victim did not add to the accuracy of the model. Again this is surprising as these two factors would increase the seriousness of the offence. This again leads me to question the accuracy of the data, as it is possible this information has not been recorded correctly in OASys.

At the offender level drug misuse was not found to be a significant factor affecting sentence severity. This is consistent with research by Spohn and Belenko (2013), who found that neither the offender's history of drug use nor their use of drugs at the time

they committed their offence influenced the sentencing decision.

9.4 Concluding remarks

This chapter has provided the result from the multilevel modelling which investigated sentencing disparity using the new measure of sentence severity - the sentence severity scale. The following chapter will interpret these results and provide a full discussion on what these results suggest and the implications these results may have on sentencing policy.

Chapter 10

Discussion and conclusions

10.1 Introduction

This doctoral research set out to achieve two broad aims. The first of these was to create a more accurate measure of sentence severity which incorporated the full distribution of sentences issued to offenders in the criminal courts of England and Wales for a variety of offences. An extended form of the Goodman RC Association analysis was used to measure the association between the offence and sentence categories using real sentencing data obtained from the Police National Computer (PNC) following sentencing. This enabled the construction of a scale while controlling for three legal variables of plea, previous history of offending and number of offences the offender was being sentenced for. The second aim was to use this measure of sentence severity to assess disparity in sentencing using data from the Offending Assessment System (OASys), which was matched to the PNC data. This provided data on the offenders socio-demographic characteristics, as well as other information relating to their offences and perceived risk of recidivism and risk of harm through actuarial assessment. Assessing disparity in sentencing involved using multilevel modelling by way of the nested nature of the data, to assess sentencing variation at the offence and offender

level. Variation at the offender level after controlling for the legal characteristics of the case would imply sentencing disparity - as discussed in Chapter 8 and 9.

The results of the multilevel modelling reported in the previous chapter (Chapter 9) identified some sentencing disparity relating to a number of extra-legal factors. Consistent with previous research by Albonetti (1997), Johnson (2006), and Mustard (2001) for example, there is evidence to suggest women are sentenced less severely than 'like-situated' male offenders even after controlling for the type of offence they were sentenced for, as well as additional legal factors such as plea, the total number of offences they were sentenced for (taking totality into consideration), and the type of court they were sentenced in. Ethnicity was another area in which this research identified subtle disparities, with Black offenders being sentenced more severely than 'like-situated' White British and Irish offenders. Again this is consistent with previous research carried out in the field, for example by Hood (1992), Mustard (2001), and Woodredge (2010). Using the new measure of sentence severity (extended Goodman RC association analysis), we can be more confident that disparity is prevalent across the full distribution of sentences meted out by the courts in England and Wales.

This research went a step further by also controlling for a number of extra-legal factors deemed to be legally irrelevant when sentencing offenders. However, these factors form the basis of the risk assessment undertaken by the Probation Service prior to sentencing. This information is included in the pre-sentence report which is given to the judge as an decision-making aid. Again this research found evidence of significant disparity relating to the employment status of the offender. Unemployed offenders are sentenced more severely than offenders who are employed prior to sentencing. Those with no fixed abode are sentenced more severely than those living in permanent accommodation prior to sentencing [23] . Additionally, when we consider the income

---

[23]Offenders that have no fixed abode, are less likely to be granted bail and therefore are more likely to be held on remand. Offenders held on remand prior to sentencing are then more likely to be sentenced more severely due to fact they were not originally granted bail.

status of offenders or where their income comes from, the research found that offenders who are in receipt of benefits are sentenced less severely than those offenders who were earning a wage through employment prior to sentencing.

Although this research has identified some degree of sentencing disparity, it is acknowledged that only a limited number of variables - both legal and extra-legal, were used in the multilevel model analysis. Furthermore, we were unable to fully account for the offenders' history of offending which plays a key role in determining sentences. If additional data had been available, such as number of previous offences, it is likely that more of the variability at the offender level would have been explained. That being said, the results from this research and levels of disparity are consistent with previous research that was able to control for additional legal variables such as the offender's previous convictions.

This chapter then discusses the contributions that this empirical research has brought to the area of sentencing in terms of developing a measure of sentence severity which encompasses the full distribution of sentences issued by the courts in England and Wales, and in turn allows sentences to be measured in terms of their severity, as well as how different sentences compare with each other. In developing this severity scale, it was then possible to assess sentencing disparity and identify which extra-legal factors are causing judges to sentence offenders more or less severely than others when controlling for additional legal and extra-legal covariates in the model. This chapter then goes on to discuss the implications of the results found in this research and how it might be best to utilise these results to achieving a more just sentencing framework. The final section outlines the possibilities of future work and how with more detailed data, it would be possible to refine not only the sentencing severity scale but also the analysis into sentencing disparity.

10.2 Methodological advances

Chapter 4 discussed the previous ways in which sentence severity has been measured and the problems associated with these methods. For example, it was discussed that by considering both the 'in-out method' and the length of custodial sentence method bias was introduced into these samples by excluding a wealth of information (Sebba, 1978). By only concentrating on sentencing at one end of the severity scale, namely imprisonment, this excludes the non-custodial sentences which are more commonly used and, are just as important when considering sentence severity and with that, disparity (Merrall et al., 2010).

A third way of measuring sentence severity is to use a single severity scale which includes the full distribution of sentences given to offenders in the courts in England and Wales. Again in Chapter 4 I explained the previous methods of constructing these types of scales, and again highlighted their shortcomings in respect to them being highly subjective, based on perceptions of severity, and also using vignette style approaches based on hypothetical examples or instances in which to determine opinions of severity instead of using real examples.

This research has been able to overcome these problems by opting to use a form of statistical modelling - Goodman Row Column Association analysis - to model real sentencing data obtained from the courts in England and Wales. By organising the data into a two-way table, this method models the association between the offence and resultant sentence categories, as discussed in Chapter 6. The Goodman RC model uses the relationship between these two categories to create a scale measuring one dimension of sentencing - severity. I then extended this model to allow for the control of additional legal factors - plea, prior court appearance and totality - to more accurately reflect the factors that judges must take into consideration when sentencing offenders. This corrects for the imbalance or the affect that these variables have on sentence severity.

In doing so, this research creates a sentence severity scale which places sentences into a rank order based on their magnitude in terms of sentence severity in comparison to the other sentence categories. To avoid the scale appearing arbitrary (McDavid and Stipak, 1981), linear interpolation was used to anchor the scale. From here it was then possible to measure disparity over the entire spectrum of sentences.

The same legal covariates used to create the scale are then also included in the multilevel modelling analysis. The extended Goodman RC analysis uses these covariates (plea, prior court appearance and number of offences the offender is being sentenced for) to control for the imbalance these factors have on sentence severity to formulate the scale. The multilevel analysis then allows us to identify the effects that the different legal factors can then have on sentence severity whilst controlling for other characteristics of the offence and the offender. For example, we are able to estimate the effect of pleading not guilty compared to pleading guilty in terms of the difference in the sentence. With offenders who plead not guilty in general serving longer or more severe sentences.

There are a number of potential uses for this work; firstly the sentence severity scale could be used as an aid to judges to identify sentences that are of equivalent severity and also to identify alternatives to some custodial sentences. This idea of interchangeable sentences is discussed in the following section. Additionally, the scale could be used, as it has been in this research, to access sentencing disparity. With the availability of more covariates - both legal and extra-legal - it may be possible to control for more of the unexplained variability in the multilevel models which will contribute to explaining what is happening within the courts of England and Wales. The following section goes on to discuss the idea of interchangeable sentences.

10.3   Interchangeability of sentences

The idea of the interchangeability of sentences has been "relatively under-conceptualised" (Marinos, 2005, p. 441) and yet judges, to a certain extent, use sentences interchangeably, especially in borderline cases (Jacobson and Hough, 2007). Interchangeable sentences allow for alternative or substitute sanctions to be used in place of another sentence of equivalent severity. Von Hirsch et al. (1989) argue this permits considerable flexibility to use non-custodial sentences in place of other non-custodial and custodial sentences. This is particularly beneficial where sentences can be given in place of, for example shorter prison sentences. There are a number of benefits for being able to use sentences interchangeably, not to mention the possibility of using alternatives to custody, which may alleviate the numbers being sent to prison and also reduce costs to the tax payer. The results from the extended Goodman RC Association analysis demonstrate that certain sentences can be considered as alternatives or as equivalent in terms of their levels of severity.

The results for the extended Goodman RC association analysis and linear interpolation demonstrated the equivalent sentence severity scores for each of the sentences categories (see Table 7.2) and following on from that, in Section 7.3.2 I suggested that sentences that are of equivalent severity can be interchangeable. In Section 7.3.2 I provided two examples of sentences which have equivalent sentence severity scores and therefore these sentences can be considered as interchangeable. The first example relates to a community order of less than 12 months (δ_s = -4.028) being roughly equivalent to a suspended sentence order of greater than one month and less than or equal to 2 months (δ_s = -4.007). Following the linear interpolation to anchor these scores in line with the custodial sentences, these two sentences are the equivalent to approximately 65 and 66 days in custody respectively.

Likewise the conditional discharge of less than or equal to one year category (δ_s = -4.115) equates to 63 days in custody following linear interpolation. The score for

immediate custody between 1 and 2 months (based on the modal number of days in the category) is 54 days.

Another example indicates, the sentence category for greater than three and less than or equal to four years immediate custody ($\delta_s = 2.886$) is relatively close to a community order of greater than two years and less than or equal to three years ($\delta_s = 2.821$). In days they equate to 1362 and 1335 respectively (following linear interpolation for the community order category), which again are quite similar.

Previous research by Marinos (2005) specifically investigated 'penal equivalents' or the interchangeability of sentences; namely whether the public viewed it as acceptable to substitute fines and community orders with imprisonment for a variety of offences - small scale theft, assault, and sexual assault. In doing so, the author argues for going beyond a focus of severity by examining how different sanctions are appropriate and vary by the nature of the offence and offender - whether the offender is an adult or a youth.

The results revealed that respondents were more likely to support the interchangeability of a fine or community order for a 30-day custodial sentence for the theft offence than the assault or sexual assault offences for the adult offenders. Additionally, community orders were significantly more likely to be supported as a substitution for a prison sentence than the fine, again particularly for adult offenders. Marinos argues that the support for substituting sentences varies by the offence type and this illustrates a "qualitative dimension of punishment" (Marinos, 2005, p. 446). Furthermore Marinos argues the results suggest the public perceive there to be difference between the two intermediate sanctions (community orders and fines) in terms of their characteristics and qualities, and this affects the nature of their support. Therefore, sanctions are not seen as equally interchangeable because of this.

I would argue, that it is not surprisingly that there is more support for the alternative sentences for the theft offence than there is for the assault and sexual assault

offences. I would argue that this result would suggest that although in exchange for a 30-day custodial sentence, neither a fine nor a community sentence is deemed to be severe enough and given that, it may be argued that in fact a 30-day custodial sentence is also not severe enough for an assault or sexual assault offence. Taking this view, it is unlikely that the public would then support either of the alternative sentences which may be seen as less severe than a longer custodial sentence, which is deemed more appropriate for an assault or sexual assault offence. Marinos acknowledges that the public identify different characteristics and qualities with the two non-custodial sentences, and this affects their support for the interchangeability of sentences, but fails to acknowledge that one of the characteristics or qualities, may in fact be their perception of the severity of those sentences.

Furthermore, this particular study fails to take into consideration retribution as a purpose of punishment and concluded that the results suggest that the support for interchangeability is not related to perceptions about severity, instead people have broader perceptions about punishment. It would seem that Marinos (2005) refuses to acknowledge a retributive approach to sentencing exists, and instead focuses purely on a more utilitarian approach to sentencing. Indeed most of the discussions in relation to this type of research frames itself in either of the two camps. There seems to be less acceptance of a more hybrid type of approach that is currently adopted in England and Wales.

Additionally, one further argument for the use of interchangeable/alternative sentencing is the evidence that suggests offenders recidivate significantly less after community service sentences than after imprisonment (Ministry of Justice, 2013a); (Wermink et al., 2010). Therefore it could be argued that using alternatives to imprisonment (where appropriate) that are of equivalent severity based on the new scale, could improve the likelihood or rate in which offenders reoffend.

10.4 Disparity

The results obtained from this research into disparity are consistent with previous research which found significant differences in the sentencing of offenders. Following the methodological problems that this research has sought to overcome (by incorporating custodial and non-custodial sentences into a single scale of sentence severity by extending the Goodman RC Association model), this research indicates that disparity occurs across all types of sentences e.g discharges, fines, community orders etc. and not just in relation to custodial sentences. Although it was only possible to control for a very small number of the legal and extra-legal variables that are likely to impact on decisions, it does identify some disparities in sentence severity relating the offender's employment status as well as their living arrangements. There was some disparity which may as a result of the extra-legal characteristics of the offender - these will be discussed in turn - but there is still a considerable amount of sentencing variation that could not be explained by the model. This suggests there are additional factors affecting decision-making that need to be explored and included in future research. However, by adopting a new way of modelling the full range of sentences we can begin to explore disparity across the distribution of sentencing and not just for those being sent to prison. Therefore by using the new sentence severity scale, this research can be seen to more accurately reflect sentencing and what is really happening in the courts in England and Wales. However, it is recognised that I have only controlled for a small number of the factors in the multilevel analysis that may ultimately affect sentencing decisions.

The results from this research indicate that there is some sentencing disparity which may lead to discrimination. Some of this is likely to be explained by the extra-legal factors of the offender, such as their gender, ethnicity, age, employment status, income status and accommodation status.

This research has identified some disparity in the sentencing of male and female

offenders once we control for the legal characteristics of the case. This disparity is apparent across the entire distribution of sentences meted out by the courts in England and Wales. Women receive sentences which are on average 9% less severe.

Hedderman and Gelsthorpe (1997) argue that women tend to play minor roles in offences and are therefore sentenced accordingly. I was able to control for the major offence type yet the disparity remained, with women being sentenced more leniently than like-situated male offenders. Although, it is recognised that this disparity could have been somewhat reduced if we had been able to also control for the offenders previous convictions. However, this result is consistent with the research conducted by for example, Spohn (2009) and Albonetti (1997), who both found significant differences in sentence severity between male and female offenders, even when the offence type and other legal characteristics were controlled for, for example previous convictions. It is important to consider that even though the offence type was controlled for, it was not possible to determine (from the information available) whether the female offenders did play lesser roles within the offence. For example, we do not know the specific details of the offence or whether the judge considered the offence to be a category 1, 2 or 3 offence in terms of the harm and culpability caused by the offender which also indicates the seriousness of the offence. Additionally, there is insufficient information relating to the mitigating and aggravating factors associated with the offence. More detailed information relating to the factors that were considered in making sentencing decisions would allow for better modelling of the data. Here data from the Crown Court Survey would be useful as the judge is required to state the significant aggravating and mitigating factors which led to their decision.

The aim of this research is not to enter into the discussion about whether women should or should not be sentenced more leniently than men. However, it is necessary to highlight the different perspectives surrounding the sentencing of female offenders. Many feminists would tend to disagree with the 'Conceptual framework for achieving

justice' model (see Section 3.2.1 Figure 3.1) which would suggest that disparate sentencing would likely result in discrimination and ultimately injustice. Instead they would argue that justice can never be achieved due to the injustice women already suffer within society (for example Corston (2007); Daly (1994)) and therefore an alternative model of justice is necessary (see Figure 3.2). This perhaps raises the question, can we really compare male and female offending? Any research into sentencing disparity perhaps needs to consider male and female sentencing separately.

Consistent with previous research looking at racial disparities, this research also found there to be significant differences between the sentencing of White British and Irish offenders and Black offenders: Black offenders are sentenced on average just over 4% more severely than like-situated White British and Irish offenders. By controlling for the additional legal covariates, we can begin to argue that black offenders are being discriminated against during sentencing. We can also rule out that Black offenders are being sentenced more severely due to the greater number of Black offenders pleading not guilty as this has been controlled for in the model, although that is not to say this is another way in which Black offenders may be indirectly discriminated against at the sentencing stage (Spohn, 2009). Again this result is consistent with previous research. Hood (1992) also found that Black offenders were sentenced 5% more severely that White offenders. These results (parameter estimates) may seem quite small but they are statistically significantly different. This approximately equates to an additional 18 days in custody (almost three weeks) for a one year custodial sentence or an additional 55 days (almost two months) in custody for a three year custodial sentence. Therefore this is not a small effect.

Few studies on ethnic disparity in sentencing have looked at other ethnic groups, the focus tends to be on Black and Asian offenders in the UK (e.g. Hood (1992); Phillips and Brown (1998)), and Black and Hispanic offenders in the US (e.g. Johnson (2006); Albonetti (1997); Engen and Gainey (2000)), compared to the White

indigenous category. In this study, it was found that the White other offenders are sentenced 7% less severely than the White British and Irish offenders, where as the 'Other' offender group are sentenced 13% more severely than the White British and Irish offenders. I am unable to offer an explanation for this result. There was however, a large amount of sentencing variation at the offender level that was not explained by the optimal model - Model 24. This may reduce the size of these effects, if other factors impacting on sentence severity are being masked. However, further investigation needs to be conducted to try and explain why there is such wide sentencing disparity.

Previous research has suggested that Asian offenders are also sentenced disproportionately compared to the White British offenders (Ministry of Justice, 2011). Although this research did not find this, it is important to treat this result with caution. Flood-Page and Mackie (1998) suggests using such broad categories may mask ethnic disparities. For example there may be differences in the way Indian, Pakistani and Bangladeshi offenders are sentenced, and by grouping these offenders together masks the true outcomes. Hood's (1992) earlier research highlighted the potential ways in which disparity may be disguised, and this can be applied to any situation. Therefore future work may benefit from using narrower categories for ethnicity.

The age of the offender at sentencing does not feature heavily in research into sentencing disparity: it remains a rather under-conceptualised factor (Mueller-Johnson and Dhami, 2009). Although, a hand-full of studies do control for age within their models, for example Johnston and Alozie (2001), Johnson (2006), Steffensmeier et al. (1995), and Wheeler et al. (1982). In this doctoral research it was found that sentence severity increases with the age of the offender even after controlling for the offence type and other legal factors. For example, those in the 25-29 category are sentenced 4% more severely than the 21-24 (reference) group, by the time the offender is over the age of 55 years, this doubles, so the 55+ group are sentenced almost 9% more severely than the youngest group. However, this result is not consistent with previous research

that has controlled for age, which have generally reported a small negative linear effect, with older offenders tending to be sentenced more leniently (e.g. (Johnston and Alozie, 2001); Woodredge (2010)). This ties in with widely held assumption that older offenders are sentenced more leniently based on the notion that older offenders are more able to reform themselves but also because they carry less dangerousness and risk (Steffensmeier et al., 1995). Although Steffensmeier et al. (1995) would suggest that the effect is actually curve-linear, with younger and older offenders being sentenced more leniently.

It is important to be wary of the age result found within this research, firstly because it is inconsistent with previous research, but also because we were not able to control for previous offending. It is possibly that age is masking previous offending, as we may observe a linear relationship between age and number of offences - as offenders get older they will likely acquire more offences/convictions. This could explain why there appears to be significant differences between offenders of different ages.

From the results of this research, it would seem that anything less than being in permanent employment increases the severity of a given sentence. This result is consistent with previous studies that have explored the relationship between employment status and sentencing, such as research by Jacobson and Hough (2007); Flood-Page and Mackie (1998); and Moxon (1988). Therefore having steady employment may be seen as a factor which may mitigate the severity of a sentence (Ashworth, 2010b). However, Ashworth (2010b) recognises that to mitigate on the grounds of employment can be seen as discriminating against those who are unemployed.

To add to this, unless this factor is specified in the sentencing legislation or guidelines, some judges may consider employment to be a mitigating factor, where as other judges may rule it out altogether. For example, the potential loss of employment following a custodial sentence, may cause the judge to use this as a grounds for mitigation (Easton and Piper, 2008). Loss of employment is also more likely to be taken

into consideration if it impacts on third parties such a employees (ibid). Additionally, judges may apply different weight to employment status. Again there is also no information relating to how employment status should be dealt with, which potentially leads to disparate sentencing. Unless there is also parity in the approach, the end result may just be more disparate sentencing.

Research suggests that those with no-fixed-abode may be more of a risk of recidivism than those living in permanent housing following incarceration (Social Exclusion Unit, 2002). The Social Exclusion Unit (2002) report that having stable accommodation can reduce reoffending by over 20%. Not only that, having no fixed abode also makes it more difficult to find a job and earn money legitimately (ibid).

Accommodation status may be taken into consideration when granting the offender bail, along with a number of other socio-demographic information, such as their marital status, family ties, and whether they are in employment (and for how long) (Dyer, 2013). Although the legislation (Bail Act 1976) does state that those with no fixed abode are not automatically prohibited bail, in practice is does cause significant difficulties (Dyer, 2013). This may also apply at the sentencing stage, as those who have no fixed abode are viewed as having greater risk of recidivism and therefore they are sentenced more severely which would explain the results found in this research.

The Prison Reform Trust (2013) state that a lack of accommodation can also severely hinder the likelihood of former prisoners finding suitable employment - almost one quarter of employers would not consider hiring a homeless person. Getting ex-prisoners into stable housing can act as an effective gateway to resettlement (Prison Reform Trust, 2013). This may explain why offenders who are in probation/supported housing prior to sentencing are being sentenced less severely. Judges may see this as offenders trying to improve their situation and therefore punish them less severely. Whereas, those with no fixed abode, remain too high a risk and therefore sentence

them more severely. This then becomes a vicious cycle that exacerbates offenders' problems and offending.

The Social Exclusion Unit (2002) report that the a high proportion of offenders sentenced to prison were in receipt of benefits. Similarly, a high proportion of offenders in this research were also in receipt of benefits prior to sentencing. However, those in receipt of benefits were sentenced less severely than those who were earning a wage. This result would appear to be inconsistent with the previous result, in which unemployed offenders are sentenced almost a fifth (18.62%) more severely than those in employment. However, this result may be connected to the offenders means to pay. For example, Flood-Page and Mackie (1998) found that unemployed offenders were less likely to receive a fine, but were three times more likely to receive a conditional discharge. In interviews with magistrates, 7 out of the 12 said they sometimes used a conditional discharge where they felt the offender could not afford a fine (Flood-Page and Mackie, 1998). One judge was quoted as saying:

"We are more often than not forced rather then led on the decision to fine or discharge by means of the defendant. You can't get blood out of a stone so, even if a fine was the most appropriate sentence, we have to resort to a conditional discharge" (Flood-Page and Mackie, 1998, p. 49)

Although fines are now means tested, this type of mentality could still be operating within the courts. It may be that some judges are taking pity on offenders who do not have sufficient resources. However, further work would need to be conducted to investigate whether this is the likely cause.

As far as I am aware, to date, there are no UK based studies which consider the income status of the offender as a factor determining sentence severity and disparity.

10.5   Risk

Although it is recognised that offenders are not a homogeneous group, a range of problems/needs are more commonly observed amongst this group than in the general population (Ministry of Justice, 2013b). These factors are understood to be associated with an increased risk of offending: the purpose of risk assessment then is to predict future criminality using an explicit set of factors (see Section 5.2.1 Figure 5.1) that correlate with reoffending, and to classify offenders into specific groups based on their likelihood of reconviction (Kleiman et al., 2007). As explained in Chapter 5, these 'criminogenic needs'/factors are recorded in the OASys pre-sentence assessment, including; employment status, income status, accommodation status, relationships, drug and alcohol misuse, attitudes, lifestyle and associations, and emotional well-being. These factors then contribute towards the general reoffending predictor (OPG) score, and which along with the pre-sentence report, is given to the judge prior to sentencing. Judges then have to make decisions based on all the information they have.

This research indicates that certain offenders or rather, offenders with certain characteristics are being sentenced more severely than others due to their extra-legal characteristics. One explanation for this is the increased risk offenders with these characteristics possess and therefore they are sentenced more severely.

The argument surrounding the risk debate generally depends on one's perspective on the purposes of punishment (Silver and Chow-Martin, 2002). In Chapter 2 by way of the sentencing schema diagram (see Figure 2.1 in Section 2.2.1), I explained that generally speaking there two main types of sentencing focus; a consequentualist or forward looking focus and a non-consequentualist or backward looking focus (Francis et al., 2001). The retributivist sentencing approach can be thought of as non-consequentualist, and this position adopts a backward looking focus that uses proportionality as the means for punishing offenders. This approach views future risk

as irrelevant (Kleiman et al., 2007): consideration of the premise of risk may in fact lead to discrimination (Ashworth, 2010b). On the other hand, the consequentualist or forward-looking utilitarian approach would justify punishment as a means of protecting the public and preventing future crime and so rely on risk as a method of predicting the likelihood of further criminality (ibid).

The justice system in England in Wales adopts a hybrid sentencing approach which places emphasis on proportionate sentencing based on the seriousness of the offence by assessing the harm and culpability caused or could potentially have been caused by the offence (retributivism)(Sentencing Guidelines Council, 2004), and also encompasses deterrence, protecting the public, rehabilitation, and reparation, as principles which also need to be considered when making sentencing decisions (Criminal Justice Act 2003). These additional principles are partly achieved through the use of actuarial risk assessment to predict the risk of recidivism but also the risk the offender may pose towards others. As explained in Chapter 2, for more serious offences (offences which pass the community order threshold) this information is provided to the courts in the form of pre-sentence reports which are then used by judges to aid the decision-making process.

Previous research from the US. has suggested that judges are guided by three focal concerns when making sentencing decisions; the offender's blameworthiness; protection of the community; and practical implications of sentencing decisions (Steffensmeier and Demuth, 2000). The offenders blameworthiness is associated with the offence and their role in the offence, their history of offending, and whether they have themselves been a victim (which tends to mitigate the perceived blameworthiness) (Steffensmeier and Demuth, 2000). Protection in the community focuses on the need to incapacitate the offender on assessment of potential future behaviour, based on the legal characteristics of the present case (whether they are being sentenced for a serious/dangerous offence), their criminal history and perhaps on the characteristics

of the offenders such as education, employment, or community ties (ibid). Organisational constrains and practical consequence include the suitability of the offender for a particular offence and the disruption of ties to children and other family members (ibid). Additionally, judges are also likely to be concerned about the repercussions of recidivism on the court in the public's eyes and on their own careers (ibid). Steffensmeier and Demuth (2000) also suggest that due to the complex interplay and nature of these concerns, judges may rely on attributes linked to the offender's gender, race, social class and other social positions. This illustrates that judges use a host of factors to reach their decision when sentencing offenders which corresponds with Johnson's (2005) point that sentencing does not occur within a social vacuum. Similar factors are also likely to impact on sentences

This focus on risk may explain why we are seeing these disparities between offenders who are being sentenced for similar offences. For example, previous research has shown that those who are unemployed prior to sentencing are at greater risk of recidivism (Social Exclusion Unit, 2002). This could be a justification for sentencing more severely even though it goes against the principle of equality before the law. However, the same justification cannot be used to explain why there is sentencing disparity based on the ethnicity of the offender. It is therefore reasonable to argue that discrimination is present where an offender's ethnicity is effecting sentencing decisions.

Further research would need to delve into the relationship between risk and disparity to begin to unpick how the two interplay. However, to reduce the possibility of disparity occurring, sentencing legislation and guidelines need to be more specific about which factors should be taken into consideration when sentencing offenders (Ashworth, 2010a) but also just how much weight they should also carry (Roberts, 2011a). If these were to be made clearer, there may be less room for disparity to occur.

10.6 A two stage sentencing process

Proponents of the utilitarian approach to sentencing, such as Morris and Tonry (1990) and Kolber (2009), may be critical of a unidimensional severity scale, insisting that this type of scale fails to take into consideration the effects that different sentences may have on individual offenders. They may also suggest that this type of scale fails to take into consideration the 'punitive bite' of the different types of sentence. However, this is not the case. The new sentence severity scale demonstrates the interchangeability of sentences which then allows us to identify sentences which are equivalent in terms of their severity.

Steffensmeier et al. (1993) explain that sentencing can be seen as a two-stage process, which firstly involves the decision as to whether or not to imprison, and secondly, if incarceration is appropriate, then a decision about the length of the sentence must be made. We can adapt this idea, so that once the adequate level of severity is determined (which ties in with Steffensmeier et al.'s (1993) idea of deciding whether a custodial sentence is appropriate, except it includes all sentences not just custodial sentences), it is then be possible to substitute sentences to take into consideration possible factors that will determine the suitability of a given sentence. These may include parental responsibilities or where for example a custodial sentence will result in others being unemployed if the offender employs staff. This would then tie in with Steffensmeier et al.'s (1993) decision on the appropriate length of sentence. This could be possible if it were to be made explicit in the sentencing legislation which factors should be taken into consideration. This is the case in Minnesota, where for example the sentencing legislation explicitly excludes employment status from the decision-making process or more specifically, as a factor to depart from the guideline sentence (Minnesota Sentencing Guidelines Commission, 2013). Equality before the law states that employment status should not be taken into consideration at the sentencing stage (Warner, 2012), but as this research suggests, as well as previous research by Jacobson

and Hough (2007), Flood-Page and Mackie (1998), and Moxon (1988) for example, being unemployed as opposed to employed, increases the severity of the sentence by almost a fifth (18.6%).

10.7   Consideration for blind sentencing

Sentencing disparity has a number of implications, namely that certain types of offenders are being discriminated against which in turn leads to injustice if we use the framework/pathway for achieving justice as discussed in Section 3.2.1. Is it possible to ever limit this disparity?

The 12 foot tall gold leaf statue which stands on top of the dome at the Old Bailey in London is the depiction of Lady Justice. Her modern representation is typically blindfolded and seen with scales in her right hand and a sword in her left. Lady Justice is said to be the figurative representation of the moral force in judicial systems symbolising fair and equal justice. Aas (2005) explains that the blindfold, scales and sword indicate justice being served regardless of skin colour, social status etc. in those who appear before the courts. However, she points out that the blindfold does appear absent at times which possibly represents what she describes as a profound duality in the nature of Western penal law. She goes on to say, on one hand there is a demand that punishment should be equal for all and that justice should be blind. On the other hand, one should know the ones that are being punished (ibid).

Could this notion of the Lady Justice be used in the courts? One extreme way around this would be for a judge to not be made aware of any extra-legal factors of the offender and therefore they could be said to be 'going in blind'. Sentencing decisions could then be made purely on the basis of the legal factors related to the case. In doing so, their decisions are agreed upon without the offender ever appearing in court. Similar practice occurs in Rio de Janeiro, whereby decisions are made on paper before

the offender attends court (Humphreys et al., 2014). The name of the offender can sometimes indicate their race or ethnicity and therefore judges may make decisions based on this extra-legal factor. There may then also be an argument to omit this information altogether. Therefore, I propose that judges should not be aware of any extra-legal characteristics of the offender, only the characteristics of the offence and the case which includes any aggravating and mitigating factors. This way, even if unconscious bias of the judge exists they will not be able to base their judgements on these factors.

10.8 Limitations of this work

This section will discuss the main limitations of this work and the following section will then discuss how this research could be improved through future work.

10.8.1 Data availability and quality

There were a number of issues surrounding the availability of data and data quality issues in the PNC and missing data in the OASys assessment. In Section 5.2.2 I explained that the complex nature of some of the newer sentences (for example, community orders and suspended sentence orders) are not adequately captured in the PNC. As well as the length of time the offender is subjected to these types of sentences, these sentences have a number of different elements or requirements which ultimately determine their severity. However, these requirements are not recorded in the PNC and therefore only the length of these sentences were available as an indication of their severity. More detailed information on these types of sentences would allow finer degradation of these types of sentences which would consequently improve the accuracy of the extended Goodman RC association analysis and ultimately the sentence severity scale. The Crown Court Survey for example records not only the

type and length of the sentence, but for suspended sentence orders and community orders it also records the types of requirements attached to the orders. However this data is not yet available at the individual level which is required to conduct this research.

By extending the Goodman RC association model, this research was able to control for the imbalance that legal factors can have on sentence severity, making a sentence more or less severe. Only three legal factors (offence plea, number of offences the offender was sentenced for, and whether the offender had previously appeared in court) were included in the extended Goodman analysis. As I explained in Section 6.5, I provisionally wanted to control for the statutory aggravating and mitigating factors such as, previous convictions, whether the offence was committed when the offender was on bail, whether the offence was aggravated in relation to race, religion, disability or sexual orientation of the victim, whether the offence involved a terrorist connection, and plea, which is considered a mitigating factor. However, a lot of data relating to the aggravating factors were not adequately recorded in the OASys assessment and were therefore essentially missing.

In light of this, two additional covariates were created. The first covariate was an indicator of prior offending due to the lack of information relating to the offenders' previous convictions. Prior court appearance was therefore used to indicate previous involvement in the court system. As explained in Section 2.3.1, previous convictions have a cumulative effect on sentence severity. This new variable is admittedly a very poor indicator of an offenders' previous convictions as it masks the true extent that prior convictions have on sentence severity. Future work would therefore greatly benefit from having a more accurate measure of offenders' conviction histories. The second variable that was created relates to the principle of totality (see Section 2.5.2). This covariate therefore identified cases where the offender was sentenced for multiple offences. Again the Crown Court Survey records both of these factors and also

asks judges to record any other aggravating or mitigating factors that they took into consideration sentencing the offender.

In the second part of this research multilevel modelling was used to measure sentencing disparity. Here a number of legal and extra-legal factors were included in the model to identify if and where disparity occurred - either at the offence or offender level. The same legal variables that were controlled for in the extended Goodman RC analysis were also included in the multilevel modelling analysis. As explained in Section 10.8.1, by including these variables at this stage allowed me to control for the imbalance that these factors have on sentence severity. Therefore, by also including for them in the multilevel modelling, it was then possible to identify how sentences altered. It was possible to identify change in sentence severity, for example, how much more severe is a not guilty plea as opposed to guilty plea.

As discussed in Section 10.8.1, the lack of prior conviction information is a limitation to this work as previous convictions increase the seriousness of offences. Prior court appearance was used as an indicator of this, but this did not improve the multilevel model fit and therefore was not included in the optimal model. If previous conviction history data had been available it would have been included in this analysis and it is anticipated that it would have been highly statistically significant. It is also likely that by including this variable it would have explained far more of the variability at the offender level.

In a similar vein, due to the amount of missing data in the OASys assessment only 14 covariates were tested in the multilevel model (offence, plea, number of offences, prior court appearance, court, weapon, direct victim, sex, ethnicity, age, employment status, accommodation status and income status). If more data had been available in terms of more aggravating and mitigating factors and additional extra-legal factors, including information relating to the OASys risk factors (see Section 5.5.1), then these would have been included in the model. Potentially far more of the variation at the

offender level would have therefore been explained.

Related to the issue of data availability is the amount of missing data also identified in the data. In fact the missing data proved to be one of the biggest limitations in conducting this research. Variables such as the court the offender was sentenced in was completely missing and it is believed this was a conscious measure to ensure the data was completely anonymous. Other variables such as the offenders ethnicity, accommodation status and income status had a lot of missing information. As explained in Section 8.4.1 this data was treated as Missing Completely at Random (MCAR) as it is thought that the data is missing because those carrying out the assessments failed to complete the information. Assessments now are fully electronic and the person completing the assessment cannot skip sections without completing all the fields. This will therefore ensure complete assessments without the missing data encountered in this research.

10.8.2   Extensions of analysis

The optimal multilevel model (Model 24) used in this research had two levels - level 1 offences and level 2 offenders. An extension of the model would have been to extend the model to account for a third level - Probation trust. This was tested but did not improve the model fit and so was not included in the model. This is potentially because there was too much sentencing variation at this aggregate level and therefore a more refined measure would have been better. As just discussed in Section 10.8.1, court information was removed from the data in a bid to ensure the offenders in the same remained anonymous. It is my belief that if this covariate had been available to use as a third level in the multilevel model, it would have improved the model fit and therefore would have been included in the analysis. For example, in the famous study by Hood (1992), he found there to be significant sentencing variation within courts. This disparity would have been masked if the analysis had not considered the court

level effects. However, in this research the availability of the data prohibited this line of analysis.

Another possible extension to the random intercept multilevel model would have been to also vary the slopes of the model as well as the intercepts as illustrated in Figure 8.1 diagram (b). This type of model is called a random coefficient model. Had the three level model with Probation Trust been the preferred model, then it would have made sense to have gone on and extended the random intercept model. By doing this, we could have identified if for example the affects of the extra-legal offender characteristics varied by Probation trust. As the two level model turned out to be the optimal modal, a random coefficient model would not have been suitable because there are 13,495 offenders at level 2 opposed to 35 probation trust areas. Therefore the random coefficient model would have been overly complicated and would not have made a lot of sense.

10.9   Future work

This research has shown that it is possible to model the association between offences and sentences to create a measure of sentence severity. This work can be seen as the first step to potentially creating a severity scale in which all sentences are arranged on a continuum. Future work would benefit from using more detailed sentence categories, for example more detailed information relating to the type and length of requirements attached to community orders and suspended sentence orders. This would allow further refinement of the sentence severity scale to more accurately measure sentence severity.

Data from the Crown Court Sentencing Survey could provide this if it were to be made available at the offender level. At present the data from the Crown Court Survey is aggregated, and at this level of aggregation, it does not permit the statistical

modelling necessary to conduct similar analysis. Additionally, if the data were to be disaggregated and matched with OASys data, further research would benefit from more detailed sentence categories and legal characteristics, as well as a more detailed list of extra-legal factors obtained from the probations data. Initial contact has been made with the Sentencing Council about the possibility of securing access to this data. However, at present the Sentencing Council are only considering the possibility of making this data more readily available.

Future work would also benefit from incorporating additional sentences into the measure of sentence severity where offenders receive more than one type or combination of sentence. In Section 5.2.2, it was explained that due to the problems identified with the way sentences were recorded in the PNC, the decision was made to only use the first sentence recorded in the data as it was understood to be the most accurate. If this information was recorded differently (or obtained through another means), this would also add to the accuracy of any future sentence severity scale. However, this research already goes beyond the previous attempts at measuring sentence severity by including all offences that offenders were sentenced for, rather than only focusing on the primary offence, as has been done in previous research.

Additionally, future work into disparity would also need to make sure there was sufficient information relating to the offender's prior history of offending, including the number of previous convictions, as well as, the previous sentences they have served. On top of this, any future analysis would also benefit from more complete data. Since receiving the data for this research the Probation Service have implemented new quality control measures which should ultimately improve the quality of the information recorded.

In spite of these suggestions for future work, this research stands alone as a competent and informative piece of research. Therefore we can be more confident that the results of this research, based on statistical modelling of real data (both the extended

Goodman RC Association and multilevel modelling analyses) evidence sentencing disparity within the courts of England and Wales. This research offers some indication that sentencing disparity and discrimination is present within sentencing in England and Wales, which must not go unnoticed. Future work using data such as the Crown Court Survey data could potentially overcome these limitations and improve this research.

Bibliography

Aas, K. F. (2005). *Sentencing In The Age Of Information: From Faust To Macintosh.* The GlassHouse Press, London.

Aitkin, M., Francis, B., Hinde, J., and Darnell, R. (2009). *Statistical Modelling in R.* Oxford University Press, Oxford.

Albonetti, C. A. (1997). Sentencing under the federal sentencing guidelines: Effects of defendant characteristics, guilty pleas, and departures on sentence outcomes for drug offenses, 1991-1992. *Law and Society Review*, 31(4):789–822.

Allen, R. B. and Anson, R. H. (1985). Development of a punishment severity scale: The item displacement phenomenon. *Criminal Justice Review*, 10(2):39–44.

Anderson, A. L. and Spohn, C. (2010). Lawlessness in the federal sentencing process: A test for uniformity and consistency in sentence outcomes. *Justice Quarterly*, 27(3):362–393.

Arnull, E., Eagle, S., Gammampila, A., Patel, S. L., Sadler, J., Thomas, S., and Bateman, T. (2007). Youth Justice Board Report: Housing needs and experiences. `https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/354901/yjb-accommodation-needs-experiences.pdf`. [Online; accessed 02-November-2014].

Ashworth, A. (1987). Disentangling disparity. In Pennington, D. and Lloyd-Bostock,

S., editors, *The Psychology of sentencing: approaches to consistency and disparity*. Centre for Social-Legal Studies, Oxford.

Ashworth, A. (1992). Sentencing reform structures. *Crime and Justice*, 16:181–241.

Ashworth, A. (1994). Towards European sentencing standards. *European Journal on Criminal Policy and Research*, 2(1):7–11.

Ashworth, A. (2005). *Sentencing and Criminal Justice*. Cambridge University Press, Cambridge.

Ashworth, A. (2010a). Coroners and Justice Act 2009: Sentenncing guidelines and the Sentencing Council. *Criminal Law Review*, 5(1):389–401.

Ashworth, A. (2010b). *Sentencing and Criminal Justice*. Cambridge University Press, Cambridge.

Ashworth, A. and Redmayne, M. (2010). *The Criminal Process*. Oxford University Press, Oxford.

Ashworth, A. and Roberts, J. V. (2013). *Sentencing Guidelines: Exploring the English Model*. Oxford University Press, Oxford.

Austin, W. and Williams, T. A. (1977). A survey of judge's responses to simulated legal cases: Research note on sentencing disparity. *Jorrnal of Criminal Law and Criminology*, 68(2):306–310.

Baayen, R. H., Davidson, D. J., and Bates, D. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(1):390–412.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7.

Blumstein, A. (1983). *Research on sentencing: The search for reform.* National Academy Press, Washington.

Borga, M. (2001). Canonical correlation: A tutorial. `http://www.imt.liu.se/people/magnus/cca/tutorial/tutorial.pdf`. [Online; accessed 03-March-2012].

Bowling, B. and Philips, C. (2003). Policing Ethnic Minority Communities. In Newburn, T., editor, *Handbook of Policing*. Willan Publishing, Devon.

Braithwaite, J. (1982). Challenging just deserts: Punishing white-collar criminals. *Journal of Criminal Law & Criminology*, 73(2):723–763.

Britt, C. L. (2009). Modeling the distribution of sentence length decisions under a guidelines system: An application of quantile regression models. *Journal of Quantitative Criminology*, 25(4):341–370.

Bryk, A. S. and Raudenbush, S. W. (1992). *Hierarchical linear models : Applications and data analysis methods.* Thousand Oaks, California.

Buchner, D. (1979). Scale of sentence severity. *Journal of Criminal Law & Criminology*, 70(2):182–187.

Bushway, S. D. and Piehl, A. M. (2001). Judging judicial discretion: Legal factors and racial discrimination in sentencing. *Law and Society Review*, 35(4):733–764.

Carlen, P. (1988). *Women, Crime and Poverty.* Open University Press, Milton Keynes.

Carter, P. (2003). *Managing offenders, reducing crime: A new approach.* UK Cabinet Office Strategy Unit, London.

Cavadino, M. and Dignan, J. (2007). *The penal system: An introduction.* Sage, London.

Chang, L (2010). lme4. `http://www.u.arizona.edu/~ljchang/NewSite/papers/LME4_HO.pdf`. [Online; accessed 10-July-2015].

Chesney-Lind, M. (1995). Rethinking women's imprisonment: A critical examination of trends in female incarceration. In Price, B. R. and Skolff, N. I., editors, *The criminal justice sytem and women*. McGraw-Hill, New York.

Chesney-Lind, M. and Pasko, L. (2012). *The female offender: Girls, women, and crime*. Sage, London.

Clogg, C. C. (1982). Association Models in Sociological Research: Some Examples. *American Journal of Sociology*, 88(1):114–134.

Corston, J. (2007). The Corston Report: A Report of a Review of Women with Particular Vulnerabilities in the Criminal Justice System: Home Office Report. `http://www.justice.gov.uk/publications/docs/corston-report-march-2007.pdf`. [Online; accessed 27-March-2014].

Council of Europe Committee of Ministers (1993). Consistency in Sentencing: Recommendation No. R(92) 17.

Courts and Tribunals Judicary (2014). Judges. `https://www.judiciary.gov.uk/about-the-judiciary/who-are-the-judiciary/judicial-roles/judges/`. [Online; accessed 09-December-2014].

Crow, I. and Simon, F. (1989). *Unemployment and Sentencing*. NACRO, London.

Crown Prosecution Service (2012). Annual report and accounts 2011-2012. `http://www.cps.gov.uk/publications/reports/2011/annex_b.html`. [Online; accessed 10-December-2014].

Croyle, J. L. (1983). Measuring and explaining disparities in felony sentences: Courtroom work group factors and race, sex, and socioeconomic influences on sentence severity. *Political Behavior*, 5(1):135–153.

Daly, K. (1989). Rethinking judicial paternalism: Gender, work-family relations, and sentencing. *Gender & Society*, 3(1):9–36.

Daly, K. (1994). *Gender, Crime and Punishment*. Yale University Press, New Haven, CT.

Danziger, S., Levav, J., and Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, 108(17):6889–6892.

Department for Constitutional Affairs (2006). *A Guide to the Human Rights Act 1998: Third Edition*. Department for Constitutional Affairs, London.

Dhami, M. K. (2013). Sentencing Guidelines in England and Wales: Missed Opportunities? *Law Contemporary Problems*, 76(1):289–307.

Dowds, L. and Hedderman, C. (1997). The sentencing of men and women. In Hedderman, C. and Gelsthorpe, L., editors, *Understanding the sentencing of women. Home Office Research Study 170*. Home Office , London.

Duff, A. (2001). *Punishment, Communication, and Community*. Oxford University Press, Oxford.

Dyer, P. (2013). Bail in criminal proceedings. `http://www.qebholliswhiteman.co.uk/articles-pdfs/bail-update.pdf`. [Online; accessed 03-November-2014].

Easton, S. and Piper, C. (2008). *Sentencing and Punishment: The Quest for Justice*. Oxford University Press, Oxford.

Eaton, M. (1986). *Justice for Women*. Open University Press, Philadelphia.

Engen, R. L. and Gainey, R. R. (2000). Modeling the effects of legally relevant and extralegal factors under sentencing guidelines: The rules have changed. *Criminology*, 38(4):1207–1230.

Equality Act (2010). Equality act 2010. *HM Stationery Office, London.*

Equality and Human Rights Commission (2011). *Services, public functions and associations: Statutory Code of Practice.* The Stationery Office, London.

Erickson, M. L. and Gibbs, J. P. (1979). On the perceived severity of legal penalties. *Journal of Criminal Law & Criminology*, 70(1):102–116.

Erikson, R. and Goldthorpe, J. H. (1992). *The Constant Flux: A Study of Class Mobility in Industrial Societies.* Claredon, Oxford.

Farrington, D. P. and Morris, A. M. (1983). Sex, sentencing and reconviction. *British Journal of Criminology*, 23(3):229–248.

Fawcett Society (2009). What about women? `http://www.fawcettsociety.org.uk/wp-content/uploads/2013/02/Fawcett-Society-What-About-Women-report-low-res.pdf`. [Online; accessed 08-June-2013].

Feeley, M. M. and Simon, J. (1992). The new penology: Notes on the emerging strategy of corrections and its implications. *Criminology*, 30(4):449–474.

Fielding, J. and Gilbert, N. (2006). *Understanding Social Statistics.* Sage, London.

Flood-Page, C. and Mackie, A. (1998). *Sentencing practice: an examination of decisions in magistrates' courts and the Crown Court in the mid-1990's. Home Office Research Study 180.* Home Office London.

Foucault, M. (1977). *Discipline and punish: The birth of the prison.* Random House LLC.

Francis, B., Soothill, K., and Dittrich, R. (2001). A new approach for ranking serious offences. the use of paired-comparisons methodology. *British Journal of Criminology*, 41(4):726–737.

Freedman, D. (2001). Ecological inference and the ecological fallacy. In Smelser, R. M. and Baltes, P. B., editors, *International Encyclopedia for the Social and Behavioral Sciences 6*. Elsevier , Oxford.

Garland, D. (1990). *Punishment and Modern Society: A Study in Social Theory.* University of Chicago Press, Chicago.

Goodman, L. A. and Hout, M. (1982). Statistical methods and graphical displays for analysing how the association between two qualitative variables differs amoung countries, or over time. part ii: Some explanatory techniques, simple models and simple examples. *Sociological Methodology*, 31:189–221.

Grimshaw, R. (2002). Accommodation and offending - What Works? An international literature review: Centre for Crime and Justice Studies Report. `http://www.crimeandjustice.org.uk/sites/crimeandjustice.org.uk/files/Accommodation_and_offending.pdf`.

Halliday, J. (2001). *Making punishments work: report of a review of the sentencing framework for England and Wales.* Home Office, London.

Harlow, R. E., Darley, J. M., and Robinson, P. H. (1995). The severity of intermediate penal sanctions: A psychophysical scaling approach for obtaining community perceptions. *Journal of Quantitative Criminology*, 11(1):71–95.

Hedderman, C. and Gelsthorpe, L. (1997). *Understanding the sentencing of women.* Home Office, London.

Helms, R. and Jacobs, D. (2002). The political context of sentencing: An analysis of community and individual determinants. *Social Forces*, 81(2):577–604.

Her Majesty's Inspectorate of Prisons (2007). The Mental Health of Prisoners: A Thematic Review of the Care and Support of Prisoners with Mental Health Needs. `http://www.rcpsych.ac.uk/pdf/TheMentalHealthPrisoners2007.pdf`.

Hogarth, J. (1971). *Sentencing as a human process.* University of Toronto Press, Toronto.

Holden, L. (2007). Offender Assessment System (OASYS). In Canton, R. and Hancock, D., editors, *Dictionary of Probation and Offender Management.* Willan Publishing, Devon.

Holleran, D. and Spohn, C. (2004). On the use of the total incarceration variable in sentencing research. *Criminology*, 42(1):211–240.

Hood, R. (1992). *Race and Sentencing.* Clarendon Press, Oxford.

Hough, M. and Jacobson, J. (2008). Creating a sentencing commission for England and Wales: Prison Reform Trust Report. `http://www.prisonreformtrust.org.uk/Portals/0/Documents/Creating%20a%20Sentencing%20commission%20for%20England%20and%20Wales.pdf`.

House of Commons Justice Committee (2013). Women offenders: after the Corston Report. `http://www.parliament.uk/documents/commons-committees/Justice/Women-offenders.pdf`. [Online; accessed 27-October-2014].

Howard, P. (2006). *The Offender Assessment System: an evaluation of the second pilot.* Home Office, London.

Howard, P. (2011). *Hazards of different types of reoffending.* Ministry of Justice, London.

Hox, J. (2002). *Multi-level Analysis: Techniques and Applications.* Lawrence Erlbaum Associates Inc, New Jersey.

Hox, J. (2010). *Multi-level Analysis: Techniques and Applications.* Routledge, London.

Hudson, B. (1993). *Penal Policy and Social Justice.* Macmillan, London.

Hudson, B. (1996). *Understanding Justice: An Introduction to Ideas, Perspectives, and Controversies in Modern Penal Theory.* Open University Press, Buckingham.

Hudson, B. (1998). Doing Justice to Difference. In Ashworth, A. and Wasik, M., editors, *Fundamentals of Sentencing Theory.* Oxford Clarendon Press, Oxford.

Humphreys, L., Francis, B., Cano, I., and Ribeiro, E. (2014). Assessing gender and racial bias in sentencing in Rio de Janeiro: Working Paper.

Hutton, N. (2013). The definitive guideline on assault offences: The performance of justice. In Ashworth, A. and Roberts, J. V., editors, *Sentencing Guidelines: Exploring the English Model.* Oxford University Press, Oxford.

Jacobson, J. and Hough, J. (2007). Mitigation: the role of personal factors in sentencing: Prison Reform Trust Report. `http://www.prisonreformtrust.org.uk/Portals/0/Documents/mitigation%20-%20the%20role%20of%20personal%20factors%20in%20sentencing.pdf`.

Johnson, B. D. (2005). Contextual disparities in guidelines departures: Courtroom social contexts, guidelines compliance, and extralegal disparities in criminal sentencing. *Criminology*, 43(3):761–796.

Johnson, B. D. (2006). The multilevel contect of criminal sentencing: integrating judge and country-level influences. *Criminology*, 44(2):259–298.

Johnston, C. W. and Alozie, N. O. (2001). The effect of age on criminal processing: Is there an advantage in being 'older'? *Journal of Gerontological Social Work*, 34(4):65–82.

Kautt, P. M. (2002). Location, location, location: Interdistrict and intercircuit variation in sentencing outcomes for federal drug-trafficing offences. *Justice Quarterly*, 19(4):633–671.

Kirkwood, S. and Richley, T. (2008). *Supported accommodation services for offenders: A research literature review*. CJSWDC Briefing Papers 11, Edinburgh.

Kleiman, M., Ostrom, B. J., and Cheesman II, F. L. (2007). Using Risk Assessment to Inform Sentencing Decisions for Nonviolent Offenders in Virginia. *Crime and Delinquency*, 53(1):106–132.

Klepper, S., Nagin, D., and Tierney, L. (1983). Discrimination in the criminal justice system: A critical appraisal of the literature. In Blumstein, A., editor, *Research on sentencing: The search for reform*. National Academy Press, Washington.

Kolber, A. J. (2009). The comparative nature of punishment. *Boston University Law Review*, 89(1):1565–1608.

Leiter, B. (2005). Americam legal realism. In Golding, M. P. and Edmundson, W. A., editors, *The Blackwell guide to the philosophy of law and legal theory*. Blackwell, London.

Luke, D. A. (2004). *Multilevel Modelling*. Sage Publications, London.

Maguire, N. (2010). Consistency in sentencing. *Judicial Studies Institute Journal*, 2(1):14–54.

Malila, I. S. (2012). Severity of multiple punishments deployed by magistrate and customary courts against common offences in botswana: A comparative analysis. *International Journal of Criminal Justice Sciences*, 7(2):618–634.

Marinos, V. (2005). Thinking about penal equivalents. *Punishment & Society*, 7(4):441–455.

Mason, T., de Silve, N., Sharma, N., Brown, D., and Harper, G. (2007). *Local variation in sentencing in England and Wales.* Ministry of Justice, London.

McDavid, J. C. and Stipak, B. (1981). Simultaneous scaling of offense seriousness and sentence severity through canonical correlation analysis. *Law and Society Review*, 16(1):147–162.

Merrall, E. L., Dhami, M. K., and Bird, S. M. (2010). Exploring methods to investigate sentencing decisions. *Evaluation Review*, 34(3):185–219.

Miethe, T. D. and Moore, C. A. (1986). Racial differences in criminal processing: The consequences of model selection on conclusions about differential treatment*. *The Sociological Quarterly*, 27(2):217–237.

Ministry of Justice (2009). Statistics on race and the criminal justice system 2007/8. *Statical Bulletin.*

Ministry of Justice (2010). Sentencing statistics: England and wales 2009 statistics bulletin. *Statistical Bulletin.*

Ministry of Justice (2011). Statistics on race and the criminal justice system 2010. *Statical Bulletin.*

Ministry of Justice (2013a). 2013 Compendium of re-offending statistics and analysis. `https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/278133/compendium-reoffending-stats-2013.pdf`. [Online; accessed 24-October-2014].

Ministry of Justice (2013b). Story of the prison population: 1993 - 2012 england and wales. `https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/218185/story-prison-population.pdf`. [Online; accessed 30-January-2014].

Minnesota Sentencing Guidelines Commission (2013). Minnesota Sentencing Guidelines Commentary. `http://mn.gov/sentencing-guidelines/images/2014%2520Guidelines.pdf`.

Morgan, R. and Russell, N. (2000). *The judiciary in the magistrates' courts*. Home Office, London.

Morris, N. and Tonry, M. (1990). *Between Prison and Probation: Intermediate Punishments in a Rational Sentencing System*. Oxford University Presst, New York.

Moxon, D. (1988). Sentencing practice in the crown court. *Home Office Research Study 103*.

Mueller-Johnson, K. U. and Dhami, M. K. (2009). Effects of offenders' age and health on sentencing decisions. *The Journal of Social Psychology*, 150(1):77–97.

Mustard, D. B. (2001). Racial, ethnic, and gender disparities in sentencing: Evidence from the us federal courts. *Journal of Law and Economics*, 44(1):285–314.

Nagel, I. and Johnson, B. (1994). The role of gender in a structured sentencing system: Equal treatment, policy choice, and the sentencing of female offenders under the united states sentencing guidelines. *The Journal of Criminal Law Criminology*, 85(1):181–221.

Padfield, N. (2013). Exploring the success of Sentencing Guidelines. In Ashworth, A. and Roberts, J. V., editors, *Sentencing Guidelines: Exploring the English Model*. Oxford University Press, Oxford.

Palys, T. and Divorski, S. (1986). Explaining sentencing disparity. *Canadian Journal of Criminology*, 82(1):347–362.

Phillips, C. and Bowling, B. (2002). Racism, ethnicity, and criminal justice. In

M. Maguire, R. M. and Reiner, R., editors, *The Oxford Handbook of Criminology.* Oxford University Press, Oxford.

Phillips, C. and Brown, D. (1998). *Entry into the criminal justice system: a survey of police arrests and their outcomes.* Home Office, London.

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.

Powers, D. A. and Xie, Y. (2000). *Statistical Methods for Categorical Data Analysis.* Academic Press, London.

Prison Reform Trust (2000). *Justice for women: The need for reform.* Prison Reform Trust, London.

Prison Reform Trust (2013). Resettlementt. `http://www.prisonreformtrust.org.uk/ProjectsResearch/Resettlement`. [Online; accessed 17-November-2014].

Rex, S. and von Hirsch, A. (1997). Community orders and the assessment of punishment severity. *Federal Sentcing Reporter*, 10:278.

Richardson, J. (2011). *Archbold: Criminal Pleading, Evidence, and Practice 2012.* Sweet and Maxwell, London.

Roberts, J. V. (2008). Aggravating and mitigating factors at sentening: towards greater consistency in application. *Criminal Law Review*, 264.

Roberts, J. V. (2009). Revisiting the recidivist sentencing premium. In von Hirsch, A., Ashworth, A., and Roberts, J., editors, *Principled Sentencing: Readings on Theory and Policy.* Hart Publishing, Oxford.

Roberts, J. V. (2011a). Punishing, more or less: exploring aggravation and mitigation at sentencing. In Roberts, J. V., editor, *Mitigation and Aggravation at Sentencing.* Cambridge University Press, Cambridge.

Roberts, J. V. (2011b). Sentencing guidelines and judicial discretion: Evolution of the duty of courts to comply in england and wales. *British Journal of Criminology*, 51:997–1013.

Samuels, A. (1987). Consistency in sentencing. In Pennington, D. and Wasik, S. L.-B., editors, *The Psychology of sentencing: approaches to consistency and disparity.* Centre for Social-Legal Studies, Oxford.

Schiff, M. F. (1997). Gauging the intensity of criminal sanctions: developing the criminal punishment severity scale (CPSS). *Criminal Justice Review*, 22(2):175–206.

Schwartz, J. and Steffensmeier, D. (2007). The nature of female offending: Patterns and explanation. *Female Offenders: Critical Perspectives and Effective Interventions (Second Edition).*

Scottish Government (2012). *Commission on Women Offenders.* Scottish Government, Edinburgh.

Sebba, L. (1978). Some explorations in the scaling of penalties. *Journal of Research in crime and delinquency*, 15(2):247–265.

Sebba, L. and Nathan, G. (1984). Further explorations in the scaling of penalties. *British Journal of Criminology*, 24:221.

Sellin, T. and Wolfgang, M. E. (1964). *The Measurement of Delinquency.* Wiley, Oxford.

Sentencing Council (2010). Analytical note: The Resource effects of increased consistency in sentencing. `http://sentencingcouncil.judiciary.gov.uk/docs/Consistency_in_sentencing.pdf`. [Online; accessed 04-April-2014].

Sentencing Council (2011). *Assault Definitive Guideline*. Sentencing Council, London.

Sentencing Council (2012a). *A short guide: Sentencing for multiple offences (Totality)*. Sentencing Council, London.

Sentencing Council (2012b). *Offences Taken into Consideration and Totality: Definitive Guideline*. Sentencing Council, London.

Sentencing Guidelines Counci (2004). New Sentences: Criminal Justice Act 2003. `http://sentencingcouncil.judiciary.gov.uk/docs/web_new_sentences_guideline1.pdf`. [Online; accessed 09-April-2013].

Sentencing Guidelines Council (2004). *Overarching Principles: Seriousness Guideline*. Sentencing Guidelines Council, London.

Sentencing Guidelines Council (2007). *Reduction in Sentence for a Guilty Plea*. Sentencing Guidelines Council, London.

Sentencing Guidelines Council (2008). *Magistrates' Court Sentencing Guidelines*. Sentencing Guidelines Council, London.

Sentencing Guidelines Council (2009). *Overarching Principles - Sentencing Youths*. Sentencing Guidelines Council, London.

Silver, E. and Chow-Martin, L. (2002). A multiple models approach to assessing recidivism risk: Implications for judicial decision making. *Criminal Justice and Behaviour*, 29(1):538–568.

Snijders, T. A. B. and Bosker, R. J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modelling, 2nd Edition.* Sage Publications Inc, London.

Social Exclusion Unit (2002). Reducing reoffending by ex-prisoners: Social Exclusion Unit Report. `http://www.bristol.ac.uk/poverty/downloads/keyofficialdocuments/Reducing%20Reoffending.pdf`. [Online; accessed 30-January-2012].

Spohn, C. (2009). *How do judges decide? The search for fairness and justice in punishment.* SAGE Publications Inc, London.

Spohn, C. and Belenko, S. (2013). Do the Drugs, Do the Time? The Efffects of Drug Abuse on Sentences Imposed on Drug Offenders in Three U.S District Courts. *Criminal Justice and Behavior*, 40(6):646–670.

Stacey, A. M. and Spohn, C. (2006). Gender and the social costs of sentencing: An analysis of sentences imposed on male and female offenders in three u.s. district courts. *Berkeley Journal of Criminal Law*, 43(1):43–76.

Steffensmeier, D. and Demuth, S. (2000). Ethnicity and Sentencing Outcomes in U.S. Federal Courts: Who is Punished More Harshly? *American Sociological Review*, 65(5):705–729.

Steffensmeier, D., Kramer, J., and Streifel, C. (1993). Gender and imprisonment decisions. *Criminology*, 31(3):411–446.

Steffensmeier, D., Kramer, J., and Ulmer, J. (1995). Age differences in sentencing. *Justice Quarterly*, 12(3):583–602.

Sweeten, G. (2012). Scaling Criminal Offending. *Journal of Quantitative Criminology*, 28(1):533–557.

Sykes, A. (1993). *An introduction to regression analysis.* Law School, University of Chicago, Chicago.

Tak, P. J. (1995). Sentencing in the netherlands: Discretion and disparity. *Federal Sentencing Reporter*, 7(6):300–304.

Tarling, R. and Weatheritt, M. (1979). *Sentencing practice in magistrates' courts.* HM Stationery Office, London.

Thomas, C. (2010). Are juries fairer? *Ministry of Justice Research Series 1/10.*

Thomas, D. (2008). *Sentencing Referencer 2008.* Sweet and Maxwell, London.

Thurstone, L. L. (1994). A law of comparative judgment. *Psychological Review*, 101(2):266–270.

Tiffany, L. P., Avichai, Y., and Peters, G. W. (1975). Statistical Analysis of Sentencing in Federal Courts: Defendants Convicted after Trial, 1967-1968. *The Journal of Legal Studies*, 4:369–390.

Tonry, M. (1994). Proportionality, parsimony, and interchangeability of punishments. In Duff, A. and Gardland, D., editors, *A Reader on Punishment.* Readings on Socio-Legal Studies, Oxford.

Tonry, M. (1996). *Sentencing Matters.* Oxford University Press, Oxford.

Turner, H. and Firth, D. (2012). *Generalized nonlinear models in R: An overview of the gnm package.* R package version 1.0-7.

UK Drug Policy Commission (2008). Reducing Drug Use, Reducing Reoffendings. `http://www.ukdpc.org.uk/wp-content/uploads/Policy%20report%20-%20Reducing%20drug%20use,%20reducing%20reoffending%20(summary).pdf.` [Online; accessed 01-November-2013].

Ulmer, J. T. and Johnson, B. (2004). Sentencing in context: A multilevel analysis. *Criminology*, 42(1):137–178.

U.S. Appeal Court (2014). UNITED STATES OF AMERICA versus QUARTAVI-OUS DAVIS,. `https://www.aclu.org/sites/default/files/assets/q_davis_opinion.pdf`. [Online; accessed 27-October-2014].

von Hirsch, A. (1992). Proportionality in the philsophy of punishment. *Crime and Justice*, 16(1):55–98.

von Hirsch, A. (1993). *Censure and Sanctions*. Clarendon press, Oxford.

von Hirsch, A., Ashworth, A., and Roberts, J. V. (2009). *Principled Sentencing: Readings on Theory and Policy*. Hart Publishing, Oxford.

von Hirsch, A., Wasik, M., and Greene, J. (1989). Punishments in the community and the principles of desert. *Rutgers Law Journal*, 20(3):595–618.

Walker, M. (1989). The court disposal and remands of white, afro-caribbean, and asian men (london 1983). *British Journal of Criminology*, 29(4):353–367.

Warner, K. (2012). Equality before the law and equal impact of sanctions: Doing justice to differences in wealth and emlpoyment status. In Zedner, L. and Roberts, J. V., editors, *Principles and Values in Criminal Law and Criminal Justice*. Oxford University Press, Oxford.

Weidner, R. R., Frase, R. S., and Schultz, J. S. (2005). The impact of contextual factors on the decision to imprison in large urban jurisdictions: A multilevel analysis. *Crime & Delinquency*, 51(3):400–424.

Wermink, H., Blokland, A., Nieuwbeerta, P., Nagin, D., and Tollenaar, N. (2010). Comparing the effects of community service and short-term imprisonment on re-

cidivism: a matched samples approach. *Journal of Experiemntal Criminology*, 6(1):325–349.

Wheeler, S., Weisburd, D., and Bode, N. (1982). Sentencing the white-collar offender: Rhetoric and reality. *American Sociological Review*, pages 641–659.

Woodredge, J. (2010). Judges' unequal contributions to extralegal disparities in imprisonment. *Criminology*, 48(2):539–567.

Xie, Y. (1992). The log-multiplicative layer effect model for comparing mobility tables. *American Sociological Review*, 57(3):380–395.

Appendices

Appendix A

Aggravating and mitigating factors

The Sentencing Guidelines Council (2004) provide a non-exhaustive list of the aggravating and mitigating factors

Aggravating factors: factors indicating higher culpability

- Offence committed whilst on bail for other offences

- Failure to respond to previous sentences

- Offence was racially or religiously aggravated

- Offence motivated by, or demonstrating, hostility to the victim based on his or her sexual orientation (or presumed sexual orientation)

- Offence motivated by, or demonstrating, hostility based on the victims disability (or presumed disability)

- Previous conviction(s), particularly where a pattern of repeat offending is disclosed

- Planning of an offence

- An intention to commit more serious harm than actually resulted from the offence

- Offenders operating in groups or gangs

- 'Professional' offending

- Commission of the offence for financial gain (where this is not inherent in the offence itself)

- High level of profit from the offence

- An attempt to conceal or dispose of evidence

- Failure to respond to warnings or concerns expressed by others about the offenders behaviour

- Offence committed whilst on licence

- Offence motivated by hostility towards a minority group, or a member or members of it

- Deliberate targeting of vulnerable victim(s)

- Commission of an offence while under the influence of alcohol or drugs

- Use of a weapon to frighten or injure victim

- Deliberate and gratuitous violence or damage to property, over and above what is needed to carry out the offence

- Abuse of power

- Abuse of a position of trust

Factors indicating a more than usually serious degree of harm:

- Multiple victims

- An especially serious physical or psychological effect on the victim, even if unintended

- A sustained assault or repeated assaults on the same victim

- Victim is particularly vulnerable

- Location of the offence (for example, in an isolated place)

- Offence is committed against those working in the public sector or providing a service to the public

- Presence of others e.g. relatives, especially children or partner of the victim

- Additional degradation of the victim (e.g. taking photographs of a victim as part of a sexual offence)

- In property offences, high value (including sentimental value) of property to the victim, or substantial consequential loss (e.g. where the theft of equipment causes serious disruption to a victims life or business)

Mitigating factors: factors indicating lower culpability:

- A greater degree of provocation than normally expected

- Mental illness or disability

- Youth or age, where it affects the responsibility of the individual defendant

- The fact that the offender played only a minor role in the offence

Offender mitigation

- Genuine remorse

- Admissions to police in interview

- Ready co-operation with authorities

Appendix B

Community order requirements

a. unpaid work requirement

An unpaid work requirement requires an offender to perform unpaid work for a number of hours - not less than 40 and not more than 300 - specified in the order. This is usually to be completed within 12 months but the requirement stays in place until the hours are completed. The offender does not have to consent to the making of the unpaid work requirement but the requirement is subject to availability (Richardson, 2011).

b. activity requirement

An activity requirement requires the offender to present themselves to a person or a place specified on such days, or to participate in activities specified in the order on the number of days. The aggregate number of days the offender must present themselves or participate in activities is 60 days and once again this requirement is subject to availability (Richardson, 2011).

c. programme requirement

A programme requirement requires the offender to take part in an accredited programme at a place or places specified on a number of days specified, or participate

in a number of activities specified in the order for the days specified in the order. The programme will have to have been recommended as being suitable for the offender by probation and that spaces are available. There are no restrictions on the number of days in which they must participate (Richardson, 2011).

d. prohibited activity requirement

A prohibited activity requirement requires the offender refrains from participating in activities or on days or periods specified in the order. This requirement can only be included after consultation with probation. The prohibited activity requirement can include requirements relating to possessing, carrying or using firearms (Richardson, 2011).

e. curfew requirement

A curfew requirement requires the offender to remain, for periods and at places specified in the order. Periods may be between 2-12 hours in any day and the requirement can specify different places or different periods for different days. Curfew requirements must start on the day the order is made and must fall within 6 months of the order being made. The court must consider information about the places and periods specified in the order (Richardson, 2011).

f. exclusion requirement

An exclusion requirement prohibits the offender from entering areas or places specified in the order, during times or periods also specified in the order (Richardson, 2011)

g. residence requirement

The court must consider the home surroundings of the offender and may provide for the offender to reside at a place other than that specified. Only with the

recommendation of probation can a hostel or other institution be specified as their place of residence (Richardson. 2011; Thomas, 2008).

h. mental health treatment requirement

A court may impose this requirement only if it is satisfied with the evidence that has been provided by a medical practitioner, and the offender is susceptible to treatment but does not warrant detention under a hospital order. The offender must express their willingness to comply with the requirement (Richardson, 2011).

i. drug rehabilitation requirement

For a court to impose a drug rehabilitation requirement, they have to be satisfied that the offender is dependent on drugs or has a propensity to misuse drugs; and that their dependency or propensity can be treated. The treatment and testing period must be at least six months but the nature of the treatment is not specified in the order. Once more, this requirement needs to be recommended by probation and the offender must express their willingness to comply with the requirement (Richardson, 2011).

j. alcohol treatment requirement

To impose an alcohol treatment requirement, the court has to be satisfied that the offender is dependent on alcohol; and that this dependency requires, and may be susceptible to treatment. The treatment must not be less than six months (Richardson, 2011).

k. supervision requirement

A supervision requirement requires the offender to addend appointments with a responsible officer and is to help promote offender rehabilitation (Richardson, 2011).

l. attendance centre requirement

An attendance centre requirement is only available for offenders aged under 25. An offender will attend an attendance centre for a total of between 12-36 hours which will be worked around commitments like school or work (Richardson, 2011).

m. electronic monitoring requirement

Where a court makes a community order imposing a curfew requirement or exclusion requirement, the court must also impose an electronic monitoring requirement as defined in the Criminal Justice Act 2000, unless it is prevented from doing so or the circumstances of the case make it inappropriate (Richardson. 2011). The Act also states that a court may not impose an electronic monitoring requirement unless the court has been notified by the Secretary of the State that electronic monitoring arrangements are in place in that area and are satisfied that the necessary provisions can be made under those arrangements.

Furthermore, by imposing an unpaid work requirement, an activity requirement, a programme requirement, a prohibited activity requirement, a residence requirement, a mental health treatment requirement, a drug rehabilitation requirement, an alcohol treatment requirement, a supervision requirement or an attendance centre requirement the court can also impose an electronic monitoring requirement unless prevented by the Criminal Justice Act 2000 from doing so (Richardson, 2011).

Appendix C

Layout of dataset for Goodman RC and extended Goodman RC analysis

Figure C.1 shows an example of the raw data displayed in R prior to the Goodman RC and extended Goodman RC association analyses: each line in the data relates to one offence. The first 11 offences in the dataset are displayed here and show the offences type, the sentence that was received for this offence, the offence plea that was entered for this offence, whether the offender had a prior court appearance and the number of offences the offender was being sentenced for. In the example shown here, the first three offences relating to the offence of Obscene publications relate to the same offender. For each offence the offender received an immediate custodial sentence of between two and three years, they entered a gulity plea for all three of their offences, had no prior court appearances and were sentenced for three of more offences indicated by the variable Numoff. The fourth offence in this example relates to a shoplifting offence and in this case the offender received a community order of less than twelve months. Lines five and six in this example relate to child neglect

offences and the resulting sentence for these offences were fully suspended sentences of between nine and twelve months. Offence seven is burglary in a dwelling and this offender received an immediate custodial sentence of between six and nine months. The last three offences in this example are various driving related offences relating to one offender and the sentences were immediate custodial sentences of varying lengths.

```
> names(DATA)
[1] "OFFENCE"    "SENTENCE"   "Plea"      "Prior"     "Numoff"
> head(DATA,11)
                                    OFFENCE                    SENTENCE       Plea    Prior Numoff
1                      Obscene publications     I> 2 years <= 3 years     Guilty No Prior     >3
2                      Obscene publications     I> 2 years <= 3 years     Guilty No Prior     >3
3                      Obscene publications     I> 2 years <= 3 years     Guilty No Prior     >3
4                               Shoplifting           C <= 12 months     Guilty    Prior     <2
5                       Neglect of children  F> 9 months <= 12 months Not Guilty    Prior     <2
6                       Neglect of children  F> 9 months <= 12 months Not Guilty    Prior     <2
7                    Bulglary in a dwelling   I> 6 months <= 9 months     Guilty    Prior     <2
8  Other criminal damage endangering life  I> 18 months <= 24 months     Guilty    Prior     <2
9                  Driving licence offences   I> 2 months <= 3 months     Guilty No Prior     >3
10                       Drink/drug driving   I> 2 months <= 3 months     Guilty No Prior     >3
11                        Dangerous driving  I> 18 months <= 24 months     Guilty No Prior     >3
>
```

Figure C.1: Layout of raw data prior to tabulation for the Goodman RC and extended Goodman RC analysis

293

Appendix D

Layout of dataset for multilevel modelling

Figure D.1 shows an example of the raw data displayed in R following the extended Goodman RC association analysis and prior to the multilevel modelling analysis. This example of the layout of the raw data is similar to that shown in Appendix C Figure C.1 but now also displays the extended Goodman analysis raw score (Raw_score), the equivalent number of days in custody (obtained through linear interpolation)(Equivdays), the log of the equivalent number of days in custody (Logdays), and the other variables used in the multilevel modelling analysis, such as type of court the offender was sentenced in, their age group and their ethnicity. Again each line relates to one offence and these are the same offences that are shown in Figure C.1, except we now have additional information relating to the offender. For example, lines five and six relate to two offences of child neglect and the sentence for this was a fully suspended sentences of between nine and twelve months. From the information displayed in the dataset, we can see that the offender was sentenced in the magistrates' court in the Probation Trust of London. The offender was a White British female, and this offender was unavailable for work at the time of sentencing and reliant on

state benefits. The offender lived in permanent housing, and was aged between 30 and 40 years old. Finally we know the offender did not use a weapon during the offence but there was a direct victim. Finally we also know that the offender was misusing drugs at the time of her offence.

```
> names(DATA)
 [1] "OFFENCE"   "SENTENCE"  "Plea"      "Prior"     "Numoff"    "Raw_score" "Equivdays" "Logdays"   "Court"     "ProbTrust" "Eth"       "Sex"
[13] "Emp"
> head(DATA, 11)
                                   OFFENCE                  SENTENCE       Plea      Prior Numoff Raw_score Equivdays  Logdays       Court
1                     Obscene publications    I> 2 years <= 3 years     Guilty   No Prior     >3     2.008       998 6.905753       CROWN
2                     Obscene publications    I> 2 years <= 3 years     Guilty   No Prior     >3     2.008       998 6.905753       CROWN
3                     Obscene publications    I> 2 years <= 3 years     Guilty   No Prior     >3     2.008       998 6.905753       CROWN
4                              Shoplifting          C <= 12 months     Guilty      Prior     <2    -4.028        65 4.174387 MAGISTRATES
5                      Neglect of children  F> 9 months <= 12 months Not Guilty      Prior     <2    -0.286       216 5.375278       CROWN
6                      Neglect of children  F> 9 months <= 12 months Not Guilty      Prior     <2    -0.286       216 5.375278       CROWN
7                   Burglary in a dwelling   I> 6 months <= 9 months     Guilty      Prior     <2     0.347       252 5.529429       CROWN
8  Other criminal damage endangering life I> 18 months <= 24 months     Guilty      Prior     <2     5.321       705 6.558198       CROWN
9                 Driving licence offences   I> 2 months <= 3 months     Guilty   No Prior     >3    -3.390        86 4.454347       CROWN
10                      Drink/drug driving   I> 2 months <= 3 months     Guilty   No Prior     >3    -3.390        86 4.454347       CROWN
11                       Dangerous driving I> 18 months <= 24 months     Guilty   No Prior     >3     5.321       705 6.558198       CROWN
                         ProbTrust                    Eth    Sex                  Emp            Inc       Acc   Age Weap Dir Drugs
1                         Cheshire WHITE - British + Irish   MALE     Employed (ft/pt)          Wages Permanent 25-29   NO  NO    No
2                         Cheshire WHITE - British + Irish   MALE     Employed (ft/pt)          Wages Permanent 25-29   NO  NO    No
3                         Cheshire WHITE - British + Irish   MALE     Employed (ft/pt)          Wages Permanent 25-29   NO  NO    No
4  Staffordshire and West Midlands WHITE - British + Irish   MALE           Unemployed State benefits Permanent 30-40   NO  NO   Yes
5                           London WHITE - British + Irish FEMALE Unavailable for work State benefits Permanent 30-40   NO YES   Yes
6                           London WHITE - British + Irish FEMALE Unavailable for work State benefits Permanent 30-40   NO YES   Yes
7  Staffordshire and West Midlands WHITE - British + Irish   MALE           Unemployed State benefits Permanent 25-29   NO  NO   Yes
8               Devon and Cornwall WHITE - British + Irish   MALE           Unemployed          Wages Permanent 30-40  YES YES   Yes
9                       Humberside WHITE - British + Irish   MALE     Employed (ft/pt)          Wages Permanent 25-29   NO  NO    No
10                      Humberside WHITE - British + Irish   MALE     Employed (ft/pt)          Wages Permanent 25-29   NO  NO    No
11                      Humberside WHITE - British + Irish   MALE     Employed (ft/pt)          Wages Permanent 25-29   NO  NO    No
>
```

Figure D.1: Layout of raw data prior to the multilevel modelling analysis

Appendix E

Multilevel modelling offences

The 99 offence categories from Model 24 are shown below. The offence of murder is the reference category which is why the coefficient and multiplicative effect is 0.00 and 1.000. The other offence categories are then listed below the reference category.

Table E.1: Multilevel modelling offences

Offence type	Coefficient (95% CI)	Multiplicative effect
Murder (reference category)	0.00	1.0000
	(-,-)	
Attempted murder	-1.657636	0.1906
	(-3.24, -0.08)	
Threats, Conspiracy or Incitement to murder	-2.751341	0.0638
	(-4.16, -1.34)	
Manslaughter	-1.442210	0.2364
	(-2.86, -0.02)	
Wounding and other acts endangering life	-1.661499	0.1899
	(-3.07, -0.26)	

Continued on next page

Offence type	Coefficient (95% CI)	Multiplicative effect
Malicious wounding and other like offences	-3.066134 (-4.47, -1.66)	0.0466
Cruelty to or neglect of children	-3.104493 (-4.51, -1.70)	0.0448
Abduction of a child	-2.095258 (-3.56, -0.63)	0.1230
Buggery or attempted buggery[*]	-1.572239 (-3.02, -0.12)	0.2076
Indecent assault on a male	-2.181715 (-3.59, -0.77)	0.1128
Indecency between males[*]	-2.905657 (-4.34, -1.48)	0.0547
Rape of male or female	-1.122453 (-2.53, 0.28)	0.3255
Indecent assault on a female	-2.152085 (-3.55, -0.75)	0.1162
Unlawful sexual intercourse with girl under 13	-1.803819 (-3.21, -0.40)	0.1647
Unlawful sexual intercourse with girl under 16	-2.051623 (-3.46, -0.65)	0.1285
Incest with girl under 13 years old	-2.069645 (-3.49, -0.64)	0.1262
Procuration of males and females	-2.418245	0.0891

Offence type	Coefficient	Multiplicative effect
	(95% CI)	
	(-3.93, -0.91)	
Burglary in a dwelling	-2.344578	0.0959
	(-3.75, -0.94)	
Aggravated burglary in a dwelling	-1.346778	0.2601
(including attempts)	(-2.78, 0.09)	
Burglary other than a dwelling	-3.106236	0.0448
	(-4.51, -1.70)	
Going equipped for stealing	-3.543753	0.0289
	(-4.95, -2.14)	
Robbery and assaults with intent to rob	-1.808225	0.1639
	(-3.21, -0.41)	
Blackmail	-2.487347	0.0831
	(-3.93, -1.04)	
Kidnapping	-1.846456	0.1578
	(-3.26, -0.44)	
Aggravated taking of a vehicle	-2.963377	0.0516
	(-4.37, -1.56)	
Concealing	-3.227058	0.0397
	(-4.63, -1.82)	
Stealing from another person	-3.348745	0.0351
	(-4.75, -1.94)	
Stealing in a dwelling other than from	-3.334129	0.0356
automatic machines and meters	(-4.74, -1.92)	

Offence type	Coefficient	Multiplicative effect
	(95% CI)	
Stealing by an employee	-3.229936	0.0396
	(-4.64, -1.82)	
Unlawfully taking away or opening a	-2.622229	0.0726
mail bag	(-4.11, -1.14)	
Abstracting electricity	-3.709321	0.0245
	(-5.13, -2.28)	
Stealing pedal cycles	-3.800126	0.0224
	(-5.22, -2.38)	
Stealing from vehicles	-3.431051	0.0324
	(-4.84, -2.03)	
Stealing from shops and stalls	-3.634850	0.0264
(shoplifting)	(-5.04, -2.23)	
Stealing from automatic machines	-3.614099	0.0269
	(-5.05, -2.18)	
Theft of motor vehicle	-2.970288	0.0513
	(-4.38, -1.56)	
Other stealing and unauthorised takings	-3.373628	0.0343
	(-4.78, -1.97)	
Frauds by agents, trustees, company	-1.983610	0.1376
directors etc.	(-3.59, -0.38)	
False accounting	-3.241391	0.0391
	(-4.66, -1.82)	
Other frauds	-3.261729	0.0383

Offence type	Coefficient (95% CI)	Multiplicative effect
	(-4.66, -1.86)	
Receiving/handling stolen goods	-3.273109	0.0379
	(-4.68, -1.87)	
Bankruptcy	-2.880239	0.0561
	(-4.36, -1.40)	
Arson	-2.436147	0.0875
	(-3.85, -1.03)	
Other criminal damage endangering life	-2.052242	0.1284
	(-3.54, -0.57)	
Other criminal damage	-3.714285	0.0244
	(-5.12, -2.31)	
Threat or possession with intent to commit criminal damage	-3.194524	0.0410
	(-4.61, -1.78)	
Forgery of prescription	-3.443108	0.0320
	(-5.05, -1.84)	
Other Forgery	-3.088323	0.0456
	(-4.49, -1.68)	
Public Order Act 1986 Sec 2 Violent Disorder	-2.443852	0.0868
	(-3.87, -1.02)	
Other offences (against the State and Public Order)	-3.200686	0.0407
	(-4.60, -1.80)	
Perjury	-3.590329	0.0276
	(-5.16, -2.02)	

Offence type	Coefficient	Multiplicative effect
	(95% CI)	
Intending to facilitate prison escape	-3.255067	0.0386
	(-4.75, -1.76)	
Child prostitution	-1.743845	0.1748
	(-3.30, -0.19)	
Abuse of trust - Sexual offences	-2.623002	0.0726
	(-4.06, -1.18)	
Gross indecency with a child	-2.571078	0.0765
	(-3.98, -1.17)	
Immigration Act including people	-2.812441	0.0601
trafficking	(-4.26, -1.36)	
Attempting to pervert the course of	-3.326293	0.0359
justice	(-4.73, -1.92)	
Absconding from lawful custody	-3.535287	0.0292
	(-4.97, -2.11)	
Firearms offences	-2.933616	0.0532
	(-4.34, -1.53)	
Offences against laws relating to	-3.129509	0.0437
Customs, Excise and Inland Revenue	(-4.68, -1.58)	
Bail offences	-4.098861	0.0166
	(-5.50, -2.70)	
Trade Descriptions and similar offences	-3.189278	0.0412
	(-4.61, -1.77)	
Obscene publications	-2.841431	0.0583

Offence type	Coefficient	Multiplicative effect
	(95% CI)	
	(-4.24, -1.44)	
Sexual Offences Act 2003 including	-2.624689	0.0725
grooming and other sexual acts	(-4.03, -1.22)	
Misuse of Drugs Act	-3.176108	4.99
.0417	(-4.58, -1.77)	
Permitting premises to be uses for	-3.370170	0.0344
unlawful (drug-related) purposes	(-4.79, -1.95)	
Other indictable offences	-3.510796	0.0299
	(-4.92, -2.10)	
Assault on a constable	-3.607455	0.0271
	(-5.01, -2.21)	
Common and other types of assault	-3.466341	0.0312
	(-4.87, -2.06)	
Cruelty to animals, ill treatment, neglect	-3.417375	0.0328
etc.	(-4.83, -2.00)	
Offences in relation to dogs Dangerous	-2.934109	0.0532
Dogs Act 1991	(-4.47, -1.40)	
Preventing child from receiving information	-3.475661	0.0309
	(-4.93, -2.02)	
Firearms Act 1968	-3.510114	0.0299
	(-4.95, -2.07)	
Offences against Public Order	-3.599987	0.0273
	(-5.00, -2.20)	

Offence type	Coefficient	Multiplicative effect
	(95% CI)	
Interference with a motor vehicle	-3.602765	0.0272
	(-5.01, -2.19)	
Stealing and unauthorised takings	-3.604700	0.0272
	(-5.01, -2.20)	
Aggravated vehicle taking	-3.177177	0.0417
	(-4.59, -1.77)	
Drunk and disorderly	-3.836318	0.0216
	(-5.25, -2.43)	
Licensing Act offences	-4.707037	0.0090
	(-6.48, -2.93)	
Criminal Damage	-3.684377	0.0251
	(-5.09, -2.28)	
Benefit offences	-3.492830	0.0304
	(-4.91, -2.08)	
Offences relating to Police regulations	-4.252614	0.0142
	(-5.75, -2.76)	
Public health - By-laws	-3.908706	0.0201
	(-5.44, -2.38)	
Transport offences	-3.251573	0.0387
	(-5.03, -1.48)	
Pick-lock related offences	-3.769305	0.0231
	(-5.23, -2.31)	
Misuse of drug related offences	-3.851218	0.0213

Offence type	Coefficient (95% CI)	Multiplicative effect
	(-5.26, -2.44)	
Other offences excluding motoring offences	-3.468924	0.0312
	(-4.87, -2.07)	
Improper use of public electronic	-3.453308	0.0316
communications network	(-4.87, -2.04)	
Dangerous Driving	-2.995642	0.0500
	(-4.40, -1.59)	
Driving or attempting to drive after	-3.506415	0.0300
consuming alcohol or drugs	(-4.91, -2.10)	
Careless driving	-3.847321	0.0213
	(-5.27, -2.43)	
Accident offence	-3.684747	0.0251
	(-5.09, -2.28)	
Driving license offences	-3.483969	0.0307
	(-4.89, -2.08)	
Vehicle insurance offences	-3.657077	0.0258
	(-5.06, -2.25)	
Vehicle Test Offences	-3.753196	0.0234
	(-5.20, -2.30)	
Fraud, Forgery associated with vehicle or	-3.809534	0.0222
driving records	(-5.30, -2.32)	
Speed limit offences	-3.696433	0.0248
	(-5.23, -2.16)	

Table E.1-*continued from previous page*

Offence type	Coefficient	Multiplicative effect
	(95% CI)	
Road traffic offence, neglect, directions etc.	-3.727891	0.0240
	(-5.17, -2.28)	

[*]These offences have now been re-classified as male rape

Appendix F

Sample code

A sample of the R code used to create the extended Goodman RC association model and optimal multilevel model (Model24) is shown below along with commentary explaining the syntax. Each new line of code is indicated by a > symbol and the commentary follows after the # symbol.

Extended Goodman RC association sample code

>`library(foreign)` # allows files to be imported from SPSS

>`upgradecase=read.spss("LONG1.sav",to.data.frame=T)` # upload SPSS file into R

>`upgtable=xtabs(∼ SENTENCE+OFFENCE+SENT+ PLEAGRP+CONV+NUMOFF2,data=upgradecase)` # create table 99x35x12

>`upgradefreq=as.data.frame(upgtable)` # convert table into frequency form

>`library(gnm)` # allows the gnm package to be used

>`ExtRC=gnm(Freq SENTENCE+OFFENCE+Mult((SENTENCE+PLEAGRP+CONV+ NUMOFF2),OFFENCE)+PLEAGRP+NUMOFF2+CONV,family=poisson, data=upgradefreq)` # fits extended Goodman RC model

Multilevel modelling sample code

>`library(foreign)` # allows files to be imported from SPSS

>`DATA=read.spss("New_SPSS_2.sav",to.data.frame=T)` # upload SPSS file into R

>`DATA=na.omit(DATA)` # omit missing data to carry out a complete case analysis

>`library(lme4)` # allows the lme4 package to be used

>`Model24 =lmer(Logdays` ∼ `OFFrecode+PLEAGRP+NUMOFF2+COURT+EMP+`

`ACC+INCOME+SEX+ETHNICITY+AGE+(1|SentOcc),data=DATA, REML=FALSE)`

fits optimal multilevel model, estimates chosen using maximum likelihood

>`print(Model24, correlation=FALSE)` # prints output

>`Model24CI=profile(Model24)` # calculates grid of profile likelihood values around maximum likelihood estimates

>`CI.fit24=confint(Model24CI)` # extracts confidence intervals from profile likelihoods

>`print(CI.fit24)` # print confidence intervals