**SENTENCING CONSISTENCY
IN THE
NEW ZEALAND DISTRICT COURTS**

By

Wayne Goodall

A thesis

**Submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Doctor of Philosophy
In Criminology**

Victoria University of Wellington

2014

Abstract

This thesis examines the consistency of sentencing between the circuits of the New Zealand District Courts. Four predictions based on a sequence or chain of theories incorporating the concept of bounded rationality from decision making theory, the influence of formal and substantive rationalities on sentencing decisions, court community theory, and personal construct psychology were tested. The circuit in which sentencing took place was expected to affect the likelihood of incarceration and to affect the length of incarceration. If these predictions were met, it was further predicted that the weight applied to some or all of the sentencing factors would vary between circuits. It is understood to be the first study controlling for a wide range of sentencing factors examining the consistency of sentencing between locations in New Zealand and one of the first from anywhere outside of the United States.

The four predictions were tested using sentencing data from the two years 2008-2009 for three high volume offences (aggravated drink driving, male assaults female and burglary). Sentencing was treated as a two part decision process, with the selection of a sentence type followed by the determination of the sentence amount. Each prediction was separately modelled for each offence. Different types of model were chosen as being more suitable for the specific predictions: logistic regression for the likelihood of incarceration; linear regression for the length of incarceration; multi-level generalised linear regression with random co-efficients to determine if the weight applied to specific factors varied by circuit in the determination of whether or not to incarcerate; and multi-level linear regression with random co-efficients to determine if the weight applied to specific factors varied by circuit in the determination of sentence length.

The logistic regression and linear regression models demonstrated that there were statistically significant and substantively significant differences between circuits in the likelihood and length of incarceration. The extent of inconsistency varied by offence type with the most marked differences occurring for aggravated drink driving and burglary. Offence seriousness and criminal history factors were found to be the principal influences on both sentence decisions for all three offences. The multi-level models for aggravated drink driving and burglary revealed a core of seriousness and criminal history factors whose influence varied across the circuits. The models for male assaults female were less informative, highlighting the likelihood that these models were limited by the omission of

key sentencing variables and the narrow scope of this particular assault type within the wider spectrum of assaults.

The findings should not be interpreted as if they are a critique of the sentence imposed in any individual case or of the sentencing by any judge or in any circuit. It is a critique of sentencing guidance in New Zealand and its ability to achieve a fundamental tenet of justice: the similar treatment of similar offenders being sentenced in similar circumstances.

In addition to testing the predictions the multi-level models were extended to address whether the observed variation in sentencing was associated with variations in circuit context. Due to the limited number of circuits (17) and multi-collinearity between the contextual variables, bivariate analyses had to be employed. The modelling revealed a consistent difference between provincial and metropolitan circuits; the two categories of circuit were distinguished from one another by many of the other more specific variables that had a significant association with sentencing approaches. The provincial circuits were more likely to incarcerate and to impose longer sentences. However, the small number of circuits and multi-collinearity between the variables precluded more detailed analysis to identify which of the contextual variables distinguishing metropolitan and provincial circuits had the greatest influence.

These findings have significant implications for the judiciary and for sentencing policy makers. Urgent attention should be given to addressing opportunities to increase the availability of sentencing guidance to reduce the degree of inconsistency. More detailed offence based sentencing guidance is required; in the current context there are two options that could be used. The Court of Appeal could issue a broader range of guideline judgments or the legislation for the Sentencing Council and the process for developing and promulgating guidelines could be implemented. For logistical and public policy reasons the Sentencing Council approach is preferred. There is a risk that failure to address the levels of inconsistency will result in the sentencing system falling into disrepute.

Acknowledgements

I owe thanks and acknowledge debts to those who have supported my endeavours. First, to my parents Margaret and Snow who recognised the importance of education and free-thought. I owe a huge debt to my wife Debbie Iversen and daughters Anna and Ruth who supported me throughout. They are the ones who saw the ups and downs most often and kept me going. Thanks too, Anna and Ruth, for the data coding and clerical assistance you provided. Without the assistance of James Wade (from the Ministry of Justice), Roger Macky and Jason Gleason I would not have been able to complete this research. They had the unenviable task of translating my specifications into an extract of raw data and transforming it into meaningful observations that I could manipulate and use to model sentencing decisions. To my friend Charles Sullivan, thanks for increasing my understanding of more advanced statistical modelling and acting as a sounding board when I needed to test ideas. This research would not have been possible without the support of the New Zealand Judiciary who gave approval for the provision of data. Thanks to the Law Commission who provided me with accommodation and the staff of the Commission for their friendship and forbearance. I reserve my greatest and everlasting thanks to Dr Russil Durrant and Dr Warren Young, my supervisors in this endeavour, and without whose help I could not have completed this thesis.

Table of Contents

	Page
Abstract	iii
Acknowledgements	v
List of Tables	xi
List of Figures	xix
Chapter 1: Sentencing Consistency In New Zealand: An Unresolved Debate	21
Chapter 2: A Framework for Understanding Sentencing Disparity	29
Exercising Discretion: The Sentencing Challenge	31
Formal and Substantive Rationalities and Court Communities	37
Summary of the Framework	43
Applying the Framework to New Zealand	44
Concluding Comment	48
Chapter 3: Sentencing Disparity by Location, the Findings from Empirical Research	49
United States Research	50
Research from Outside the United States	65
Summary	69
Chapter 4: Predictions, Research Questions & Statistical Methods	71
Predictions	72
Research Questions	74
Two Stage Modelling Strategy	75
Assembling and Interpreting the Models	77

Chapter 5: Scope, Data Sources & Variables	85
The Scope of the Sentencing Models	85
The Data	95
The Variables Modelled	99
 Chapter 6: Consistency between Circuits in the Use of Incarceration	 107
Aggravated Drink Driving: Descriptive Statistics	108
Aggravated Drink Driving: In-out Model	113
Male Assaults Female: Descriptive Statistics	119
Male Assaults Female: In-out Model	125
Burglary: Descriptive Statistics	131
Burglary: In-out Model	137
Residual Statistics, Influential Cases and Outliers	143
Discussion of the Model Results	145
 Chapter 7: Consistency between Circuits in the Length of Incarceration	 151
Aggravated Drink Driving: Descriptive Statistics	152
Aggravated Drink Driving: Sentence Length Model	156
Male Assaults Female: Descriptive Statistics	159
Male Assaults Female: Sentence Length Model	163
Burglary: Descriptive Statistics	167
Burglary: Sentence Length Model	171
Residual Statistics, Influential Cases and Outliers	177
Discussion of the Model Results	179
 Chapter 8: The Interpretation and Use of Variables between Circuits	 185
Development of the Models	185
Aggravated Drink Driving: In-out Model	187
Aggravated Drink Driving: Sentence Length Model	190
Male Assaults Female: In-out Model	192
Male Assaults Female: Sentence Length Model	194
Burglary: In-out Model	196
Burglary: Sentence Length Model	199
Summary	201

Chapter 9: The Influence of Circuit Context on Sentencing Outcomes	203
Development of the Models	204
Descriptive Statistics	205
The Multi-level Models	209
Summary	218
 Chapter 10: Summary of Findings, Discussion & Conclusion	 219
Inconsistency of Sentencing between Circuits	223
Variation between Circuits in the Interpretation and Use of Sentencing Factors	226
Association between Circuit Context and Inconsistency of Sentencing	229
Sentencing Heuristics	230
The Policy Implications	231
Future Research	234
Conclusion	237
 Appendices	 239
Appendix A: Sentencing Variability Associated with Socio-Demographic Factors	241
Appendix B: Comparison of Parameter Estimates for Sentencing Variables in Single Level Regressions Models and Multi-level Random Intercept Models	251
Appendix C: Sentence Notes as a Data Source	259
Appendix D: Summary of the Tests Employed to Determine Primary Offence Status	263
Appendix E: The Predictor Variables Included in the Modelling	265
Appendix F: The Impact of Cumulative Sentencing and Home Detention on the Aggravated Drink Driving and Male Assaults Female Sentence Length Models	275
Appendix G: Correlation Matrices for Circuit Contextual Variables for Multi-level Modelling	283
Appendix H: Using the Sentence Notes to Identify Omitted Variables	285
 References	 295

List of Tables

	Page
Table 3.1: United States sentencing research actively addressing the influence of location	51
Table 3.2: United States sentencing research with controls for location but that did not actively address the influence of location	52
Table 5.1: Mean annual sentencing volumes for selected property/dishonesty and drug offences 2004-2006	88
Table 5.2: Mean annual sentencing volumes for selected violent and sexual offences 2004 – 2006	89
Table 5.3: Mean annual sentencing volumes for selected serious driving offences 2004-2006	90
Table 6.1: Descriptive statistics for the demographic variables for aggravated drink drivers sentenced in 2008-2009 ($n = 9,282$)	108
Table 6.2: Descriptive statistics for the circuit variable for aggravated drink drivers sentenced in 2008-2009 ($n = 9,282$)	109
Table 6.3: Descriptive statistics for the offence seriousness variables for aggravated drink drivers sentenced in 2008-2009 ($n = 9,282$)	110
Table 6.4: Descriptive statistics for the case processing variables for aggravated drink drivers sentenced in 2008-2009 ($n = 9,282$)	111
Table 6.5: Descriptive statistics for the criminal history variables for aggravated drink drivers sentenced in 2008-2009 ($n = 9,282$)	112
Table 6.6: Logistic regression parameter estimates for the offence seriousness, criminal history and offender demographic variables for the in-out model for aggravated drink drivers sentenced in 2008-2009 ($n = 9282$)	115
Table 6.7; Logistic regression parameter estimates for the circuit dummy variables for the in-out model for aggravated drink drivers sentenced in 2008-2009 ($n = 9,282$)	117
Table 6.8: Descriptive statistics for the demographic variables for male assaults female offenders sentenced in 2008-2009 ($n = 6,166$)	119
Table 6.9: Descriptive statistics for the circuit variable for male assaults female offenders sentenced in 2008-2009 ($n = 6,166$)	120

Table 6.10: Descriptive statistics for the offence seriousness variables for male assaults female offending sentenced in 2008-2009 ($n = 5,166$)	121
Table 6.11: Descriptive statistics for the case processing variables for male assaults female offending sentenced in 2008-2009 ($n = 6,166$)	123
Table 6.12: Descriptive statistics for the criminal history variables for male assaults female offenders sentenced in 2008-2009 ($n = 6,166$)	124
Table 6.13: Logistic regression parameter estimates for the offence seriousness, case processing, criminal history and offender demographic variables for the in-out model for male assaults female offenders sentenced in 2008-2009 ($n = 6,166$)	126
Table 6.14: Logistic regression parameter estimates for the circuit dummy variables for the in-out model for male assaults female offenders sentenced in 2008-2009 ($n = 6,166$)	128
Table 6.15: Descriptive statistics for the demographic variables for burglars ($n = 5,740$)	131
Table 6.16: Descriptive statistics for the circuit variable for burglars sentenced in 2008-2009 ($n = 5,740$)	132
Table 6.17: Descriptive statistics for the offence seriousness variables for burglary included in the final model ($n = 5,740$)	133
Table 6.18: Descriptive statistics for the offence seriousness variables for burglary not included in the final model ($n = 5,740$)	134
Table 6.19: Descriptive statistics for the case processing variables for burglary ($n = 5,740$)	135
Table 6.20: Descriptive statistics for the criminal history variables for burglars ($n = 5,740$)	136
Table 6.21: Logistic regression parameter estimates for the offence seriousness, case processing, criminal history and offender demographic variables for the in-out model for burglars sentenced in 2008-2009 ($n = 5,740$)	139
Table 6.22: Logistic regression parameter estimates for the circuit dummy variables for burglary offenders sentenced in 2008-2009 ($n = 5,740$)	141
Table 6.23: Order of circuits by odds ratio from highest to lowest by offence type	149
Table 7.1: Demographic profile of offenders incarcerated for aggravated drink driving in 2008-2009 ($n = 2,312$)	152

Table 7.2: Number of offenders incarcerated for aggravated drink driving in 2008-2009 by circuit ($n = 2,312$)	153
Table 7.3: Descriptive statistics for offence seriousness variables linked to offenders incarcerated for aggravated drink driving in 2008-2009 ($n = 2,312$)	154
Table 7.4: Descriptive statistics for case processing variables linked to offenders incarcerated for aggravated drink driving in 2008-2009 ($n = 2,312$)	154
Table 7.5: Descriptive Statistics for criminal history variables linked to offenders incarcerated for aggravated drink driving in 2008-2009 ($n = 2,312$)	155
Table 7.6: Linear regression parameter estimates for the criminal history, offence seriousness, case processing and demographic variables for the sentence length model for aggravated drink drivers sentenced in 2008-2009 ($n = 2,312$)	157
Table 7.7: Linear regression parameter estimates for the circuit variables for the sentence length model for aggravated drink drivers sentenced in 2008-2009 ($n = 2,312$)	159
Table 7.8: Demographic profile of offenders incarcerated for male assaults female in 2008-2009 ($n = 723$)	160
Table 7.9: Number of offenders incarcerated for male assaults female in 2008-2009 by circuit ($n = 723$)	160
Table 7.10: Descriptive statistics for offence seriousness factors linked to offenders incarcerated for male assaults female in 2008-2009 by circuit ($n = 723$)	161
Table 7.11: Descriptive statistics for case processing variables linked to offenders incarcerated for male assaults female in 2008-2009 by circuit ($n = 723$)	162
Table 7.12: Descriptive statistics for criminal history variables linked to offenders incarcerated for male assaults female in 2008-2009 by circuit ($n = 723$)	163
Table 7.13: Linear regression parameter estimates for the criminal history, offence seriousness, case processing and demographic variables for the sentence length model for male assaults female offenders sentenced in 2008-2009 ($n = 722$)	165
Table 7.14: Linear regression parameter estimates for the circuit variables for the sentence length model for male assaults female offenders sentenced in 2008-2009 ($n = 722$)	166

Table 7.15: Demographic profile of offenders incarcerated for burglary in 2008-2009 ($n = 2,338$)	167
Table 7.16: Number of offenders incarcerated for burglary in 2008-2009 by circuit ($n = 2,338$)	168
Table 7.17: Descriptive statistics for offence seriousness variables linked to offenders incarcerated for burglary in 2008-2009 by circuit ($n = 2,338$)	169
Table 7.18: Descriptive statistics for case processing variables linked to offenders incarcerated for burglary in 2008-2009 by circuit ($n = 2,338$)	170
Table 7.19: Descriptive statistics for criminal history variables linked to offenders incarcerated for burglary in 2008-2009 by circuit ($n = 2,338$)	171
Table 7.20: Linear regression parameter estimates for the criminal history, offence seriousness, case processing and demographic variables for the sentence length model for burglary offenders sentenced in 2008-2009 ($n = 2,338$)	173
Table 7.21: Approximate effects of offence seriousness and case processing variables for burglary found to increase sentence length	174
Table 7.22: Linear regression parameter estimates for the circuit variables for the sentence length model for burglary offenders sentenced in 2008-2009 ($n = 2,338$)	176
Table 7.23: Order of circuits by co-efficient from largest to smallest by offence type	182
Table 7.24: Order of circuits by odds ratio from highest to lowest and by co-efficient (sentence length) from longest to shortest by offence type	183
Table 8.1: Aggravated drink driving in-out model simplification of case level variables	187
Table 8.2: Case level random co-efficient model for aggravated drink driving in-out decision fixed effects component (for random effects see Table 8.3) ($n = 9,282$)	188
Table 8.3: Case level random co-efficient model for aggravated drink driving in-out decision random effects component (for fixed effects see Table 8.2) ($n = 9,282$)	189
Table 8.4: Case level random co-efficient model for aggravated drink driving sentence length decision fixed effects component (for random effects see Table 8.5) ($n = 2,312$)	191
Table 8.5: Case level random co-efficient model for aggravated drink driving sentence length decision random effects component (for fixed effects see Table 8.4) ($n = 2,312$)	191

Table 8.6: Male assaults female in-out model simplification of case level variables	192
Table 8.7: Case level random co-efficient model for the male assaults female in-out decision fixed effects component (for random effects see Table 8.8) ($n = 6,166$)	193
Table 8.8: Case level random co-efficient model for the male assaults female in-out decision random effects component (for fixed effects see Table 8.7) ($n = 6,166$)	194
Table 8.9: Case level random co-efficient model for the male assaults female sentence length decision fixed effects component (for random effects see Table 8.10) ($n = 722$)	195
Table 8.10: Case level random co-efficient model for the male assaults female sentence length decision random effects component (for fixed effects see Table 8.9) ($n = 722$)	195
Table 8.11: Burglary in-out model simplification of case level variables	196
Table 8.12: Case level random co-efficient model for the burglary in-out decision fixed effects component (for random effects see Table 8.13) ($n = 5,740$)	198
Table 8.13: Case level random co-efficient model for the burglary in-out decision random effects component (for fixed effects see Table 8.12) ($n = 5,740$)	199
Table 8.14: Case level random co-efficient model for the burglary sentence length decision fixed effects component (for random effects see Table 8.15) ($n = 2,338$)	200
Table 8.15: Case level random co-efficient model for the burglary sentence length decision random effects component (for fixed effects see Table 8.14) ($n = 2,338$)	201
Table 9.1: Generic demographic, socio-economic and caseload and caseflow contextual variables	205
Table 9.2: Offence specific caseload and caseflow contextual variables	207
Table 9.3: Means and standard deviations for the metropolitan and provincial groupings for the generic contextual variables	208
Table 9.4: Means and standard deviations for the metropolitan and provincial groupings for the offence specific contextual variables	208
Table 9.5: Parameter estimates for circuit level variables individually added to the case level model for the in-out decision for aggravated drink drive sentencing	210

Table 9.6: Odds ratios one standard deviation from the mean for circuit level variables individually added to the case level model for the in-out decision for aggravated drink drive sentencing	211
Table 9.7: Parameter estimates for circuit level variables individually added to the case level model for the in-out decision for burglary drink drive sentencing	213
Table 9.8: Odds ratios one standard deviation from the mean for circuit level variables individually added to the case level model for the in-out decision for burglary drink drive sentencing	214
Table B1: Comparison of parameter estimates for logistic regression and hierarchical generalised linear models for aggravated drink driving ($n = 9,282$)	252
Table B2: Comparison of parameter estimates for linear regression and hierarchical linear models for aggravated drink driving ($n = 2,312$)	253
Table B3: Comparison of parameter estimates for logistic regression and hierarchical generalised linear models for male assaults female ($n = 6,166$)	254
Table B4: Comparison of parameter estimates for linear regression and hierarchical linear models for male assaults female ($n = 722$)	255
Table B5: Comparison of parameter estimates for logistic regression and hierarchical generalised linear models for burglary ($n = 5,740$)	256
Table B6: Comparison of parameter estimates for linear regression and hierarchical linear models for burglary ($n = 2,338$)	257
Table C1: The number of cases and mean sentence lengths by circuit for the administrative and sentence note datasets for aggravated drink driving	260
Table C2: The number of cases and mean sentence lengths by circuit for the administrative and sentence note datasets for male assaults female	261
Table C3: The number of cases and mean sentence lengths by circuit for the administrative and sentence note datasets for burglary	262
Table E1: Demographic variables and the circuit variable	265
Table E2: Drink driving offence seriousness, case processing and criminal history variables	266
Table E3: Male assaults female offence seriousness, case processing and criminal history variables	268
Table E4: Burglary offence seriousness, case processing and criminal history variables	270

Table F1: Relative proportions of imprisonment and home detention sentences for aggravated drink driving by categories of sentence length	276
Table F2: Relative proportions of imprisonment and home detention sentences for male assaults female by categories of sentence length	277
Table F3: The proportion of incarcerated aggravated drink drivers sentenced to home detention or who received a cumulative sentence of incarceration	278
Table F4: The proportion of offenders incarcerated for male assaults female sentenced to home detention or who received a cumulative sentence of incarceration	279
Table G1: Correlation matrix for the generic contextual variables	283
Table G2: Correlation matrix for the offence specific contextual variables and including the correlations between the generic variables and each offence specific variable	284
Table H1: Descriptive statistics for the sentence length decision sentencing variables sourced from sentence notes for aggravated drink driving ($n = 1,308$)	285
Table H2: Parameter estimates for the statistically significant sentence note variables added to the sentence length model for aggravated drink driving ($n = 1,308$)	286
Table H3: Descriptive statistics for the sentence length decision sentencing variables sourced from sentence notes for male assaults female ($n = 415$)	287
Table H4: Parameter estimates for the statistically significant sentence note variables added to the sentence length model for male assaults female ($n = 415$)	289
Table H5: Comparison of the regression co-efficients for the circuit variables for models including and excluding sentence note variables for male assaults female	290
Table H6: Descriptive statistics for the sentence length decision sentencing variables sourced from sentence notes for burglary ($n = 1,386$)	291
Table H7: Parameter estimates for the statistically significant sentence note variables added to the sentence length model for burglary ($n = 1,386$)	293
Table H8: Comparison of the regression co-efficients for the circuit variables for models including and excluding sentence note variables for burglary	294

List of Figures

	Page
Figure 6.1: Odds ratios and 95% confidence intervals for incarceration of aggravated drink drivers by circuit (Circuit 0 is the reference)	117
Figure 6.2: Relative risk of incarceration for aggravated drink driving by circuit	118
Figure 6.3: Odds ratios and confidence intervals for incarceration of male assaults female offenders sentenced in 2008-2009 by circuit (Circuit 0 is the reference)	129
Figure 6.4: Relative risk of incarceration for male assaults female offenders sentenced in 2008-2009 by circuit	130
Figure 6.5: Odds ratios for incarceration of burglary offenders sentenced in 2008-2009 by circuit (Circuit 0 is the reference)	142
Figure 6.6: Relative risk of incarceration for burglary offenders sentenced in 2008-2009 by circuit	143

Chapter 1: Sentencing Consistency in New Zealand: An Unresolved Debate

For a short time in 2006 and 2007 sentencing consistency between regions of the New Zealand District Courts was the subject of discussion and debate. This was spurred by a report from the New Zealand Law Commission. In 2006 the Law Commission recommended that “a Sentencing Council should be established with a mandate to draft sentencing guidelines” (Law Commission, 2006, p. 13). While there were multiple arguments raised by the Law Commission in favour of the recommendation, one substantive argument raised issues about the fairness of sentencing. Informed by a comparison of sentencing decisions between regions the Law Commission claimed “there is significant inconsistency in the sentences imposed between judges and between courts, particularly in relation to lower end offences” (Law Commission, 2006, p. 13). The Law Commission attributed the variations to the wide discretion entrusted to judges and the inadequacy of appellate review and other guidance as an effective check on the exercise of the discretion.

The recommendation was accepted by the Government and a Bill to establish a sentencing council was introduced but it was not universally supported. Despite the opposition the Sentencing Council Act was assented in July 2007 and came into force on 1 November 2007. However, following a change of government in 2008 work on the establishment of the Council and the development of guidelines was halted. Subsequently there has been little interest in the consistency of sentencing.

Opposition to the introduction of guidelines was founded on two arguments: rejection of the notion that the existing sentencing framework was inadequate; and doubts about the claimed inconsistency, which was dismissed as being the result of differences in the composition of the cases between locations. In Parliament opponents of the legislation argued that: the Sentencing Act provides a comprehensive guide to judges on the factors to be considered in sentencing; the appeal regime provides an opportunity for any necessary correction; and the accrued body of appeal decisions, especially guideline judgments from the Court of Appeal, provide an additional more specific framework guiding the severity of sentence (New Zealand Parliament, 2006 & 2007).

Doubts were raised about the degree to which the perception of inconsistency and the comparison of sentencing decisions constituted proof of real inconsistency. These doubts were succinctly expressed in a speech by Justice Priestley after the decision to discontinue work on the Council and guidelines, he argued that: “Sentencing guidelines

would undoubtedly have curbed the discretion of sentencing judges to do justice in the infinite variety of cases over which they preside” (Priestley, 2009, p.34).

A substantial body of literature addressing sentencing consistency (or variation or disparity) has been assembled in the United States but little work has been done elsewhere (Ulmer, 2012) and little work has been done in New Zealand. The lack of research interest in sentencing consistency in New Zealand has not been matched by a lack of attention to it in sentencing policy. For more than two decades the Court of Appeal has issued guideline judgments as a means of attempting to achieve greater consistency in the sentencing of certain offences. In 2002 the Sentencing Act consolidating sentencing legislation and codifying the case law in a single statute was passed. For the first time sentencing was informed by a cogent and comprehensive legislative framework enhancing the likelihood of greater consistency of outcomes through greater consistency of process and approach. Much of the framework came from the case law; the principal change was to draw it together in a single statute. The Sentencing Act 2002 explicitly incorporated consistency into the framework as one of the principles that judges must take into account when sentencing. Section 8(e) of the Act requires judges to “take into account the general desirability of consistency with appropriate sentencing levels and other means of dealing with offenders in respect of similar offenders committing similar offences in similar circumstances”. A review of the Sentencing Act in 2004 described the consistency principle as according with the prevailing general expectations of fairness and equality before the law (Chhana, Spier, Roberts & Hurd, 2004). Earlier, Hall (1991) described consistency as inherent in the rule of law and argued that consistency must be assumed in the interests of fairness and justice. Ashworth (1998) described inconsistency as a manifest injustice capable of bringing a sentencing system into public disrepute.

The policy aim of this research was to advance the uncompleted debate on the adequacy of the sentencing framework and the extent of inconsistency. The debate on adequacy was advanced by examining how inconsistency can occur despite the provision of guidance such as that provided by the Sentencing Act and appellate authorities, and the corrective effects of the appeal regime on individual cases. The extent of inconsistency was addressed by an empirical examination of sentencing decisions in the New Zealand District Courts in 2008-2009. Regression models based on a specially assembled dataset of sentencing decisions by District Court Judges resident in the circuits and a set of contextual variables based on census data and Ministry of Justice caseload and sitting time reports were developed. The empirical examination also provided a practical test of the

adequacy of the sentencing framework. The results are important to fill gaps in the theoretical framework used to understand sentencing, to advance understanding of the real levels of consistency or inconsistency in sentencing, and to inform future policy developments. The research adds an additional perspective to the international sentencing research literature. It makes a contribution to a research literature that has otherwise been described as being dominated by studies of the federal and some state jurisdictions in the United States, with a dearth of research from other jurisdictions (Ulmer, 2012). Much of the United States research is from sentencing under guideline systems. The current research will enable the results of research from a non-guideline system to be contrasted with the results from guideline systems.

In Chapter 2 a framework for understanding sentencing disparity is advanced. The framework can be likened to a chain whose links span the gap between sentencing based on a strict application of the law across judges and locations to a patchwork of sentencing approaches differing between judges and locations resulting in inconsistency between locations. In the United States the framework for understanding the occurrence of disparity stems from a consideration of the effects of judicial discretion in sentencing and the difficulties confronting judges in the exercise of discretion. Sentencing frameworks aim to bring consistency to sentencing outcomes but at the same time afford differing degrees of discretion to judges. Discretion is allowed to enable sentences to be fitted to circumstances, to be individualised to particular offending and offender. The exercise of sentencing and reaching decisions is complicated and challenging. In any given case a judge is confronted by a broad and complex set of factors to be taken into account, incomplete information, the need to speculate about the future, and must reach a decision quickly. According to Simon (1976) decision makers confronted with these challenges (complexity, incomplete information, uncertainty and time pressure) are limited by “bounded rationality” and restricted to “satisficing” or making good enough decisions. In response decision makers, in this case judges, develop heuristic strategies or rules of thumb, to simplify and speed up decision making. Sentencing discretion enables concerns outside formal sentencing frameworks to influence outcomes. According to Savelsberg (1992) sentencing is a balance between formal rationalities based on the law or guidelines and substantive rationalities based on concerns, ideologies, attitudes, biases and individual interests. It follows that if sentencing is a balance of rationalities that the content of any sentencing heuristic can be categorised into formal rationalities and substantive rationalities. These concepts explain how sentencing might vary between similar cases and

between judges but it does not explain variation between locations. Court community theory (Eisenstein & Jacobs, 1978; Eisenstein, Flemming & Nardulli, 1988) and the commonality and sociality corollaries from personal construct psychology (Ostrom, Ostrom & Kleiman, 2004) bridge the gap. Court community theory suggests that sentencing is influenced by the concerns and attitudes of the local courthouse community (the judges, prosecutors and defence counsel who regularly appear in the same courtrooms) (Eisenstein & Jacob, 1977) and by the concerns and attitudes of the wider community that the court serves (Eisenstein, Flemming & Nardulli, 1988). The commonality corollary suggests that those who share a common environment are likely to make sense of it in similar ways. Sentencing approaches converge between the judges in a locality because they share a similar set of experiences and observations. Convergence is reinforced by the sociality corollary with those who work together on a regular basis sharing their experiences and approaches and anticipating and adopting the behaviour of others.

The discussion of the framework is followed by a review of empirical research of sentencing consistency between locations (Chapter 3). Research into sentencing disparities between locations is a relatively recent development. Disparities between ethnic groups, genders and, to a lesser extent, other extra-legal factors have been of greater interest. The bulk of the research has been undertaken in the United States with little research elsewhere. Furthermore only United States researchers have utilised multivariate analytical techniques to control for sentencing factors and to isolate the influence that location has on sentencing decisions. The review focuses heavily on the United States but also summarises the findings from elsewhere. Almost all the studies included in the review found evidence of disparities or inconsistency between locations. In a limited number of multi-level models partitioning the influence of sentencing factors between cases and locations, the research found that the weight given to the factors that influence sentencing varies between locations. The research was extended one step further to identify differences between locations in terms of socio-demographic and court caseload variables associated with the disparities in sentencing decisions. The outcome of the research examining these contextual variables has been more equivocal. The size of the locality (judge numbers or population served) and caseload pressure have commonly been found to be associated with disparities between locations. There have been mixed results across the reported research for variables based on the proportion of ethnic minorities in the population, economic disadvantage, the political context, trial rates and available jail capacity.

The balance of the thesis applies the framework and learning from the research to the New Zealand context. In Chapter 4 four predictions and two research questions founded on the framework and prior research are set out. The primary predictions are that, after controlling for offence seriousness, criminal history and offender demographics, similar offenders will be incarcerated at different rates and subject to different terms of incarceration between circuits. Subject to two conditions it was further predicted that the interpretation and use of some offence seriousness, criminal history and offender demographic variables in the determination of incarceration and incarceration sentence lengths will vary between circuits. The conditions were that the predictions of different rates and lengths of incarceration between circuits would be proven, and secondly that the models would demonstrate that offence seriousness and criminal history variables were the principal determinants of incarceration and incarceration sentence lengths. The research questions were specified on the assumption that the predictions would be supported. The research questions examine whether any circuit level contextual variables are associated with the variation in sentence outcomes between circuits.

The balance of Chapter 4 addresses a series of higher level methodological choices affecting the specification of the empirical research. Single level regression models were chosen for the testing of the principal predictions of inconsistency of sentencing decisions between circuits. Multi-level regression models were required to test the predictions that different weights would be given to some case level variables in some circuits and to identify any circuit level contextual variables associated with inconsistency of sentencing decisions between circuits. The selection of variables and their specification for inclusion is discussed. This process was challenging because of the detailed data available and the capacity for some variables to be expressed in a multitude of ways. No adjustment for sample selection bias was made in the sentence length models. This may be seen as a controversial choice. The decision to make no adjustment was based on a concern that an adjustment may distort the findings, and in particular that it might increase the estimated level of inconsistency and result in criticisms that the model specification had “cooked the books”.

Decisions affecting the detailed specification of the models are set out in Chapter 5. The decisions cover: the courts and circuits to be included in the models; the time span; the choice of offences; the choice between individual offence based models or collective models for the selected offences; determination of the unit of analysis; the appropriate treatment of cumulative sentences; restriction of the research to sentencing by resident

judges; the choice between the sentence imposed at first instance or the final sentence after reviews/appeals; and, the exclusion of cases when the most serious charge was for an offence not covered by this research. There is a description of the data sources for the research and the outcome or dependent variables for the models and the control variables.

The results of the modelling are reported in Chapters 6 - 9. Logistic regression models (Chapter 6) show significant and substantial levels of inconsistency between circuits in the use of incarceration. Likewise, linear regression models (Chapter 7) show significant and substantial levels of inconsistency between circuits in the length of incarceration. The extent of inconsistency differed between the three offences modelled; in particular there were greater levels of inconsistency between circuits in the sentencing of offenders for aggravated drink driving and burglary than male assaults female. The pattern of results by circuit also varied between the offences. For example the rates of incarceration for aggravated drink driving were spread across a wide spectrum whereas the rates for burglary form two broad clusters. In addition to demonstrating inconsistency between circuits the logistic and linear regression models confirmed that offence seriousness and criminal history variables were the principal influences on the determination of sentences. The conditions necessary for the predictions that the interpretation and use of some sentencing factors would vary between locations were confirmed. The results of multi-level (hierarchical linear and hierarchical generalised linear) models (Chapter 8) separating the fixed and random effects (by circuit) of sentencing variables demonstrated significant variations in the weights applied to a broad range of seriousness and criminal history factors for aggravated drink driving and burglary. The predictions were also supported for male assaults female but the support was weaker.

In Chapter 9 the multi-level models are expanded to examine the association between circuit context and sentencing inconsistency. Due to the limited number of circuits (17) and multi-collinearity between the contextual variables, bivariate analyses had to be employed. The modelling revealed a consistent difference between provincial and metropolitan circuits; the two categories of circuit were distinguished from one another by many of the other more specific variables that had a significant association with sentencing approaches. The provincial circuits were more severe. However, the small number of circuits and multi-collinearity between the variables precluded more detailed analysis to identify which of the contextual variables distinguishing metropolitan and provincial circuits had the greatest influence.

The final chapter draws together a discussion of the results, the limitations of the research, future research directions, and the implications of the findings for sentencing policy in New Zealand. The strong support for the predictions of inconsistency between circuits and differences in the weights applied to seriousness and criminal history factors support the efficacy of the framework for understanding sentencing. In particular the formally rational sentencing rules dominate the determination of sentences but the application of the factors varies due to the influence of local substantive sentencing concerns. The circuits are distinct communities varying from one another. In particular there is a difference in sentencing approaches between metropolitan and provincial circuits but the differences in conditions between the two categories is difficult to specify. If the level of inconsistency is considered to be a concern then any solution must rebalance the influence of formal and substantive rationalities. Judges need more guidance on the sentences to be applied for a given range of offending – the determination of the approach should not be dependent on the judge or the judges in a particular circuit. The current sentencing framework is manifestly inadequate. It is doubtful that this could be directly rectified by providing more detailed generic guidance in legislation. More detailed offence based sentencing guidance is required; in the current context there are two options that could be used. The Court of Appeal could issue a broader range of guideline judgments or the legislation for the Sentencing Council and the process for developing and promulgating guidelines could be implemented. For logistical and public policy reasons the Sentencing Council approach is preferred. There is a risk that failure to address the levels of inconsistency will result in the sentencing system falling into disrepute.

The findings should not be interpreted as if they are a critique of the sentence imposed in any individual case or of the sentencing by any judge or in any circuit. It is a critique of sentencing guidance in New Zealand and its ability to achieve a fundamental tenet of justice: the similar treatment of similar offenders being sentenced in similar circumstances.

Chapter 2: A Framework for Understanding Sentencing Disparity

The occurrence of sentencing disparities between locations has been the subject of a substantial body of research in the United States in recent years. The research has found widespread evidence of disparity (see Johnson, 2011; Ulmer, 2012 for reviews). Before examining whether the Law Commission's (2006) research indicating variation in sentencing for like offenders in different locations in New Zealand withstands rigorous scrutiny the more immediate question is why inconsistency between locations occurs. Those who believe that it will not withstand scrutiny base their argument on the proposition that the legal framework guiding the sentencing process in New Zealand is sufficient to achieve consistency. The supporters of the current regime rejected the empirical argument suggesting that it was misconceived by failing to properly recognise differences between cases and compositional differences in caseload between locations. Conversely the Law Commission argued that the level of discretion accorded to judges almost makes inconsistency inevitable and pointed to the differences in rates of imprisonment and sentence lengths as proof.

The arguments for and against the existence of a problem and the need for change both lack depth. They are little more than assertions of a view, they do not provide a framework explaining how the sentencing of the myriad cases results in consistent sentencing between locations or how it results in disparity. A framework for understanding sentencing disparity is developed in this chapter. The framework can be likened to a chain whose links span the gap between sentencing based on a strict application of the law across judges and locations to a patchwork of sentencing approaches differing between judges and locations resulting in inconsistency between locations.

In the United States the framework for understanding the occurrence of disparity is based on sentencing discretion and the difficulties confronting judges in the exercise of discretion. In simple terms judges have discretion because sentencing is challenging. Judges must evaluate the widely varying circumstances of offending and offenders against a complex set of factors, and in doing so they are often faced with incomplete information, take into account uncertain future prospects and must arrive at a decision quickly. Each case has to be considered separately, a sentence crafted to the circumstances of the case is needed to ensure justice is done. The exercise of discretion and the consequential challenges form the first two links in the chain.

In decision making contexts characterised by complexity, uncertainty and time pressure, Simon (1976) suggested decision makers are limited by “bounded rationality” and restricted to “satisficing” or making good enough decisions. To manage the challenges decision makers develop heuristic strategies or rules of thumb to simplify and speed up decision making. Bounded rationality and the use of heuristic strategies are the third and fourth links in the chain.

Based on the availability of sentencing discretion Savelsberg (1992) argued that sentencing can be understood as a balance between formal and substantive rationalities. Formal rationality involves the strict application of externally developed rules to the determination of sentences. Examples of externally developed rules include legislation, sentencing guidelines and appellate authority. Substantive rationality is the influence of concerns, ideologies, attitudes, biases and individual interests on sentencing. Substantive rationalities are not expressed as formal rules. If sentencing is a balance of rationalities it follows that the content of any sentencing heuristic can be categorised into formal rationalities and substantive rationalities. Savelsberg’s (1992) construct is the fifth link in the chain.

To this point the chain explains how sentencing might vary between similar cases and between judges but it does not explain variation between locations. Substantive rationalities, like formal rationality, may be system wide or they may reflect concerns and attitudes based on local conditions and result in disparities between locations in the same legal system. Earlier researchers made a similar argument suggesting that sentencing was influenced by the concerns and attitudes of the local courthouse community (the judges, prosecutors and defence counsel who regularly appeared in the same courtrooms) (Eisenstein & Jacob, 1977) and by the concerns and attitudes of the wider community the court served (Eisenstein et al., 1988). Court community theory is the sixth link in the chain establishing a base for variation between locations in the recognition of the influence of the immediate court community and the wider community around the court.

Using constructs borrowed from personal construct psychology Ostrom et al. (2004) explained how judges’ personal sentencing approaches converge into a local sentencing approach. In simple terms they argued that under the “commonality corollary” those who share a common environment are likely to make sense of it in similar ways. Secondly, they argued that this is reinforced by the “sociality corollary” with those who work together on a regular basis sharing their experiences and approaches and anticipating

and adopting the behaviour of others. The commonality and sociality corollaries form the seventh link in the chain.

In the United States much of the research into sentencing disparity has been influenced by Steffensmeier, Ulmer and Kramer's (1998) focal concerns models founded on the exercise of discretion, the challenges of sentencing, and the use of heuristic devices drawing on the influence of substantive rationalities. Although the focal concerns model is a general theory potentially applicable in New Zealand and elsewhere, its specific applicability to New Zealand is questionable in light of the specific nature of the statutory sentencing regime.

In the remainder of this chapter the seven links in the chain spanning the gap from sentencing discretion to inconsistency between locations will be elaborated. The chapter concludes with a discussion of the applicability of the theory to sentencing in New Zealand.

Exercising Discretion: The Sentencing Challenge

Ostrom et al. (2004) identified sentencing as being a complex decision making task despite it being a routine activity for judges. Judges are often called on to frequently sentence offenders for the same high volume offence type (e.g. burglary or lower level assaults). However, sentencing remains complex despite the frequency of the task because each case must be dealt with on its merits. Describing judges as frequently sentencing burglars or offenders who committed minor assaults masks the reality that the offending and offenders within these broad categories differ widely. The merits of a particular case can be challenging to determine. Judges must evaluate the varying circumstances of offending and offenders against a complex set of factors, and in doing so they are often faced with incomplete information, have to take into account uncertain future prospects and arrive at a decision quickly. Separate consideration of each case resulting in a sentence crafted to the circumstances of the case (individualised sentencing) has been held to be necessary to ensure justice is done.

Outwardly similar cases can and sometimes should properly result in substantially different outcomes. The transformation of the evaluation of the circumstances into a sentence may require a finely balanced choice between sentence types and almost always requires the determination of the quantum of sentence from a continuum or between two points on the continuum.

The ordinary complexity of weighing and balancing a multiplicity of factors is compounded by having to do so with incomplete information and having to overcome the uncertainty associated with the application of some factors. Some factors pose little difficulty, for example, whether an offender pleaded guilty, was a disqualified driver, had prior convictions for the offence or offended while on bail can be established with certainty. The sincerity of a claim of remorse or the balance between competing versions of the ferocity of an admitted assault cannot be so readily determined; the absence of complete information on the assault requires the exercise of a judgement. Some factors are more challenging because they require a judgement about the future: for example, whether an offender is likely to re-offend; or, conversely, whether an offender can be rehabilitated. The challenge does not end with a determination of the factual basis for sentencing; it must be weighed up against any sentencing guidance and transformed into a sentence. The challenge of doing all this is exacerbated by high workloads and time pressure on decisions. Quicker decisions help to ensure that the consequences are felt by the offender as close to the offending as practicable and assist victims to get early closure. What does this combination of discretion and sentencing challenges mean for sentencing decision making?

Bounded rationality

Many researchers who have sought to understand sentencing decision making have focussed on the sentencing challenges (complexity, incomplete information, uncertainty, and time limitations) as critical characteristics affecting the decision process (Albonetti, 1991; Hupfeld-Heinemann & von Helversen, 2009; Ostrom et al., 2004; Steffensmeier, Ulmer & Kramer, 1998; von Helversen & Rieskamp, 2009). These challenges are central to Simon's (1976) theory of bounded rationality. Bounded rationality has been a key element in theoretical frameworks explaining sentencing decision making.

Bounded rationality was labelled in the 1970s (Simon, 1976) but it was first described in Simon's original work published in 1945. Simon criticised the classical notion of unbounded rationality and the assumption that decision makers sought to assess all options and seek information until they reached the point of being able to recognise and choose the optimal decision. Simon (1976, pp. xxvi-xxvii) argued that: "economists attribute to economic man a preposterously omniscient rationality. Economic man has a complete and consistent system of preferences that allows him always to choose among the alternatives open to him; he is always completely aware of what these alternatives are;

there are no limits on the complexity of the computations he can perform in order to determine which alternatives are best...”. According to Simon (1976, p xxvii), economic man had “little discernible relation to the actual or possible behaviour of flesh-and-blood human beings”.

Simon (1976) recognised that the ability of humans to reach optimal decisions when faced with complex decisions having to be made under time constraints is limited. Due to limited computational powers humans cannot always choose the best alternative (the one that has greatest utility). Nor do they always have complete information and, due to time constraints, it is often impractical to seek to rectify all information deficits.

Within the limitations imposed by bounded rationality humans must make do with imperfect assessment of alternatives and incomplete information and therefore accept outcomes that are “good enough” or “satisficing” (a decision that is hoped to be close to the best possible answer) but which may not be optimal. According to Simon (1992), problem solving tends to involve a selective search through large numbers of alternatives. The search is selective because an undirected search would require an enormous amount of time. Heuristics provide a means of being selective.

Any method to make decision making or problem solving quicker and more manageable is a heuristic. Heuristics have a common set of features. They simplify the problem (decision), reduce the time required to reach a decision, and do not require a judgement that the decision is optimal, just that it is good enough (Gigerenzer, 2006). Heuristics vary on a range of dimensions with no universal heuristic (Gigerenzer, 2006). When it is not possible to optimise, decisions are made by a reliance on “routines, habits, intuition and rules of thumb” to quicken the decision making process and find a good enough solution (Gigerenzer, 2006. p. 31). Gigerenzer and Todd (1999) list three broad building blocks that can be found in any heuristic. They are: rules guiding the search for cues (factors or information) relevant to the decision/problem at hand, rules for stopping the search for cues, and rules for decision making based on the cues.

Heuristics may be compensatory or non-compensatory. In a non-compensatory heuristic the search for cues ends once the decision maker finds the first cue that discriminates between the alternative decisions. No combination of other cues can change the decision. A compensatory heuristic requires the weighing of a series of cues to achieve a balanced decision, and the search only stops when the decision maker is confident that no combination of other cues could change the indicated decision.

Heuristics can be general, or they can be specific to a particular domain. The recognition heuristic is an example of a general heuristic. It can be employed to choose between options in any domain (Gigerenzer & Goldstein, 1999). The recognition heuristic relies on the decision maker's ability to recognise one of the solutions to a problem but not any other. Gigerenzer and Goldstein (1999) gave the example of someone faced with choosing which of two cities is the larger. They hypothesise that if the person has only heard of one of the two cities they would choose that city. If they do not recognise either city or if they recognise both cities they need to find a different discriminator. A heuristic used in the health system that relies on blood pressure, age and sinus tachycardia to determine if a heart attack patient is low or high risk is an example of a domain specific heuristic (Gigerenzer & Todd, 1999) because it relies on cues particular to the heart health care domain.

Depending on the search, stopping, and decision making rules, heuristics may be faster or slower in terms of the time taken and more or less frugal in terms of the fraction of information used. Some heuristics are known as fast and frugal heuristics (Gigerenzer & Todd, 1999) because they are very fast and use only a small fraction of the available information. The fastest and most frugal heuristics require only one reason for a decision to be made (Gigerenzer & Goldstein, 1999). The "take the best" heuristic is an example of a fast and frugal heuristic. Cues are considered in descending order according to the frequency with which they have enabled a decision to be taken. A decision is made as soon as the first factor that discriminates between the options is found. Only one factor is needed for a decision. It is a non-compensatory heuristic as it cannot be over-ridden by any other factor or combination of factors.

Other heuristic strategies are more complex, requiring more information and more computation. Simon (1992) described two more complex heuristic approaches. First, the planning heuristic: the decision maker omits or puts to one side some of the detail of the problem and focuses on its essential features. The simplified problem is solved to find a favoured solution which is used as a guide to solve the full problem and adjusted as necessary. Simon (1992, p. 68) claimed use of the planning heuristic "can increase the speed of the solution by many orders of magnitude". Factoring was the second heuristic approach described by Simon: the decision maker breaks the problem into sub-problems and addresses the sub-problems with the smallest number of unknowns first.

In sum, heuristics are any approaches employed to overcome the challenges of complexity, information deficits and uncertainty, and time pressure on decision making.

There are many types of heuristic and they vary in complexity and the speed of decision making and the amount of information required.

Sentencing heuristics

The axiom that sentencing is an art rather than a science suggests that sentencing is a search for “good enough” or “satisficing” decisions (decisions hoped to be close to the best possible) and that it is not possible to optimise. The discretion entrusted to sentencing judges, the limited guidance provided to assist the exercise of this discretion, the conditions under which sentencing commonly occurs (complexity, uncertainty and time pressure) combine to suggest that the sentence decision making domain is one that is likely to be subject to bounded rationality and the use of heuristics. The use of heuristics in sentencing (or other legal decision making) has been assumed to be reasonable because of the nature of the decision making environment (Gigerenzer, 2006; von Helversen & Rieskamp, 2009).

Some general propositions can be made about any sentencing heuristics. First, it is assumed that they will be domain specific. The statutory framework, the common law and the specific qualifications and background of judges point toward domain specific heuristics. Sentencing decisions are explained in terms of factors that are particular to the criminal justice and sentencing domains, and are often specific to the particular offence being sentenced. The need to justify sentences in terms of domain specific factors is reinforced by the appellate process.

Second, it is likely that any heuristics will be compensatory rather than non-compensatory. The expectation in New Zealand that the sentence represents the net consequence of considering the effects of aggravating and mitigating factors suggests a compensatory approach. As a result faster and more frugal heuristics are unlikely to be prominent in judicial decision making. The multiplicity of factors expected to be taken into account in any sentencing decision suggests that there will be few circumstances in which a one reason decision could be made. The differing roles of the prosecution and defence counsel and the opportunities they have to influence the sentencing decision, the requirement on the judge to call for and take account of various pre-sentence reports and the disciplining effect of appeals can be viewed as mechanisms likely to encourage the balancing of multiple factors. Within the overall decision process there may be some decisions (problems) that are solved more quickly and frugally than others. For higher volume, less serious offences, the sentencing alternatives tend to be more limited. For

example, in New Zealand most offenders sentenced for a first drink drive offence are fined. In the three years from 2004 to 2006 89.2% of 24,993 offenders sentenced for a first drink drive offence with no accompanying charges had a fine as their most serious sentence (this information was derived from a dataset sourced from data supplied to the Sentencing Establishment Unit and obtained from the Law Commission under the Official Information Act). Even though a fine was imposed in a high proportion of cases the judge would still check for cues indicating that a fine is impracticable or otherwise inappropriate. The particular attributes of any heuristics employed are likely to vary by offence type, with some employing fewer cues and being faster and more frugal than others.

In addition to the system settings that militate against fast and frugal heuristics in sentencing, the principle of individualised justice, the determination of the sentence having regard to the particulars of the offending and the offender, and the need for the judge to have sufficient comfort with the “good enough” sentence also militate against faster and more frugal heuristics. The weight of the responsibility entrusted to the judge to do justice means that judges cannot too readily sacrifice accuracy (effectiveness) in decision making for greater speed (efficiency). The risk or threat of an appeal may also discipline judges to not be too fast or frugal in their approach.

A busy judge considering whether or not an offender should be imprisoned might be expected to make a decision once he or she has identified a factor or a combination of factors that he or she is convinced resolve whether imprisonment is or is not appropriate. The search for additional information ends once the judge is “satisfied” (comfortable that the choice is good enough). Sometimes this might occur relatively quickly and easily. For example, a judge may determine that imprisonment appears to be the only reasonable sentence for a recidivist drink driver, a scan of other cues may not reveal anything that warrants re-consideration of that initial judgement, and the judge moves on to the task of determining the term of imprisonment. On other occasions it may be more difficult. For example, the number of prior convictions may place the offender on the cusp of imprisonment. The judge then seeks more information on factors such as the offender’s driving behaviour, the recency of prior convictions, previous compliance with community sentences, other indicators of the risk of re-offending and the potential for rehabilitation, before weighing up the pros and cons and coming to a decision.

The development of heuristic strategies to overcome the challenges of sentencing is a key link in the chain. The focal concerns theory developed by Steffensmeier et al. (1998), and discussed in more detail later, is a broad description of a heuristic strategy.

However, it is difficult to define the content of specific heuristics used in sentencing. The content of any heuristic strategy is likely to vary from offence to offence and depends on views about which factors are important and the seriousness of the offence. It is sufficient for the purposes of this study to recognise the applicability of heuristic strategies as a means of responding to the challenges of sentencing.

Formal and Substantive Rationalities and Court Communities

In an influential article Savelsberg (1992) described sentencing as a balance between formal and substantive rationalities. At the time Savelsberg was discussing why the uniformity of sentencing sought by the development and implementation of sentencing guidelines may fail to be achieved to the extent expected by the architects of the guidelines. The concepts have wide application and have been used widely in subsequent research (e.g., Dixon, 1995; Johnson, 2005).

Savelsberg (1992) distinguished formal rationality, the making of decisions according to a set of rules, from substantive rationality, the influence on those decisions of circumstances and considerations not contained in the rules but perceived to be relevant to sentencing. The concepts were developed in light of the widespread movement toward sentencing guidelines in the United States. Savelsberg (1992) described the guidelines as an attempt to impose greater uniformity on sentencing by prescribing rules intended to be applied universally.

Sentencing guidelines with defined ranges for sentence based on explicit legal criteria are a good example of formal rationality (Johnson, 2005). Formal rationality has been characterised as resulting in sentences based on legal rules and criteria applied equally in all cases (Dixon, 1995). Formal rationalities are externally imposed and apply across a jurisdiction. Conversely, substantive rationality is said to be guided by non-legal factors applied on an individual basis to specific cases (Johnson, 2005). The capacity built into the guidelines for judges to depart from the guidelines and impose a sentence different from that indicated by strict application of the guidelines was identified by Savelsberg (1992) as an opportunity for substantively rational sentencing concerns to modify the formally rational sentence. The critical feature of the framework is that the balance is dependent on the extent of judicial discretion. As such, the proposition that sentencing can be understood as a balance between formal and substantive rationality is applicable to virtually any criminal justice system.

Formal rationality involves the strict application of externally developed rules to the determination of sentences. Examples of externally developed rules include legislation, sentencing guidelines and appellate authority. Substantive rationality is the influence of concerns, ideologies, attitudes, biases and individual interests on sentencing. Substantive rationalities are not expressed as formal rules and are not externally imposed. If sentencing is a balance of rationalities it follows that the content of any sentencing heuristic can be categorised into formal rationalities and substantive rationalities.

The concepts of legal and extra-legal factors frequently used in the sentencing research literature correspond to formal and substantive rationality. Legal factors are mandated by the formally rational rules; they are factors that are intended to influence sentencing. Conversely, extra-legal factors are not intended to influence sentencing; they do not spring from the formally rational rules and are the manifestation of substantive rationality. Ethnicity and gender are the most common examples of extra-legal factors that may influence sentencing. Moreover, the influence of substantively rational sentencing concerns is not limited to taking account of extra-legal factors. The influence may be expressed via the interpretation and application of formally rational rules.

Court community theory

The concepts of formal and substantive rationality complement two competing metaphors for understanding courts and sentencing described by Eisenstein et al. (1988). According to the traditional legal metaphor, courts and outcomes (like sentencing) are simply understood as the result of the application of the law (formal rationality). The law is described “as a set of rules transcending time, geography and the circumstances surrounding specific cases” (Eisenstein et al., p.5) expected to afford all individuals equality before the law. The competing community metaphor, preferred by Eisenstein et al. (1988), identifies courts as communities and locates them as political bodies. They are political because the discretion afforded to judges affords them the opportunity to interpret the law and to make policy within the broad boundaries established by the law. In doing so they are influenced by local conditions (substantive rationalities).

Eisenstein et al. (1988) adopted two usages of community from the social sciences in the development of the community metaphor. The two usages were, first, the notion of communities as neighbourhoods and, second, the notion that “undercover” communities form within organisations and institutions. The two usages of community correspond to Eisenstein & Jacob’s (1977) earlier conception of local context as comprising the influence

of two distinct, sometimes complementary and sometimes competing, environments: the influence of the courtroom workgroup (an undercover community); and the wider influence of the community the court is located within (Eisenstein & Jacob, 1997; Eisenstein et al., 1988).

Eisenstein and Jacob (1977) conceived of the courtroom workgroup as having at its core the judge and the prosecutor(s) and defence counsel who appear in the judge's courtroom. Courtroom workgroups were theorised as having four goals: doing justice; disposing of the caseload; maintaining group cohesion; and reducing uncertainty. Differences in the characteristics of workgroups between locations could lead to variation in outcomes (including sentencing) between locations (Eisenstein & Jacob, 1977). Workgroup members' familiarity with one another, dependence on one another, stability, and the relative importance attached to the four goals and task specialisation could all vary between locations and influence outcomes. Individual roles may vary from location to location depending on the control exercised by sponsoring organisations. For instance the role of prosecutors may be affected by the way they are assigned to courts, by the control extended over them by their organisation and the existence of policies affecting their discretion. In instances where prosecutors, such as the state attorney, were elected or appointed their behaviour or control over their prosecutors may be influenced by political parties or other organisations instrumental in their election or appointment. Outcomes can also be influenced by pressures from groups within the wider environment including the police, appellate courts, media and political organisations. Variability in the weights placed on the goals and the impact of the characteristics theorised to influence outcomes could result in variation in sentencing between locations.

As substantive rationalities are not externally imposed they can vary from judge to judge and place to place. Substantive rationalities based on local conditions, concerns, ideologies, and biases and individual interests can subvert the formal rationality of guidelines (Kramer & Ulmer, 2002). This is not to say that they must necessarily be unique to a location as they may be the same or similar to the extent that locations share common characteristics and concerns. In some instances particular conceptions of substantive rationality may apply across most localities. For example, concerns or biases based on stereotypical views of personal characteristics, such as gender or ethnicity, that are not incorporated into or founded on the formally rational rules may be pervasive across localities.

Ostrom et al. (2004) identified why judges personal sentencing constructs formed within the bounds of discretion and influenced by local conditions and substantive rationalities could be expected to converge into local sentencing constructs. Using two concepts from personal construct psychology, the commonality corollary and the sociality corollary, Ostrom et al. (1998) explained the patterned responses exhibited by judges within a locality. They postulated that individual judges develop personal constructs as the means by which they make sense of their world. To deal with complexity and uncertainty the judges develop personal theories based on their observations, and test them against actual experience and replace or refine the construct in light of that experience. Under the commonality corollary “people who share a similar work environment and similar cultural norms also likely share similarities in how they “make sense” of their experience” (Ostrom et al., 2004 p.45) promoting convergence of approaches locally. Although the construct is personal to the individual, in this case the judge, the common features of the environment and the common elements of the backgrounds of the judges tend to cause the constructs of local judges to converge. The convergence of sentencing approaches will be enhanced to the extent that all judges have some common background from their training and to the extent that their life experiences are more or less similar. However, if their life experiences are different, or if they live in different communities and have had different types of legal practice, their personal constructs can diverge. The sociality corollary strengthens the influence of the locality on personal sentencing constructs. As summarised by Ostrom et al (2004, p.46) “individuals who work together on a regular basis will naturally begin to anticipate the behaviour of others. In doing so, individuals develop and adopt practices based on group expectations and established ways of understanding the world. Unwritten rules and practices within the group tend to be reinforced by the give and take of everyday observation and interaction. Therefore, we should not be surprised that as judges anticipate, and sometimes emulate, the conventional sentencing practices of their colleagues, patterned responses in sentencing outcomes emerge.”

Focal concerns theory

In the United States focal concerns theory (Steffensmeier et al., 1998) has been used by many researchers as the foundation for examining sentencing disparities. Ulmer (2012) described focal concerns as an elaboration of Savelsberg’s (1992) theory. Based on interviews with judges, prosecutors and defence counsel in Pennsylvania Steffensmeier et al. (1998) argued that judges determine sentences based on judgements about three “focal

concerns”: blameworthiness; protection of the community (dangerousness/rehabilitative potential); and, the practical implications of sentencing decisions. Sentences are the outcome of an assessment of offence seriousness and the offender’s criminal history against these three concerns.

Blameworthiness takes account of the offender’s culpability and the harm done by the offending. It is associated with retributive sentencing or just deserts, and focuses on the severity of the offending ensuring that the punishment fits the crime. The judgement about severity may be modified by offender factors. Relevant prior offending may increase blameworthiness, whereas personal circumstances (physical and mental health, an abusive upbringing) may decrease it.

Community protection draws on similar factors but is conceptually distinct because it focuses on the need to protect the community by either incapacitating the offender or deterring would-be offenders. To achieve community protection the sentence based on blameworthiness may need to be varied. The risk that an offender will re-offend is inherently uncertain. As a result, Judges rely on attributions based on the nature of the offence (e.g., violent or property), case information, relevant prior offending, particular facts of the crime (e.g., the use of a weapon), and offender characteristics such as drug dependency, education, employment, or family history to identify those expected to re-offend (Steffensmeier et al., 1998).

Practical implications cover a range of distinct potential organisational and individual impacts arising from a possible sentence. The organisational impacts include “maintaining working relationships among courtroom actors, ensuring the stable flow of cases, and being sensitive to local and state correctional crowding and resources” (Steffensmeier et al., 1998 p. 767). Individual impacts include the ability of the offender to serve the sentence and the impact of sentence on an offender’s family (Steffensmeier et al., 1998).

The focal concerns provide a structure for organising the considerations relevant to determining sentence. It does not eliminate the practical challenges of sentencing but it does organise the decision process and to that extent reduces complexity. When the practical challenges of sentencing leave judges uncertain about the sentence based on an assessment of legal or formally rational factors they are said to resort to “perceptual shorthand” (substantive rationality) based on the gender, race, and age of offenders.

The idea of perceptual shorthand followed Albonetti’s (1991) bounded rationality/uncertainty avoidance perspective. Albonetti assumed that uncertainty about

the likelihood of re-offending caused judges the greatest difficulty in determining the sentence. Albonetti (1991) argued that in order to overcome the effects of bounded rationality on the judging of the likelihood of re-offending judges rely on patterned responses to reach a decision (sentence). The patterned responses relied on an attribution process influenced by causal judgements of the propensity of different types of offender to re-offend across a range of factors. The range of factors was a mix of formal and substantive rationalities. Factors such as prior record (expected to increase sentence severity because it indicates a disposition to re-offend) or the use of a weapon (expected to increase sentence severity because it indicates a disposition toward more serious offending and a higher likelihood of re-offending) can be categorised as formally rational. If the judge remained uncertain of how to proceed based on formally rational factors Albonetti (1991) suggested they rely on attributions based on substantive rationalities focussed on race and sex. Being black would increase sentence severity because of stereotypical views that blacks are more likely to re-offend than whites. Conversely, being female would reduce sentence severity because females are viewed as caregivers and less likely to re-offend than males.

Based on previous research Steffensmeier et al. (1998) predicted, like Albonetti (1991), that racial and gender stereotypes influence sentencing judges, with a perceived greater pre-disposition on the part of blacks to offend channelled through the focal concerns to decisions to imprison and to imprison for longer. They also predicted that sentencing would be influenced by the age and gender of the offender in addition to race, such that sentencing would be more severe for blacks, males, and younger offenders. Furthermore they predicted that interactions between demographic factors would result in more severe sentencing for some sub-groups, and that in particular young, black males would be sentenced more severely than any other race-age-gender grouping because they were perceived to be the most dangerous or crime-prone group (requiring the community to be protected) and less likely to be harmed by incarceration (suffering no practical consequence from the sentence). Irrespective of the influence of perceptual shorthand the authors also expected that offence severity and prior record would be strong predictors of sentence severity net of other factors.

Focal concerns theory describes a broad heuristic simplifying the determination of sentences. Sentencing is based on a three part structure for organising the consideration of sentencing factors and a means (perceptual shorthand) for resolving difficult cases. It is a

heuristic specific to the sentencing domain, using the language of sentencing to describe how satisficing or good enough decisions are reached.

Researchers (see Crow & Gertz, 2008; Johnson, 2005; Kramer & Ulmer, 2002; Ulmer, Eisenstein & Johnson, 2009) who have employed the focal concerns framework to study sentencing disparity have found disparities between locations in addition to gender and ethnic disparities. Kramer and Ulmer (2002) suggest that the focal concerns are generally applicable across regimes they are not given the same meaning or emphasis in all locations. “The meaning, relative emphasis and priority, and situational interpretations” afforded to the focal concerns can vary between locations because they are “embedded in local court community culture, organisational contexts, and politics” (Kramer and Ulmer, 2002, p.903). Seemingly local substantive rationalities within the court community influence sentences within the focal concerns framework.

Summary of the Framework

Sentencing disparity and the formation of distinct local sentencing approaches can be understood as the outcome of sequence of effects explaining how sentencing based on a common legal framework applying across judges and locations can be transformed into a patchwork of sentencing approaches differing between judges and locations.

Sentencing frameworks aim to bring consistency to sentencing outcomes but at the same time afford differing degrees of discretion to judges. Discretion is allowed because of the challenges of sentence decision making to enable sentences to be fitted to the circumstances of individual cases. When decision makers are confronted with the challenges (complexity, incomplete information, uncertainty and time pressure) facing judges they are limited by “bounded rationality” and restricted to “satisficing” or making good enough decisions (Simon, 1976). In response, decision makers, in this case judges, develop heuristic strategies or rules of thumb to simplify and speed up decision making. Sentencing discretion enables concerns outside formal sentencing frameworks to influence outcomes; sentencing is a balance between formal rationalities based on the law or guidelines and substantive rationalities based on concerns, ideologies, attitudes, biases and individual interests (Savelsberg, 1992). It follows that if sentencing is a balance of rationalities then the content of any sentencing heuristic can be categorised into formal rationalities and substantive rationalities. These concepts explain how sentencing might vary between similar cases and between judges but do not directly explain variation between locations. Court community theory (Eisenstein & Jacobs, 1978; Eisenstein et al.,

1988) and the commonality and sociality corollaries from personal construct psychology (Ostrom et al., 2004) bridge the gap. Court community theory suggests that sentencing is influenced by the concerns and attitudes of the local courthouse community (the judges, prosecutors and defence counsel who regularly appeared in the same courtrooms) (Eisenstein & Jacob, 1977) and by the concerns and attitudes of the wider community the court serves (Eisenstein et al., 1988). The commonality corollary suggests that those who share a common environment are likely to make sense of it in similar ways. Sentencing approaches converge between the judges in a locality because they share a similar set of experiences and observations. Convergence is reinforced by the “sociality corollary” with those who work together on a regular basis sharing their experiences and approaches and anticipating and adopting the behaviour of others.

Applying the Framework to New Zealand

Is the framework for understanding sentencing disparity applicable to sentencing in New Zealand? The applicability of the framework is conditional on three factors. Do judges have discretion in the determination of sentences? Do the challenges confronting judges in determining sentences apply to sentencing in New Zealand? Are the criminal justice and court systems configured in a way that is consistent with the opportunity for local sentencing constructs to form?

Sentencing discretion in New Zealand

It is widely recognised that judges have very broad sentencing discretion. In his opposition to increased guidance and denial of inconsistency Priestley (2009) recognised the breadth of discretion. The judge explained that it is the role of judges to do justice in individual cases by sentencing offenders within the broad framework established by Parliament. Parliament proscribed specific behaviours as crimes and set maximum penalties for those crimes, and guided the sentencing process by laying down broad purposes and principles in the Sentencing Act 2002. The courts were described as holding the balance between the Crown and the citizen. The judge argued that any curbing of sentencing discretion beyond these limits would have reduced the ability (and discretion) of judges to do justice.

The Law Commission (2006) relied on broad discretion to explain sentencing disparity. The Commission pointed out that the statutory regime left it to the judge(s) to determine where the threshold lies in terms of sentencing an offender to imprisonment or imposing a less severe sentence, and that for the most part the term of imprisonment was

only limited by the maximum penalty. While acknowledging that the extent of discretion has been limited by the development of a body of appellate decisions relied on as precedent for sentencing decisions the Commission saw severe limitations in the guidance. Counsel were said to be able to find competing precedents for the same offence and counsel and judges can often readily distinguish the immediate case from any claimed precedents. While the introduction of guideline judgments providing guidance that is not fact-specific and able to be generally applied to the run of cases addressed in the judgment has strengthened appellate authority the impact has been limited because the guidance only applies to a restricted number of the most serious offences.

The Sentencing Act 2002, which largely codified case law (Robertson, 2007; Hall, 2008), primarily places a framework of purposes and principles around sentencing. It supports the process by informing judges about what to take into account at a generic level but provides no substantive guidance on custodial thresholds or sentence lengths beyond a requirement to impose the least restrictive sentence appropriate. The Law Commission (2006, p.18) described the legislation as “highly permissive” and “characterised by substantial judicial discretion as to the way in which the purposes and principles of sentencing should be translated into sentencing levels”. The purposes range across accountability, reparation, deterrence, community protection, and rehabilitation. The selection is made case by case, the purposes can be used as a mandate for more punitive approaches or for therapeutic justice approaches depending on the preferences of judges. The choice is guided by the principles of sentencing. This is significant because the Courts have long interpreted the principles as a framework for proportionate sentencing (Robertson, 2007). The requirements that the judge take account of: the gravity of the offending, including the offender’s culpability; the seriousness of the type of offence in comparison with other types of offences, as indicated by the maximum penalties; and the general desirability of consistency has been interpreted as a recipe for proportionate sentencing (Robertson, 2007).

Strict proportionality does not always prevail. Robertson (2007) lists four exceptions defined in case law allowing departure from proportionality: 1) an offender who poses a particularly high risk of reoffending may receive a longer sentence for the protection of the community; 2) a reduced sentence may be imposed if justified by the offender’s potential for rehabilitation; 3) a sentence that would otherwise be proportionate may be reduced for merciful reasons; and 4) in exceptional cases, where a sentence will have a disproportionate impact and cause exceptional hardship to either the offender or the

offender's family as a result of ill health or other personal circumstances, a reduced sentence may be justified.

The breadth of judicial sentencing discretion in New Zealand is arguably greater than the United States. The development since the early 1970s of sentencing guidelines by state governments and the Federal government (Johnson, 2011) has set thresholds for the use of imprisonment and sentence ranges that are tied to specific offence seriousness and criminal history factors. The guidelines limit discretion in ways that are not mirrored in New Zealand.

Sentencing challenges

Although it may be tautological, or circular reasoning, it can be inferred from the wide discretion afforded to judges in New Zealand that the judges face the challenges of discretionary sentencing. In simple terms judges have discretion because sentencing is challenging. Judges must evaluate the widely varying circumstances of offending and offenders against a complex set of factors, and in doing so they are often faced with incomplete information, take into uncertain future prospects, and need to arrive at a decision quickly.

The judge must determine the purpose(s) of sentencing relevant to the case and choose the sentence type(s) and quantum of the sentence. These decisions have to be taken after evaluating the seriousness of the offending, the culpability of the offender, the significance of criminal history on sentence, the adjustments necessary to take account of aggravating and mitigating factors personal to the offender and the reduction to be made for a guilty plea. Depending on the offence there can be a substantial number of factors required to be taken into account in the evaluation of each component of the sentencing equation. The models reported in later chapters and the lists of potentially relevant variables set out in Appendix E demonstrate the wide range of factors that may be pertinent in any specific case.

The configuration of the criminal justice and court systems

New Zealand has a unitary system of government. The criminal justice statutes (including the Sentencing Act) and appellate authority apply to everyone equally and do not depend on location; there are no states or provinces that may enact different laws in different places.

Adult sentencing occurs in the District Courts and the High Court but it is monopolised by the District Courts. Described as the workhorses of the courts by Joseph (2007), the District Courts undertake 99% of adult sentencing. The 63 District Courts are organised into 18 circuits. The circuits are relatively autonomous. Each circuit, apart from one, has a cadre of judges resident in the circuit. In almost all instances the resident judges have their chambers in the largest court in the circuit. The resident judges in each circuit undertake most of the sentencing (>80%) in the circuit. However, the effect of the structural separation of the courts by circuit is reduced by the use of visiting judges who undertake the balance of sentencing.

The separation of the District Courts into circuits creates an opportunity for courtroom workgroups (Eisenstein & Jacob, 1977) and the wider circuit community (Eisenstein et al., 1988) to influence local sentencing norms. The judges in each circuit share a common environment, they sentence in the same set of courts and they are co-located. There is an opportunity for commonality of experience and socialisation of that experience between the judges for personal sentencing constructs to converge and result in distinct sentencing patterns between locations.

The applicability of focal concerns theory and other heuristic strategies

Despite widespread acceptance and utilisation of focal concerns theory in the United States the theory is not readily applicable in the New Zealand context because of the role of the Sentencing Act. The purposes and principles in the Act guide the process. However, the difference is more semantic than substantive. The purposes and principles in the Sentencing Act and the focal concerns can be reconciled. Blameworthiness corresponds to three purposes: holding the offender to account; instilling a sense of responsibility in the offender; and denunciation of the offending. Protection of the community corresponds to three purposes: deterrence; protection of the community; and providing for the rehabilitation and reintegration of the offender. The third focal concern, practical constraints, is not matched by any of the purposes. It is reflected in the sentencing principles requiring the personal circumstances of the offender to be taken into account and, in the exceptions to strict proportionality, allowing for merciful sentencing and reduction of a sentence that would have a disproportionate impact on the offender or their family. Two purposes from the Sentencing Act do not readily reconcile to the focal concerns, they are providing for the interests of the victim, and providing reparation. These purposes do not set up a substantive difference between the Sentencing Act and

focal concerns. The Sentencing Act, via the purposes and principles, is a heuristic device; it provides a constant framework, albeit not one that is likely to greatly increase the speed of decision making.

The specification of a preferred sentencing methodology for the determination of sentence length by the Court of Appeal (*R v Taueki*, 2005) is a specific example of a planning heuristic (Simon, 1976). Sentencing is broken into three steps (*R v Taueki*, 2005; Hall, 2008; Robertson, 2007). In the first step the judge determines a starting point sentence based on the seriousness of the offence and the offender's culpability. Any personal aggravating and mitigating factors are then taken into account to adjust the sentence in the second step. Finally, in the third step the sentence is further adjusted to take account of a guilty plea, if there was one, and any assistance provided to the authorities. At each step a limited selection of factors is used to progressively determine the sentence.

From time to time judges' sentencing decisions indicate more informal use of the planning heuristic. For example, in reading sentence notes for this research the following statements represent indications of sentencing approaches used by judges for drinking and driving. One judge, referring to prior convictions for drink driving, said "my practice is, unless there are exceptional reasons after three, you go to jail". Another judge offered a slightly more complex approach, saying "the general rule in this area is that if you offend three times or more in the period of ten years we start with imprisonment"¹. The first judge began by looking at a single factor to determine sentence type; the second judge used the same factor (prior convictions) but took into account the age of those convictions. In both examples the judge simplified the sentencing task by focussing on one or two factors before testing the indicative decision against the wider set of factors relevant to the case.

Concluding Comment

The framework for understanding sentencing disparity between locations is applicable to New Zealand. Judges have wide discretion in the determination of sentences and confront the challenges of sentencing. The criminal justice and court systems are configured in a way that is consistent with the opportunity for local sentencing constructs to form. Given the relevance and applicability of the theory to the New Zealand context disparity would be expected between District Court locations.

¹ These quotes are not referenced because this research proceeded on the basis that it would not directly or indirectly result in the identification of a particular judge.

Chapter 3: Sentencing Disparity by Location, the Findings from Empirical Research

There is only a limited body of empirical research investigating sentencing disparity between locations. Almost all of the research has been undertaken in the United States (see Johnson, 2011; Ulmer, 2012; for reviews), with only limited research in other jurisdictions. The United States research has been more sophisticated than elsewhere using multivariate regression techniques to isolate the influence of location and other factors on sentencing outcomes. The research carried out elsewhere has been relatively unsophisticated, often relying on comparisons of sentencing outcomes by location with no controls for legal or extra-legal factors that might explain differences between locations.

Research into sentencing disparities between locations is a relatively recent development. Disparities between ethnic groups, genders and, to a lesser extent, other extra-legal factors have been of greater interest. Kautt (2002, p. 635) expressed surprise at the lack of research focussed on location, noting that: “since the [sentencing] guidelines were expressly intended to minimise the impact of extra-legal factors over the sentencing decision, variation in sentences by geographic location (which is also extra-legal) intuitively should be of particular interest. Yet, this potentially significant source of extra-legal disparity is inexplicably given short shrift by most studies of post guidelines sentencing.” Johnson et al. (2008, p. 741) echoed Kautt, suggesting that “Concern over interjurisdictional variation in federal sentencing is not new; yet interdistrict variation in such disparity remains seriously under-studied”, and suggested researchers had focussed more on controlling for jurisdictional variation than investigating it.

Recent reviews highlighted the need for further research into all aspects of sentencing disparity (Johnson, 2011, Ulmer, 2012), with Ulmer (2005, p. 1501) noting that “the most glaring gap in the literature is that almost all of the research on sentencing disparity is limited to the contemporary North American – particularly U.S. – context.” Influenced by the volume and methodological imbalances between the United States and elsewhere in the empirical research this review of findings is split between the United States research and the research elsewhere. With a small number of exceptions that will be highlighted, the review covers studies that employed multivariate regression analysis. The exceptions arise when there is a paucity of research. Only findings that are statistically significant at the conventional standard (i.e. $p < .05$) are treated as significant in this review, although in limited instances some researchers treated findings as significant if $p < .10$.

United States Research

The review of United States research addresses studies that either focussed directly on location or included it as a control factor in research with a different focus. The review is presented in four parts. The first part is a presentation of the findings of the influence of location on sentencing outcomes. The second part addresses the influence of other variables, both legal and extra-legal, on sentencing outcomes. The third part addresses research findings on the contextual variables that may account for disparity between locations. The fourth part draws together the explanations offered for the influence of location and the influence of contextual variables.

Research examining disparity in sentencing between locations has varied in terms of the legal system examined (federal, state, multi-state), the level of geographic unit explored (circuits, districts, counties or aggregations of these units), the range of offences included (all offences, just felonies or a single offence type), and the extent of controls for relevant variables. Many earlier studies suffer from a failure to adequately measure offence seriousness (e.g., Eisenstein et al., 1988; Myers & Talarico, 1987), thus making any disparity in sentencing between locations difficult to interpret.

The research has primarily focussed on two sentencing decisions. The first decision focuses on the choice of sentence type, which has typically been reduced to a dichotomous choice between incarceration (custody) and other (non-custodial) sentences. This decision is often described as the “in-out” decision, referring to whether the defendant is put into custody or not. The second focus has been on the length of incarceration. To a lesser extent researchers have addressed departure decisions. The decision to depart from sentencing guidelines is most often addressed as a dichotomous choice between departing and not departing but sometimes the magnitude of the departure is the focus of the study.

The variables included in the models have been labelled as “legal” (flowing from formally rational rules) or “extra-legal” (the influence of substantive rationality). Legal factors are those that are relevant to sentencing and *intended* to influence it. Conversely, extra-legal factors are supposed to be irrelevant and should have no influence (Hofer, Blackwell & Ruback, 1999). Demographic variables are typically described as extra-legal because they are not intended to have a direct influence on sentencing decisions.

The influence of location on sentencing outcomes

A literature search identified 31 relevant studies from the United States that included location as a sentencing variable. Sixteen of these studies (listed in Table 3.1) either had location as the prime focus or at least addressed it as more than simply a control factor.

Table 3.1: United States sentencing research actively addressing the influence of location

Author(s)	Year of Publication	Year(s) studied	Offences	Jurisdiction
Britt	2000	1991-1994	General	Pennsylvania
Crow & Gertz	2008	1994-2003	General	Florida
Dixon	1995	1993	Felonies	Minnesota
Iles, Bumphus & Zehel	2011	2006	Drug	Federal (Michigan)
Johnson	2005	1999-2000	General	Pennsylvania
Johnson	2006	1999-2000	Felonies	Pennsylvania
Johnson, Ulmer & Kramer	2008	1997-2000	General	Federal
Kautt	2002	1998-1999	Drug	Federal
Kautt & Spohn	2007	1997-1998	Drug	Federal
Ostrom, Ostrom & Kleiman	2004	1995	General	Michigan
Pasko	2002	1995	Drug	Federal
Sharp	2007	1993-2003	General	Federal
Ulmer	2005	1996-2000	General	Federal
Ulmer & Johnson	2004	1997-1999	Felonies	Pennsylvania
Ulmer & Kramer	1996	1985-1991	General	Pennsylvania
Wiseman & Connolly	2008	2003-2006	Selected offences	Wisconsin

Twelve studies (listed in Table 3.2) treated location as a control factor but did not substantively address the effect it had on sentence outcomes. These studies add depth to the measurement of disparity and the degree to which it is pervasive across state and federal jurisdictions.

Table 3.2: United States sentencing research with controls for location but that did not actively address the influence of location

Author(s)	Year of Publication	Year(s) studied	Offences	Jurisdiction
Albonetti	1997	1991-1992	Drug	Federal
Albonetti & Baller	2010	1998-1999	Drug	Federal
Austin	1981a	1976	Felonies	Iowa
Austin	1981b	1976	Felonies	Iowa
Everett & Wotjkiewicz	2002	1991-1993	General	Federal
Farrell, Ward & Rousseau	2010	2000-2002	General	Federal
Kautt and Delone	2006	1997-1998	Drug	Federal
Myers & Talarico	1987	1976-1985	Felonies	Georgia
Ostrom, Ostrom, Hanson & Kleiman	2008	2002	Varied ^a	Michigan, Minnesota and Virginia
Steffensmeier, Ulmer & Kramer	1998	1989-1992	General	Pennsylvania
Ulmer & Light	2010	2001-2007 (in 3 sections)	General (excl immigration)	Federal
Wheeler, Weisburd & Bode	1982	1976-1978	White collar	Federal

^a The analysis covered all offences for Michigan and Minnesota and assaults and burglary for Virginia.

The 28 studies listed in Tables 3.1 and 3.2 all confine their analyses to a single legal system; that is, they either analyse the federal courts or they analyse a single state, so that the same law applies across all cases. One study, Ostrom et al. (2008), addressed three states but analysed separate models for each state. The studies excluded from this review addressed sentencing variability at county level but did so across a number of different states (Fearn, 2005; Weidner & Frase, 2003; Weidner et al., 2005). None of these studies controlled for the state in which the county was situated and nor was there any discussion of the sentencing systems across the states. Any differences between counties from different states may have been due to differences in legal frameworks. Any such differences could have been, at least partially, controlled for by a categorisation based on the differences in sentencing systems. In the absence of controls it is possible that the

differences found were attributable to sentencing law and policy rather than to the sentences imposed by the local judges or to some combination of the two.

The concept of location appears to have one key common feature. Each location has a unique complement of judges responsible for sentencing. Beyond that observation there is substantial variation in the degree of aggregation or disaggregation that applies. The focus of federal studies ranges from all circuits (e.g., Pasko, 2002), or all districts (e.g., Sharp, 2007), to a selection of districts (e.g., Ulmer, 2005) or an aggregation of circuits into larger regional bodies (e.g., Everett & Wojtkiewicz, 2002). The state based studies focus on counties and range from all counties (e.g., Crow & Gertz, 2008), to a representative selection based on particular criteria (e.g., Ulmer & Kramer, 1996), or an aggregation of counties into groups (e.g., Steffensmeier et al., 1998), and some drew a limited comparison between one or more specific counties and an aggregation of the rest of the counties in the state formed into a single unit (e.g., Wiseman & Connolly, 2008).

With three exceptions the studies do not make any allowance for elements of the sentencing regime. Kautt and Spohn (2002) and Kautt and Delone (2006) separately modelled and compared the results for federal drug offenders who were subject to mandatory minimum sentences and those who were subject to the federal sentencing guidelines with no mandatory minima. Ulmer and Light (2010) separately modelled and compared federal sentencing across three different time periods characterised by differences in the degree to which judges were obliged to follow the sentencing guidelines.

Much of the research has been reported in the 2000s. Fourteen of the 16 studies with an overt focus on location and six of 12 studies controlling for location were reported in the 2000s. There is a relatively even spread between single level (9) and multi-level models (7) among the 16 studies with an overt focus on location.

The studies vary according to the type of sentencing decision analysed. The most common focus is on the in-out decision and incarceration sentence length, with comparatively fewer studies focussed on or including decisions to depart from the guidelines. Eighteen of the 28 studies analysed both the in-out and sentence length decisions separately (two of these also analysed departures). Five studies analysed sentence length (including one that also analysed departures). Two studies analysed the in-out decision (including one that also analysed departures). Two studies focussed solely on departures. Overall, 23 studies included sentence length, 20 included the in-out decision and six included departures. There was one study that could not be readily categorised in these terms: Everett and Wojtkiewicz (2002) formed the in-out decision and length into a

single scale and reported an overall concept of severity. All others focussed on a single decision or took the approach that the in-out decision and the determination of length were separate decisions that may reflect different drivers and should therefore be modelled separately.

The sixteen studies in Table 3.1 all hypothesised that sentencing would vary by location. The hypotheses were commonly based on two propositions. The three focal concerns were theorised to be uniform across locations but the emphasis and subjective interpretation of each concern and the factors used to evaluate the offending would vary between locations. Variation was expected because “the meaning, relative emphasis and priority, and situational interpretations of them is embedded in local court community culture, organisational contexts and politics” (Kramer & Ulmer, 2002, p. 903). In other words, sentencing decisions were expected to vary between locations because the balance between the influence of formally rational rules (guidelines) and substantively rational sentencing concerns varies by location.

In sum there is widespread evidence of variation or inconsistency between locations in support of the hypothesis. The hypothesis is supported across all the sentencing decisions modelled: the decision to incarcerate, the determination of the length of incarceration and the decision to depart from the sentence indicated by the guidelines (e.g., Johnson, 2005; Sharp, 2007; Steffensmeier et al., 1998; Ulmer & Johnson, 2004). Statistically significant parameter estimates for location variables, indicating inconsistency or disparity between locations, were found in 63 of 71 contrasts across the 28 studies that either overtly focussed on or controlled for location. The ratios were highest for departures (100% of contrasts or 14/14) and the in-out decision (92% of contrasts or 24/26), compared to sentence length (81% of contrasts or 25/31). The higher rate for the in-out decision may reflect the fact that it and most of the departure contrasts are binary choices, whereas sentence length is more subtle with the judge fixing the sentence along a continuum.

More than half the contrasts come from three settings. In studies based on Pennsylvania, 100% of contrasts (13/13) found some degree of disparity. In studies based on the Federal District Court covering offending in general, 100% of contrasts (21/21) found some degree of disparity. In studies of Federal District Court sentencing of drug offenders, 85% of contrasts (11/13) found some degree of disparity. The two non-significant contrasts in the sentencing of drug offences in the Federal District Courts were found by Kautt and Spohn (2007) in the limited circumstance of sentence length models addressing the sentencing of offenders subject to mandatory minimum sentences.

Overall there was only one study that did not find any significant effects. Wiseman and Connolly (2008) found no effects for the in-out or sentence length decisions in a study of sentencing in Wisconsin. Their findings may be the result of data limitations that reduced the study down to two locations, Milwaukee versus a composite of the rest of the state. A similar limitation applies to three non-significant contrasts reported by Ostrom et al., (2008) in their comparison of the effectiveness of guidelines in Michigan, Minnesota and Virginia. The examination of disparity was limited to the largest counties versus rest of state aggregations. More nuanced differences may have been lost due to the levels of aggregation in these studies.

Some of the significant results may have been reduced in magnitude by similar issues to those that affected Wiseman and Connolly's (2008) study. For example, Steffensmeier et al. (1998) grouped Pennsylvania's 67 counties into three categories representing large, medium and small courts. The odds of incarceration were a statistically significant 1.38 times greater in small compared to medium courts, and in large courts were only 0.32 times the odds of incarceration in medium courts. Dixon (1995) split the 73 counties in Minnesota into two groupings based on the level of bureaucratisation of the courts and prosecutors' offices. Had all the counties in these studies been represented individually a wider spread of significant differences may have been found and clusters of counties with similar results may have been identified.

The influence of legal and extra-legal factors on sentencing outcomes

A common finding across all studies listed in Tables 3.1 and 3.2 was that legal factors describing aspects of the seriousness of offending and criminal history were the predominant influence on sentencing decisions. A substantial proportion of researchers also found that extra-legal factors representing ethnicity and gender have also influenced sentencing.

The influence of ethnicity across studies is demonstrated in a narrative review (Spohn, 2000) and a meta-analysis of effects (Mitchell & Mackenzie, 2004). Based on a review of 40 studies published from 1983 to 2000, Spohn (2000) found blacks were sentenced more severely than whites 41 times, more leniently 4 times and there was no significant difference in 50 of 95 contrasts. Hispanic offenders were sentenced more severely than whites 8 times, more leniently twice and there was no significant difference in 19 of 29 contrasts. At Federal level blacks were sentenced more severely in 15 of 22 contrasts and Hispanics more severely in 10 of 21 contrasts. Neither blacks nor Hispanics

were sentenced more leniently in any of the Federal level studies. Based on the totality of these results Spohn (2000) concluded that black and Hispanic offenders sentenced at state level and in the Federal system face higher odds of incarceration and in some jurisdictions longer sentences and lower chances of downward departure.

Mitchell and Mackenzie (2004) took the assessment of the wider body of research a step further, undertaking a meta-analysis of sentencing studies at state and Federal level. The analysis covered 122 distinct sentencing contexts drawn from 71 studies. The results across the studies were mixed, with a high proportion finding that minorities were sentenced more severely. African Americans were sentenced more severely than whites in 83% of the contrasts across the Federal system and 77% of contrasts across state systems and Latinos were sentenced more severely than whites in 72% of the contrasts across both Federal and state systems.

Twenty-four of the studies from the United States included in this review controlled for gender². Females were significantly less likely to be incarcerated in 29 out of 33 contrasts in 20 studies; were sentenced to significantly shorter terms in 22 out of 26 contrasts in 21 studies; and more likely to receive downward departures in 3 studies (Johnson, 2005; Johnson et al., 2008; Sharp, 2007) and less likely to receive upward departures (Johnson, 2005). Johnson et al. (2008) found that the magnitude of downward departures was larger for females.

A summarised review of the influence of extra-legal factors on sentencing is attached at Appendix A. The effects of gender and ethnicity are typically small by comparison to offence seriousness and criminal history factors (Johnson, 2011).

The combination of findings of disparity between locations, the predominant influence of legal factors and the influence of extra-legal factors led some researchers to hypothesise that the interpretation and use of legal and extra-legal factors varies between locations. To test the hypothesis researchers used multi-level models to separate the fixed and random effects components for both legal and extra-legal factors. Six of the studies with an overt focus on location addressed random effects (Crow & Gertz, 2008; Johnson, 2005 & 2006; Johnson et al., 2008; Kautt, 2002; Ulmer & Johnson, 2004). Without

² The studies were: Albonetti, 1997; Albonetti & Baller, 2010; Britt, 2000; Crow & Gertz, 2008; Everett & Wojtkiewicz, 2002; Farrell et al, 2010; Fearn, 2005; Johnson, 2005 & 2006; Kautt, 2002; Kautt & Delone, 2006; Ostrom et al., 2004; Pasko, 2002; Sharp, 2007; Spohn, 2000; Steffensmeier et al., 1998; Ulmer, 2005; Ulmer & Johnson, 2004; Ulmer & Kramer, 1996; Ulmer & Light, 2010; Weidner & Frase, 2003; Weidner, Frase & Schultz, 2005; Wheeler, Weisburd & Bode 1982.

exception the research found that the influence of individual legal and extra-legal factors varied significantly between locations. The findings span a number of jurisdictions and decision types. In the Federal District Courts significant variation in the influence of legal and extra-legal factors on sentence length (Kautt, 2002) and the likelihood of departure (Johnson et al., 2008) was reported. In Florida significant variation in the influence of legal and extra-legal factors on the likelihood of incarceration and sentence length was recorded by Crow and Gertz (2008). In Pennsylvania significant variation in the influence of legal and extra-legal factors on the likelihood of incarceration and sentence length (Johnson, 2006; Ulmer & Johnson, 2004) and the likelihood of departure (Johnson, 2005) was reported.

Modelling the influence of context on sentencing variation between locations

Taken as a whole it is apparent from the United States literature that there are significant disparities between locations. According to Ulmer (2012, p. 14) “*In sum, substantial evidence exists that what kind of sentence one gets, and the factors that predict why one gets it, in significant part depends on where one is sentenced*” (emphasis in the original).

Having determined that sentencing varies by location and that the variation is embedded in the local context some researchers (Britt, 2000; Crow & Gertz, 2008; Johnson, 2005 & 2006; Johnson et al., 2008; Kautt, 2002; Ulmer & Johnson, 2004) extended their models to identify contextual variables associated with disparity between locations.

There are common themes across the research. The following discussion is organised thematically drawing together similar theoretical bases for hypotheses, the measures used to test the hypotheses and the results from across the studies. The discussion begins with an examination of effects from the wider community; followed by an examination of the influence of the size of the location; and finally an examination of effects associated with the courthouse community. Size has been separated out because it relates to both the wider community, with larger populations associated with larger courts, and the courthouse community, with differing dynamics between judges, prosecutors and defence counsel depending on numbers.

Racial threat: the influence of ethnic minorities on sentencing

The most pervasive theory addressing wider community contextual influences across the studies is the proposition that sentencing may vary between locations according to perceptions of racial threat (Britt, 2000; Crow & Gertz, 2008; Johnson, 2005 & 2006; Johnson et al., 2008; Kautt, 2002; Ulmer & Johnson, 2004). It is hypothesised that ethnic or racial minorities are perceived to be a threat to the established order and as a result are sentenced more severely. Black and Hispanic offenders feature more in media coverage of crime enhancing fear of crime and racial typification of crime (Britt, 2000; Johnson et al., 2008). Racial threat hypotheses suggest that the greater the proportion of black people or Hispanic people in a location the more severe sentencing will be in that location. Some researchers tested the hypothesis by including dummy variables for both the percentage of black people in the location and the percentage of Hispanic people in the location (Crow & Gertz, 2008; Johnson, 2005; Kautt, 2002; Ulmer & Johnson, 2004), or just the percentage of black people (Britt, 2000; Johnson et al., 2008). Britt (2000) also employed a measure of ethnic heterogeneity³. Britt hypothesised that more ethnically heterogeneous locations (those with larger minority populations) would sentence more severely. In addition Johnson (2006) included the percentage of Hispanic people as a control variable.

The results were mixed with statistically significant effects found in four of the seven studies. Contrary to the racial threat hypothesis, Crow & Gertz (2008) found that the likelihood of incarceration was lower in Florida counties with a higher percentage of black people in the population but had no effect on sentence length. The authors queried whether this was due to the proposition that once a racial minority reaches a certain threshold it starts to gain political power and hence racial disparity may disappear. A limited examination of counties with larger proportions of racial minorities did not support the threshold proposition and the authors suggested that further research was required. Britt (2000) found that the likelihood of incarceration was higher but sentence lengths were shorter in Pennsylvanian counties with a higher proportion of black people in the population. The sentence length finding was unexpected. Britt (2000) suggested it may be a selection effect as counties with higher rates of incarceration may have shorter sentence

³ Ethnic heterogeneity was measured by subtracting the sum of the squares for the proportion of the population made up by each of five ethnic groups from 1. The highest possible score with five groups is .8, this occurs when each of the five groups is an equal proportion of the population and means that the population is fully ethnically heterogeneous. The lowest possible score is 0; this score occurs when a single group makes up the entire population and the other groups are not represented and means that the population is fully ethnically homogenous.

lengths because less serious offending has resulted in incarceration. Britt (2000) found the level of racial heterogeneity had no significant influence on either the likelihood of incarceration or on sentence length. Ulmer & Johnson (2004) in a later study did not replicate Britt's (2000) findings, instead reporting that the percentage of black people had no significant influence on the likelihood of incarceration or sentence length between counties in Pennsylvania. Johnson (2005) found that sentencing was more severe in Pennsylvanian counties with a higher percentage of Hispanic people in the population. In addition, Johnson (2006) found that sentence lengths were longer in Pennsylvania in locations with a higher percentage of Hispanic people in the population but there was no effect on the likelihood of incarceration. However, the percentage of black people had no effect on sentence lengths imposed on drug offenders in the Federal District Courts (Kautt, 2002), and no effect on rate or magnitude of departure between Federal District Courts (Johnson et al., 2008). Similarly, two studies found no effect for the percentage of Hispanic people on the likelihood of incarceration or sentence length in Pennsylvania (Ulmer & Johnson 2004), and Kautt (2002) found no effect on sentence lengths for drug offenders sentenced in the Federal District Courts.

The combined results are difficult to reconcile. There is evidence that the size of racial minorities has an influence on the severity of sentencing between locations. Three studies provided evidence of greater severity (Britt, 2000; Johnson, 2005 & 2006), although Britt (2000) found evidence of greater leniency but suggested it may be a selection effect. One study (Crow & Gertz, 2008) did find greater leniency. The findings may reflect a jurisdictional difference with greater severity demonstrated in studies of sentencing in Pennsylvania and lenience in a study of Florida sentencing. Crow and Gertz's speculation of a threshold effect may explain the difference between the states. The most common effect across the studies was that the size of racial minorities had no significant influence on sentencing severity between locations. Further research is required.

Economic threat: the influence of relative levels of disadvantage on sentencing

The second theory addressing wider community contextual influences across the studies is the proposition that sentencing may vary between locations according to perceptions of economic threat (Britt, 2000; Crow & Gertz, 2008; Johnson, 2005 & 2006; Johnson et al., 2008; Kautt, 2002). Economically disadvantaged groups are perceived to be a threat to the established order and as a result are sentenced more severely. Johnson et al., (2008)

theorised, based on a neo-Marxist perspective, that greater socio-economic stratification increases social conflict leading to more punitive sentencing to control and contain the higher proportion of less advantaged members of the community. Economic threat hypotheses suggest that the greater the proportion of economically disadvantaged people in a location the more severe sentencing will be in that location. Most researchers (Crow & Gertz, 2008; Johnson, 2005⁴ & 2006; Kautt, 2002) tested the hypothesis by including the unemployment rate in models. Johnson et al. (2008) used a composite measure of socio-economic disadvantage that combined median income per capita and the percentage of the population below the poverty line. Britt (2000) employed four measures: per capita income, racial income inequality, the unemployment rate, and the trend in the unemployment rate.

The results were mixed with statistically significant effects supporting the hypothesis in two of the six studies. Crow and Gertz (2008) found that higher unemployment rates in Florida counties were associated with higher rates of incarceration and longer sentences. Johnson et al. (2008) found that there was a decreased likelihood of departure in those districts of the Federal District Court with higher levels of socio-economic disadvantage but that there was no effect on the magnitude of departures. Britt (2000) found no effects on the likelihood of incarceration or sentence lengths between Pennsylvania counties based on differences in per capita income, racial income inequality, the unemployment rate, or the trend in the unemployment rate. The unemployment rate had no effect on sentence lengths for drug offenders sentenced in the Federal District Courts (Kautt, 2002), the likelihood of incarceration or sentence length in Pennsylvania (Johnson, 2006), or the likelihood of departure in Pennsylvania (Johnson, 2005). The results provide partial but limited support for the hypothesis that sentencing is more severe in locations with greater economic disadvantage.

Political context: the influence of local political leanings on sentencing

The third perspective on the wider community context was the political context, described by Crow & Gertz (2008) as the local democratic sub-culture, the politics, attitudes and major values of the people in the community. It was hypothesised that courts in more conservative locations would sentence more severely than courts in more liberal locations. The measures of conservatism used varied between the studies. Crow and Gertz (2008)

⁴ The unemployment rate was included as a control rather than as an explanatory variable.

and Johnson (2006) used the percentage of electors registered as Republicans, Ulmer and Johnson (2004) used the percentage of Pennsylvanian electors voting for the Republican candidate in the 1996 presidential election, and Johnson et al. (2008) used a liberalism score, developed by the American Civil Liberties Union, based on voting patterns in the United States Senate.

The results were mixed with statistically significant effects supporting the hypothesis in two of the four studies. Crow and Gertz (2008) found that higher percentages of registered voters identified as Republican were associated with higher rates of incarceration in Florida counties but had no effect on sentence length. Johnson et al. (2008) found that the more liberal a district of the Federal District Courts was the more likely it was to grant downward departures (i.e., to sentence more leniently) but there was no effect on the magnitude of departure. Measures of conservatism did not have any significant effects on the likelihood of incarceration or sentence lengths in Pennsylvania (Johnson, 2006; Ulmer & Johnson, 2004). The results provide partial support for the hypothesis that sentencing is more severe in more politically conservative locations.

The effects of the size of the location

Five studies included hypotheses that court size (judge numbers or population served) is related to sentence severity, with larger courts hypothesised as being more lenient (Kautt, 2002; Johnson, 2005 & 2006; Johnson, et al., 2008; Ulmer & Johnson, 2004). In four of these studies size was conceptualised based on measures of the courthouse community. Johnson (2005) split courts into three categories (large, medium and small) but did not describe the measures or thresholds used to separate the categories. Johnson (2006) and Ulmer and Johnson (2004) used the same three categories according to a mix of judge numbers and proportion of cases. Johnson et al. (2008) conceptualised size as judge numbers. Kautt (2002) diverged from the other researchers and conceptualised size as population served. Despite the difference in the measures used the choice potentially makes little difference if population and judge numbers rank courts in a similar way. Britt (2000) did not address size directly but applied a measure of urbanisation based on the proportion of the population living in a census defined urban area. The study has been included in the discussion of size for two reasons. Britt (2000) noted that the more rural courts tended to be smaller, and secondly used similar arguments to link it to a sentencing effect as the others used to link size to sentencing. Britt (2000) did not offer a specific

hypothesis of the effect of urbanisation on sentence outcomes but just hypothesised that there would be an effect.

The hypotheses are based on a range of inter-related propositions. Based on propositions first put forward by Eisenstein et al. (1988) larger and urban courts are said to be more bureaucratic, have less social interaction between members of the courthouse community, are less visible in dealing with routine offending, offer greater ease of plea bargaining, and are exposed to a larger, more diverse amount of deviant behaviour and hence are more tolerant of crime (Johnson et al., 2008). Kautt (2002) argued that larger populations are likely to be associated with a desire for faster disposals to maintain the viability of the court leading to greater leniency in order to induce resolution of cases.

The results were mixed with statistically significant effects found in four of the six studies. Three studies supported the hypothesis. Ulmer and Johnson (2004) and Johnson (2006) found that offenders were less likely to be incarcerated in larger counties in Pennsylvania but found no effect on sentence length. Johnson (2005) found that offenders were more likely to be granted downward departures in larger counties in Pennsylvania. Conversely, Britt (2002) found offenders sentenced in more urbanised (larger) counties in Pennsylvania were more likely to receive longer sentences but found no effect in relation to the likelihood of incarceration. Neither of the studies focussed on districts of the Federal District Court found any effect associated with size. Kautt (2002) found no effect on sentence length for drug offenders, and Johnson et al. (2008) found no effect on the likelihood of departure or the magnitude of departure for offenders across all offences.

The overall results provide moderate support for the hypothesis that sentencing is more lenient in larger locations. The hypothesis was supported in three of five studies that directly addressed size (Johnson, 2005 & 2006; Ulmer & Johnson, 2004) but not in one study which used a measure of urbanisation rather than size (Britt, 2000). Support for the hypothesis in the five studies focussed on direct measures of size was limited to decisions involving a binary choice (to incarcerate or not, or to depart or not) and not decisions affecting the quantum of the sentence. The hypothesis was supported by studies based on sentencing in Pennsylvania but not studies of federal sentencing.

The effects of contextual variables from the courthouse community

A range of measures were applied to identify contextual variables from within the courthouse that may form part of the explanation for variation of sentencing outcomes between locations. The identification of variables was informed by two overlapping perspectives, organisational maintenance and practical constraints.

Organisational maintenance focuses on the smooth running of the courts. Courts that are under pressure are hypothesised to sentence more leniently in order to induce case disposals. This hypothesis mirrors one of the rationales for the influence of size. It is logical to expect a relationship between the size of the caseload and the size of the court. Examination of caseload pressures is not limited to the size of the caseload but in some studies also takes into account the trial rate. Courts with higher trial rates are assumed to be under more pressure than those with lower trial rates.

Practical constraints focus on factors that limit or influence the choice of sentencing outcomes. It is one of the three focal concerns posited by Steffensmeier et al. (1998) that may influence sentence outcomes. Caseload pressure can be construed as a practical constraint (see Ulmer & Johnson, 2004) affecting organisational efficiency linked to the maintenance of the organisation (court) as an effective operation. In addition researchers have also addressed the influence of available jail capacity as a practical constraint on the choices available to judges.

The influence of caseload pressures on sentencing

Caseload pressure has been conceptualised in two ways: the pressure due to the volume of new cases per judge; and the pressure due to the rate of cases proceeding to trial. In three studies, courts with a higher number of cases per judge were hypothesised to be more lenient (Johnson, 2005; Johnson et al., 2008; Ulmer & Johnson, 2004). All three studies found support for the hypothesis. Johnson (2005) found that higher caseloads increased the likelihood of downward departure but had no effect on the likelihood of upward departures between counties in Pennsylvania. Johnson et al. (2008) found that higher caseloads increased the likelihood of downward departure between districts of the Federal District Court but conversely that the magnitude of departures reduced. The authors suggested that these contrary effects may be the result of a selection effect with higher rates of departure associated with less meritorious cases and hence reduced magnitudes. Ulmer and Johnson (2004) found that higher caseloads reduced the likelihood of

incarceration but had no effect on the length of incarceration between counties in Pennsylvania.

When caseload pressure is conceptualised as cases per judge there is evidence to support a hypothesis that as the number of cases per judge increases in a location sentencing will be more lenient in that location. All three studies (Johnson, 2005; Johnson et al., 2008; Ulmer & Johnson, 2004) found evidence of greater leniency in the rate of departures or in the rate of incarceration. While Johnson et al. (2008) also found evidence of greater severity - smaller deductions for downward departures - the result may have been due to a selection effect. Higher levels of departure after controlling for sentencing factors means that less deserving cases resulted in departures and smaller reduction in sentence length. Support for the hypothesis in the three studies was limited to decisions involving a binary choice (to incarcerate or not, or to depart or not) and not decisions affecting the quantum of the sentence.

Trial rates are also expected to influence sentencing outcomes between locations. Johnson et al. (2008) and Johnson, (2005) hypothesised that offenders in locations with higher trial rates would be less likely to receive downward departures and the magnitude of departures would be lower. The rationale for this hypothesis was that judges would be more likely to reward offenders who entered guilty pleas than those who imposed greater pressure on the courts by going to trial. The hypothesis was not supported; trial rates had no significant influence on either the likelihood of downward departure or the magnitude of departure in either study. Ulmer and Johnson (2004) included county level trial rates as a control variable in multi-level models of the likelihood of incarceration and sentence length in Pennsylvania. Trial rates did not have a significant influence on either decision. When caseload pressure was conceptualised as trial rate, there was no evidence to support the hypothesis that as the trial rate rises in a location sentencing will be more severe in that location.

The influence of available jail capacity on sentencing

Available jail capacity has been hypothesised to increase the rate of incarceration and sentence length (Johnson, 2006; Ulmer & Johnson, 2004), and reduce the likelihood of downward departures in locations with greater capacity (Johnson, 2005). In all three studies, available jail capacity was calculated as the ratio of the number of beds to the number of cases. Both Johnson (2006) and Ulmer and Johnson (2004) reported that the rate of incarceration was significantly greater in locations with greater jail capacity but that

there was no effect on sentence length. Johnson (2005) found there was no effect on departure rates from jail capacity. There was moderate support for the hypothesis but, in keeping with the findings on size and caseload pressure, the support was limited to decisions involving a binary choice (to incarcerate or not, or to depart or not) and not decisions affecting the quantum of the sentence.

Summary of the research investigating the influence of contextual factors

The research found support for a range of hypotheses that specified contextual variables will influence (be associated with) variations on sentencing decisions between locations. The findings support contextual variables that originate from within the courthouse and the wider community.

Because the findings come from just three jurisdictions (Florida, Pennsylvania and the Federal District Court) with four of the seven studies examining sentencing in Pennsylvania the degree to which the conclusions can be generalised is limited. The identification of the contextual factors that may be associated with sentencing disparities between locations requires more research. The reported studies suggest that court size and caseload pressure (cases per judge) should certainly be included and a moderately strong case can be made for inclusion of the ethnic composition of the population. None of the variables tested can be firmly rejected.

The contextual variables investigated in the studies reported here provide an indication that context has greater influence on binary decisions (to incarcerate or not, or to depart or not) than on sentence lengths or the magnitude of departures. Binary decisions accounted for 15 of the significant findings compared to four for sentence length type decisions. It may be that selection effects, as mooted by Britt (2000), confound the effects for sentence length type decisions. Higher rates of incarceration in some locations mean that less serious cases attract incarceration with a concomitant downward effect on sentence lengths in locations expected to sentence more severely.

Research Outside of the United States

Few studies from outside the United States have addressed the influence of location on sentence outcomes. Of the studies that have been completed there is a lack of adequate controls for other sentencing factors. The most significant body of sentencing disparity research has been a series of studies investigating sentencing of indigenous offenders in Australia (Bond, Jeffries & Weatherburn, 2011; Bond & Jeffries, 2011; Jeffries & Bond

2009; Snowball & Weatherburn, 2006, 2007). Although these studies have not controlled for location they do provide useful conclusions on the influence of legal and extra-legal factors on sentencing. The findings from the studies are included in Appendix A.

Canadian and United Kingdom research on the effect of location on sentence

Although location was not addressed or controlled for in any of a series of Canadian studies of the sentencing of indigenous offenders a review of one study shed light on the effects of location suggesting that the authors may have erred in not accounting for location. Latimer and Foss (2005) addressed the sentencing of aboriginal and non-aboriginal youth using a sample of youth court cases drawn from five large cities. They found that aboriginal offenders were significantly more likely to be sentenced to longer custodial terms. Doob and Sprott (2007) were surprised by this finding and re-analysed the data, finding that the effect was due to an uneven distribution of aboriginal offenders across the five cities. The distribution mattered because sentencing severity varied between the cities. Aboriginal youth were only present in significant numbers in the three cities that imposed the longest custodial sentences for any offender irrespective of ethnicity. According to Doob and Sprott (2007) the real position was that young offenders were receiving longer sentences in some cities compared to others. There was disparity between cities not ethnicities. At the very least these findings highlight the necessity of controlling for location when data is pooled across disparate locations.

No multivariate studies could be located for the United Kingdom. Tarling (1979, 2006) has twice reviewed sentencing practice in the Magistrates' Courts in England and Wales. A substantial range of imprisonment rates was found. In 1975 the range was 3.1% to 19.1% (Tarling, 1979) and in 2000 it was 8.6% to 27.7% (Tarling, 2006). Tarling (2006) recognised that the variation may be due to differences in the type of offences and offenders sentenced. A limited attempt was made to investigate this by separately focussing on two offence types, burglary and shoplifting. The resultant range of imprisonment rates for each offence was wider than for all offending. These results indicate that geographic disparity may be present but did not establish that it is.

In 2007 the Ministry of Justice in the United Kingdom completed a study of local variation in sentencing in England and Wales (Mason et al., 2007). The study addressed the use of imprisonment and average custodial sentences by the Magistrates' Courts and by the Crown Court in 2006 across 42 criminal justice areas. In the Magistrates' Courts imprisonment rates varied from 6% to 16% and average custodial sentence lengths from

2.4 to 3.6 months (this was against a cap of six months, the maximum sentence a magistrate may impose). In the Crown Court imprisonment rates varied from 45% to 68% and average custodial sentence lengths from 19.5 to 28.3 months. The study used bivariate analyses to determine the degree to which the imprisonment rates and average sentence lengths were associated with offence mix, the rate of committal of cases to the Crown Court (an indication of greater perceived seriousness), crime rates, and changes in the crime rate. The study found that variation in the mix of cases partially explained the Crown Court range and that variation in the crime rate provided a limited explanation for variation in the imprisonment rate in the Magistrates' Courts. Mason et al. (2007) concluded that factors like offence seriousness and offender characteristics only explained part of the variation in sentencing and that the balance may be due to variations in local court cultures. The results indicate that disparity between locations may be present but do not conclusively establish that it is.

New Zealand research on the effect of location on sentence

There has only been limited research into the effects of location on sentencing in New Zealand (Law Commission, 2006; Spier, 1989). In 1989 the Department of Justice undertook a study of regional differences in the use of custodial sentences in the District Courts and found "significant variations between the district court areas in the proportion of offenders given a custodial sentence and in the average custodial sentence length imposed" (Spier, 1989, p. 49) for seven of the eight offence types studied. The study, however, employed minimal controls due to a lack of available data.

Furthermore, Spier's research was undertaken over 20 years ago and may not be a good indicator of the current position due to changes to the sentencing regime. First, the Sentencing Act 2002 codified the sentencing framework and made it more accessible. This may have resulted in a greater convergence between circuits. Secondly, the range of sentences was broadened in 2007. Two new sentences (community detention and intensive supervision) were introduced and home detention became a sentence available to a judge rather than a way of serving a sentence of imprisonment at the discretion of the Parole Board. The new sentences filled a gap in the sentence structure between community work and imprisonment, providing judges with greater choice in respect of offenders on the cusp of imprisonment, and potentially reducing the degree of variation.

The degree to which the expectation of consistency is being met in New Zealand was questioned by the Law Commission in 2006. The Commission claimed there was a great deal of anecdotal evidence of inconsistency between locations. Empirical research undertaken for the Commission compared imprisonment rates and terms of imprisonment for a series of higher volume offences sentenced in 2004 and 2005 across regions of the District Courts. The Commission believed that the results revealed variations so large it was unlikely they reflected differences in the seriousness of the cases and concluded that some regions were systematically more severe than others (Law Commission, 2006). The research undertaken for the Law Commission was methodologically limited. It was a comparison between regions on the use of imprisonment and the length of imprisonment. The only form of control was to limit the comparison to regions with higher volumes. Volume was relied on as a means of enabling a “robust analysis”, leading to the conclusion that the variations were “unlikely to be explicable on the basis of differences in offence or offender variables” (Law Commission, 2006, p. 20). The Law Commission (2002) attributed the observed variations between locations to the wide discretion afforded to judges when sentencing.

The influence of factors other than location on sentencing outcomes

The findings from the United States on the primary influence of legal factors on sentencing outcomes were replicated in the multi-variate research in other jurisdictions. While much of the research in the other jurisdictions that employed controls for legal factors were addressing other forms of disparity or sentencing in general the findings are instructive. Canadian research which focussed on the influence of indigeneity on sentencing in a variety of contexts found that offence severity and criminal history explained all or a substantial part of the variation in sentences (Latimer & Foss, 2005; Weinrath, 2007; Welsh & Ogloff, 2008). Similarly, the series of Australian studies across several states on the influence of indigeneity found that offence severity and criminal history explained all or most of the variation in sentence decisions (Bond, Jeffries & Weatherburn, 2011; Bond & Jeffries, 2011; Jeffries & Bond 2009; Snowball & Weatherburn, 2006, 2007).

In New Zealand, a major statistical study of sentencing (Triggs, 1999) found that seriousness and criminal history were the most significant factors, ethnicity was not significant but gender was. A study comparing sentencing of men and women for more serious offences in Christchurch found that a series of offence seriousness and criminal history factors explained apparent differences in the in-out decision (Jeffries, Fletcher &

Newbold, 2003). The same factors also had a substantial influence in determining sentence length but the sex of the offender remained significant (Jeffries et al., 2003).

Summary

The empirical research almost universally supports the proposition that sentencing outcomes within the same jurisdiction in the United States are dependent on where an offender is sentenced. The research universally finds that formally rational or legal factors, predominantly associated with offence seriousness and criminal history, are the most influential determinants of sentence. A significant proportion of studies also find that extra-legal factors, particularly gender and ethnicity, also influence sentencing decisions in some jurisdictions.

In combination these findings suggest that the interpretation and use of both legal and extra-legal factors varies between locations. In turn this suggests that sentencing decisions are influenced by contextual differences between locations. Using multi-level models a subset of the researchers universally found support for the proposition that the interpretation and use of both legal and extra-legal factors varies between locations. Researchers subsequently expanded the models to explore which variations in the circumstances or contexts of the locations are associated with sentencing variation and therefore act as influences on local sentencing decisions. The results from this research have been much more equivocal. Within the limited range of studies the strongest results were for the size of the location and caseload pressure (expressed as cases per judge). With the exception of caseload pressure (expressed as trial rate) there were varying degrees of support for all the factors considered. It is evident that more research is required before firm conclusions can be drawn on the contextual variables that have the greatest influence on variation in sentencing decisions between locations. The results of the research exploring contextual variables suggests that the influence of substantively rational sentencing concerns do not just vary between locations but that they also vary between sentencing systems. This possibility is evidenced by the differences in the findings between Florida, Pennsylvania and the Federal District Courts.

Chapter 4: Predictions, Research Questions & Statistical Methods

The United States research and the limited research from New Zealand and elsewhere suggests there will be inconsistency in sentencing decisions between locations in New Zealand. Although the sentencing regimes differ between New Zealand and the United States the difference means that inconsistency is arguably more likely in New Zealand. The key difference is the widespread use of sentencing guidelines in the United States. The guidelines provide guidance on sentence levels that are directly related to offence seriousness and criminal history. As a consequence judicial discretion is more restricted in the United States than in New Zealand. Multi-variate research is required to determine the extent, if any, of variation in sentencing between locations in New Zealand.

In keeping with the pattern of research that has evolved in the United States the research will proceed in a sequence of three steps. First, the research will address whether there are statistically significant and substantively significant differences in the severity of sentencing between locations after controlling for compositional differences in caseloads. Second, the research will examine whether the influence of legal and extra-legal factors on sentencing decisions varies between locations. Third, the research will explore associations between variations in context and variations in sentencing between locations.

Sentencing consistency will be addressed via a series of regression models separately applied to the sentencing of aggravated drink drive, male assaults female and burglary offenders in the circuits of the New Zealand District Courts in 2008-2009. These offences were chosen because large numbers (for New Zealand) of offenders are sentenced for each offence and they represent different categories of offending. A more detailed discussion of the choice of courts, years, selected offences, the unit of analysis, other inclusions and exclusions, data sources and the variables to be included in the models is found in Chapter 5.

The modelling will include controls for gender and ethnicity as extra-legal factors. However, as disparities between the genders and ethnicities are not the focus of this research they will not be the subject of detailed analysis. However, both variables are likely to be of interest to readers and will be included in the modelling and reported whether they are statistically significant or not.

Predictions

Consistent with the framework for understanding disparity in sentencing and prior research it is predicted that sentencing will be inconsistent between District Court circuits and there will be variations in the interpretation and use of sentencing factors between circuits.

Judges in New Zealand have wide sentencing discretion and are confronted by the challenges of sentencing (complexity, incomplete information, uncertainty and time pressure) that suggest judges could be subject to bounded rationality. The organisation of the courts into circuits, with each circuit having a separate cadre of judges enables distinct courthouse communities to form and places the judges within distinct communities differentiated by socio-economic, political and caseload factors. To the extent that, consistent with theory, sentencing is a balance between formal rationality and substantive rationality, the influence of substantive rationalities can vary between circuits.

The combination of court community theory (Eisenstein & Jacob, 1977; Eisenstein et al., 1988) and personal construct psychology put forward by Ostrom et al. (2004) explains how substantive rationalities based on local context interact with the formal rationality of the law and other rules resulting in the development of different sentencing approaches between locations. The commonality and sociality corollaries from personal construct psychology explain how personal sentencing constructs can converge into distinctive local sentencing constructs. Decision makers who share a similar work environment and similar cultural norms share similarities in how they “make sense” of their experience (the commonality corollary). The similarities are likely to be strengthened by the sociality corollary. Judges who sit in the same courts and share the same common room may, like any group of individuals who work together, anticipate the behaviour of others and develop and adopt practices based on group expectations and beliefs. The tendency for sentencing approaches to converge locally is likely to be reinforced by the behaviours of prosecutors and defence counsel who with the judges comprise the courthouse community. The expectations and submissions of local prosecutors and defence counsel are likely to be based on local conditions and sentencing practices thus informing and reinforcing local norms and going rates.

The framework for understanding sentencing disparity is generalisable and applicable to New Zealand. The underlying conditions (complexity, uncertainty and time pressure) that trigger the application of Simon’s theory of bounded rationality are present in New Zealand. Judicial discretion is substantial in New Zealand. The expression of formal rationality in statute (the Sentencing Act 2002) focuses more on the sentencing

process, sets out broad purposes, non-exhaustive lists of factors and relies on appellate authority to provide more specific guidance. Thus the law provides little direct offence specific guidance. Unlike guideline systems the law does not point to offence specific factors expected to be used to determine sentences. Instead the law lists broad concepts and generalised guidance on what to take into account. With the exception of a limited range of guideline judgments, which do not apply to higher volume offences, the influence of appellate authority is limited. Judges are largely left to determine their own sentencing approaches, these approaches are influenced by substantive rationalities based on local context potentially leading to a convergence of personal approaches into local circuit level approaches. Two predictions flow from the theory and the empirical research.

Prediction 1:

The circuit in which an offender is sentenced will affect the likelihood of incarceration. After controlling for seriousness of offence, criminal history and offender demographic variables, similar offenders will be sentenced differently between circuits.

Prediction 2:

The circuit in which an offender is sentenced will affect the length of incarceration sentences. After controlling for seriousness of offence, criminal history and offender demographic variables, similar offenders will be sentenced to different lengths of sentence between circuits.

Subject to one further condition being met and assuming that the primary predictions of inconsistency between circuits are confirmed, two further predictions arise. The additional condition is that, consistent with prior research, the testing of the primary predictions demonstrates that seriousness of offence and criminal history variables are the principal influences on the determination of sentence. If both conditions are met it follows that the interpretation and use of offence seriousness and criminal history factors varies between circuits. The formal rationality of the law and appellate authority is conditioned by local substantively rational sentencing concerns. If the testing demonstrates that sentencing is significantly influenced by any extra-legal factors the influence of these factors may also vary between circuits. This reasoning leads to two predictions.

Prediction 3:

If the circuit in which an offender was sentenced affected the likelihood of incarceration and the seriousness of offence and criminal history variables were the principal influence on incarceration, the interpretation and use of some or all of the seriousness of offence, criminal history or extra-legal variables will vary between circuits.

Prediction 4:

If the circuit in which an offender was sentenced affected the length of incarceration and the seriousness of offence and criminal history variables were the principal influence on length, the interpretation and use of some or all of the seriousness of offence, criminal history or extra-legal variables will vary between circuits.

Research Questions

If the predictions are supported the theories based on courthouse communities (Eisenstein & Jacob, 1977) and wider court communities (Eisenstein et al., 1988) suggest that the variation is likely to be associated with contextual differences between circuits. The results of prior research have found the identification of community level factors to be elusive. Identification of the contextual factors linked to sentencing variations provides some insight into the otherwise invisible influences on the exercise of substantive rationalities on sentencing. Predictions on what factors might be associated with sentencing outcome variations between circuits in New Zealand cannot be made. However, the influence of contextual differences on sentencing variation between circuits can be explored. Two research questions addressing the association between circuit context and sentencing approach will be addressed if either Prediction 1 or Prediction 2 is supported.

Research Question 1:

What circuit level contextual variables, if any, are associated with any circuit level variation in the likelihood of incarceration?

Research Question 2:

What circuit level contextual variables, if any, are associated with any circuit level variation in the length of incarceration sentences?

Two Stage Modelling Strategy

A two stage modelling strategy was developed to test the predictions and answer the research questions. The first stage addressed the two primary predictions that the circuit an offender was sentenced in would influence the likelihood and length of incarceration. The second stage addresses the predictions that the influence of some or all of the offence seriousness, criminal history and other variables vary between circuits, and the research questions exploring contextual differences between circuits associated with sentence variation between circuits.

Stage one: single level modelling

Stage one sought to bring a more definitive empirical perspective to the debate about the existence (and extent) of sentencing inconsistency between circuits. Single level logistic regression (in-out decision) and linear regression (sentence length) models, using SPSS version 18, with case level variables and circuit as explanatory variables, were analysed to determine whether and to what extent there is inconsistency of sentencing between circuits.

At the design stage consideration was given to the use of multi-level random intercept models with case level variables as the first level and the circuits as the second level. Because the cases are nested within circuits there were good theoretical reasons for the use of a multi-level modelling approach. The concept of nesting is often explained by reference to the education sector and research involving students, classes and schools (e.g., Raudenbush & Bryk, 2002). In education research focussing on students across classes and schools it is recognised that the achievement of students will be partially dependent on the classes they are in or the schools they attend. Failure to account for the influence of classes or schools can result in models that are mis-specified and produce erroneous findings. This arises because statistical dependencies are likely to arise among clustered observations and they potentially share similarities that are not taken into account. As a consequence residual errors may be correlated and the assumption of independence of errors violated, resulting in underestimation of standard errors and overestimation of statistical significance. Multi-level models have increasingly been used in sentencing research (e.g., Britt, 2000; Johnson, 2005; Johnson et al., 2008; Kautt, 2002) to address the “nesting” effect when the research seeks to include contextual variables that pertain to a community (or in this case a circuit) in addition to case-level variables.

Despite these risks single level modelling was considered to be preferable. A single level model was chosen because the immediate focus was on the degree to which

similar offenders were sentenced to similar sentences at circuit level. The focus was on the case level variables that attached to either the offending or the offender being sentenced, and the circuit the sentencing occurred in. In the first stage the modelling did not incorporate any contextual measures applying to or describing the circuits. The model was limited to measuring the impact of the circuit on sentencing having controlled for case level variables, and measuring the relative influence the variables included in the models have on sentencing decisions. The single level models had the advantage of producing a more detailed view of the differences between the circuits. The failure to account for judicial preferences for more severe or more lenient sentences in single level models was not relevant to the immediate examination of sentencing consistency between circuits. These propositions do not of themselves negate the risk of underestimation of standard errors and overestimation of statistical significance. However, because the second stage of the modelling was a multi-level approach it was possible to proceed on the basis that the potential under-estimation of standard errors and over-estimation of statistical significance could be tested by comparing the estimates reported for the single level models against multi-level random intercept models utilising the same data and variables. The results of the comparison are reported in Appendix B. No problems were found and the single level models were retained as the preferred means of assessing and depicting the consistency of sentencing between circuits.

In retrospect the single level models had another advantage over multi-level models. The scope of the variables able to be included in the single level models was substantially greater than the multi-level models. The variables used in the single level models had to be simplified in four of the multi-level models to enable the models to converge.

Stage two: multi-level modelling

Stage two aimed to increase understanding of inconsistency, assuming inconsistency was found after controlling for case level factors and circuit at stage one. Stage two involved two different modelling approaches using HLM7 software.

Multi-level random co-efficient models comprised of case level variables were used to address Predictions 3 and 4. A hierarchical generalised linear model was used to address Prediction 3 by identifying factors whose influence on the likelihood of incarceration varied between circuits. A hierarchical linear model was used to address Prediction 4 by identifying factors whose influence on the length of incarceration varied

between circuits. The sentencing variables found to be significant in the single level models formed the basis for these models. All these variables could be expected to have significant fixed effects but would not necessarily have random effects.

The random co-efficient multi-level models were extended to include contextual variables to explore the degree to which there was any association between sentencing per circuit, after controlling for case level factors, and any circuit level contextual factors (Research Questions 1 & 2). The multi-level models incorporating circuit level variables were severely limited due to the small number of circuits. The rule of thumb that a model can only support one higher level variable for every 10 units, limited the analysis to one circuit level variable at a time. As a result the modelling was limited to a series of bivariate models examining the influence of circuit level contextual variables.

Assembling and Interpreting the Models

The single level logistic and linear regression models were assembled on an iterative basis using the available predictor variables that sentencing research, sentencing texts, appellate authority, the draft sentencing guidelines and the sentencing notes suggested were relevant. The assumption of linearity between each of the continuous predictor variable included in the logistic regression models and the log (logit) of the outcome variable was tested, and all predictor variables were assessed for any evidence of multi-collinearity.

The assumption of linearity was addressed by determining if the interaction term between the continuous predictor variable and its log transformation was significant; a significant interaction term indicated that the assumption of linearity of the logit was violated for the continuous predictor variable (Field, 2009). The assumption was breached by a number of continuous variables. The affected variables were re-configured as categorical variables for the logistic regression models.

Multi-collinearity occurs when two or more predictors are highly correlated with one another. If they are, the estimates of the regression co-efficients become unreliable and it is difficult to assess the importance of each predictor (Field, 2009). The presence of multi-collinearity can be detected by examining the variance inflation factors (Chatterjee & Hadi, 2006). The variance inflation factor measures the degree to which the regression co-efficient of a predictor variable is increased due to its linear association with other predictor variables, with values greater than 10 being taken as an indication of multi-collinearity (Chatterjee & Hadi, 2006). The examination of the variables did not reveal any problems.

Once the suitability of the variables was confirmed the modelling commenced. The initial model was formed from the available predictor variables that sentencing research, sentencing texts, appellate authority, the draft sentencing guidelines and the sentencing notes suggested were relevant, plus the demographic variables and the circuit variable.

The case processing variables were incorporated into the offence seriousness category because the case processing variables were at least partial proxies for seriousness. The time from charge to charge outcome can be an indication of remorse. A shorter time suggests the early entry of a guilty plea, and early plea is often interpreted as the offender taking responsibility for their actions and being remorseful. Conversely a later plea or a finding of guilt at trial is interpreted as a failure to take responsibility and a lack of remorse. The charge laid type (whether the charge was laid summarily or indictably) was a proxy indicator of seriousness – summarily laid charges are generally less serious.

Each model was progressively refined by excluding non-significant variables and by simplifying some variables. Exclusion of non-significant variables helped to ensure the models were as parsimonious as possible. This was particularly important given the wide range of variables available. The exclusion of non-significant variables was not followed for the demographic variables because of the wider interest some readers may have in the specific results for these variables. Statistical non-significance does not mean that the excluded variables do not exert any influence on sentencing, they may be important on a case by case basis. Non-significance indicates that at an aggregate level the variable did not play a significant role in the determination of sentence.

For a range of variables, especially the criminal history variables found to breach the assumption of linearity, there was a need to find the best expression of the variable via repeated running of models with variations in the definition of the variable. The sources used to identify relevant variables provided a general description but not a detailed specification of the variables. This is a critical difference distinguishing this research from much of the prior research. For instance in prior research criminal history is often reduced to a simple dichotomous variable splitting no prior offending from some prior offending. For this research the data enabled criminal history to be expressed in polychotomous categorisations. The variables could be expressed in more or less detail, for example the criminal history of a burglar could be split into separate variables for burglary, receiving and other serious dishonesty. Alternatively the three variables could be combined into a single serious dishonesty variable. In addition the age of prior convictions could be taken into account. For example, the age of aggravated drink driving convictions was

represented by three dummy variables representing different counts of convictions in the previous ten years against a reference category of no convictions in the previous ten years. Elapsed time since the last conviction for aggravated drink driving was also represented with two dummy variables (a last conviction within 18 months of the current conviction, and over 18 months but within 5 years) against a reference category of a last conviction over 5 years ago.

The effects on parameter estimates of excluding some variables and the trialling of different specifications for some variables was monitored by comparing the parameter estimates between models to ensure there was limited change to other variables in the models.

The single level models cannot be easily categorised as either explanatory or exploratory. The selection of the variables based on sentencing research, sentencing texts, appellate authority, the draft sentencing guidelines and sentencing notes supports a description of the models as explanatory. However, the trialling of some variables, especially those capable of a range of polychotomous definition, supports a description of the models as exploratory. As these variables were based on generic descriptions of the variables and required specific definition to enable effective modelling the models are closer to the explanatory end of the spectrum but nonetheless the models are best described as a hybrid form.

The data could be conceived of as a population but that is misleading. While the models are based on almost all sentencing by resident judges in their home circuits for a given time span they are a random subset of the cases the judges could have been called on to sentence. There is no systematic basis for the allocation of cases between resident and visiting judges.

Once each model was refined a confirmatory check was made to ensure there was no multi-collinearity and residual statistics were reviewed to identify influential cases and potential outliers. To ensure that there were no individual cases exerting an undue influence on the model as a whole the Cook's distance statistics were reviewed to ascertain that all values were below the threshold level of 1 (Field, 2009). The standardised DFBeta measures were scrutinised to determine if there were any cases that substantially influenced the estimate of any of the model parameters. Standardised residuals were reviewed to identify problems with potential outliers. As a first step, counts of the number of cases with absolute values more than 1 and 2 standard deviations from the mean were taken to test whether the numbers were within acceptable limits. Secondly, all cases with a

standardised residual with an absolute value in excess of 3 were investigated as potential outliers (Field, 2009). The results of these checks are reported in conjunction with each model.

The choice of the reference circuit

The selection of the reference circuit for logistic regression or the circuit to be omitted from the creation of dummy variables for linear regression was a significant decision. The same circuit was chosen as the reference circuit across all three offences and for both decisions to facilitate comparisons of the parameter estimates and relative risks for each circuit between the models. Three parameters guided the choice. A circuit with a higher volume of offenders sentenced was preferred. Second, any circuit with a known or suspected distortion or truncation of caseload was discounted. This meant that two circuits that did not host jury trials were ruled out. Third, for ease of presentation and discussion of the results from the logistic regression analyses, a circuit from the lower end of the spectrum of incarceration rates was chosen. This meant that it was likely that for the most part any significant differences would be in the same direction and more easily presented. Ultimately the choice was relatively simple, as only one larger volume circuit was toward the lower end of the incarceration rate spectrum for all three offences. The choice of the reference circuit affects the parameter estimates for the circuits and the representation of inconsistency but it does not make a substantive difference to the results.

No adjustment for sample selection bias

Sample selection bias arises “when a researcher is limited to information on a non-random sub-sample of the population of interest” (Bushway, Johnson & Slocum, 2007, p.152). It has become increasingly common for sentencing researchers to include a corrective factor for sample selection bias, particularly in models of sentence length.

The example is pertinent to much sentencing research and in particular to this research because it models sentencing as a two stage process. Determination of sentence type (typically non-custodial versus custodial) and the quantum of the sentence (typically custodial or incarceration sentence lengths) are often modelled separately. Many researchers have accepted that the offenders who are imprisoned or incarcerated are a non-random subset of convicted offenders and that there is a risk of biased inferences regarding sentence length (see Bushway et al., 2007 for a general review and specific examples in Johnson, 2006, Kautt & Delone, 2006; Ostrom et al., 2004; Ulmer & Kramer, 1996).

The common solution to the non-random or biased selection of a sample has been to include a correction factor estimating the probability of imprisonment/incarceration in sentence length models, but not every researcher has done so. Some researchers (see for example Iles et al., 2011; Kautt, 2002) recognised the threat but made no correction as the risk of bias was limited because an overwhelming proportion of offenders were incarcerated. Others assembled models with and without correction factors and compared the results, and if the models were similar proceeded with the uncorrected model (e.g., Fearn, 2003; Kautt & Mueller-Johnson, 2009). In these studies the rationale for choosing the uncorrected model was the risk posed by collinearity between the correction factor and other variables in the model. Collinearity arises because the correction factor incorporates the other variables in the sentence length model that are influential in the choice of sentence type.

For the purposes of this research the choice to not correct for sample selection bias was informed by an examination of how non-random selection occurs in the data (Stolzenberg & Relles, 1997). In this case it results from the decisions of judges in some circuits systematically choosing to not incarcerate offenders who would be incarcerated in other circuits. Consequently the model is potentially missing the sentence length values for those cases that should but did not result in incarceration. There is no way to know for sure what sentence lengths would be imposed if these cases were to result in incarceration (Stolzenberg & Relles, 1997). The primary concern is that the mean sentence lengths for circuits with higher rates of incarceration are understated because the less serious instances of offending resulting in incarceration are associated with shorter terms of incarceration, possibly suggesting that there is no significant difference when there is a real difference. However, a review of the incarceration rates and sentence lengths indicates that was not the case, circuits with higher rates of incarceration also tended to impose longer terms for each of the three offences. In this circumstance any correction would increase the degree of difference between circuits. The choice was made not to add a correction factor to the models and recognise that the degree of difference revealed could be considered to be a conservative view of the difference. This approach also avoids any suggestion that the models represent a “cooking of the books”.

Reporting the results from logistic regression: Odds ratios and relative risk

The odds ratio and the 95% confidence interval for the odds ratio are among the most important statistics for interpreting the results. However, the odds ratio can be difficult to comprehend and is sometimes confused with relative risk. The following fictitious example using female and male offenders seeks to make the meaning of the terms clear and describe the difference between them.

The odds ratio is an estimate of the odds of incarceration for one category of offenders (in this example, females) compared to a reference category (males). If females were incarcerated at a rate of 9 per 10 offenders sentenced the odds would be 9:1. If males were incarcerated at a rate of 3 per 4 offenders sentenced the odds would be 3:1. The odds ratio is the odds for females divided by the odds for males, in this example the odds ratio:

$$(OR) = (9 / 1) / (3 / 1) = 9 / 3 = 3.$$

The odds of incarceration for females are three times greater than the odds for males. This *does not* mean females are three times more likely to be incarcerated. The number of times more likely is a different measure known as relative risk. Relative risk is the probability of occurrence for females divided by the probability of occurrence for males, in this example the relative risk:

$$(RR) = (9 / 10) / (3 / 4) = .9 / .75 = 1.2.$$

For small effects the odds ratio approximates the relative risk. When the odds ratio is applied in a setting with larger effects, as can be the case with sentencing, it is misleading to describe it in terms of the likelihood of occurrence. The odds ratio increases more rapidly than the relative risk as the probability of an event increases (Localio, Margolis & Berlin, 2007). Many readers may prefer to be able to view the comparisons in terms of relative risk rather than how many times greater the odds are. Relative risk is a more intuitively understandable measure. To that end the odds ratio was converted to a measure of relative risk using the following conversion formula:

$$RR = OR / ([1 - p_0] + [p_0 * OR])$$

p_0 is an estimate of the unadjusted risk for the reference group (Localio et al., 2007). The relative risk calculated via the formula cannot fully replace the odds ratio because the confidence intervals cannot be reliably estimated (Localio et al., 2007).

The calculation of the relative risk facilitates a more ready comparison of the impact of the controls applied in the model and the materiality of the differences between circuits. The relative risk after application of the controls can be compared to the raw or unadjusted relative risk. It can be used to indicate how many more or fewer offenders per

100 are likely to be incarcerated in a given circuit relative to the reference circuit assuming an incarceration rate for one of the circuits.

Incidental insights: the influence of offender gender and ethnicity

Consistent with the findings of the empirical research supporting the theories it is expected that sentencing approaches in New Zealand will have a common core based on offence seriousness and offenders' criminal histories. It is evident from the research elsewhere that variables for gender and ethnicity should be incorporated into this research. Gender and ethnicity have often been the focus of research on sentencing disparity. Consequently the particular results for these variables will of interest in their own right to readers of this thesis. Because of this interest these variables have been included in all models irrespective of their degree of influence. It is important that those interested in the influence of gender and ethnicity note that this research is limited to sentencing by resident judges, and excludes the 15-20% of sentencing undertaken by visiting judges. The exclusion of sentencing by visiting judges matters because it is known that there are systematic differences between circuits in the level of support provided by visiting judges.

Interpreting the results

The models are based on relatively large numbers of sentencing observations. One consequence of this is that small effects are likely to achieve statistical significance. It is important not just to focus on the evidence of difference but also to examine the substantive effect of the difference. A difference of a few per cent in the likelihood of incarceration or a small span of time in sentence length may be statistically significant but of minimal real impact system wide even though there may be a significant impact on individuals who receive a more or less favourable sentence as a result. If the differences are not substantive then the existence of a statistically significant but substantively small difference might be considered to be a satisfactory "price" to pay for the benefits of an individualised sentencing system with substantial judicial discretion. Therefore the interpretation of the models takes account of both statistical and substantive significance.

Chapter 5: Scope, Data Sources & Variables

To create the models described in the previous chapter, decisions were required on the scope of the models, the data that was required, and the structure and detail of the dependent and independent variables. This chapter works through the determination of the scope of the models and the data and variables for the models.

The Scope of the Sentencing Models

The scope of the sentencing models was determined by addressing the alternative settings available for a range of parameters. The following questions were addressed. Which courts should be included? What time span should the research cover? Which offences should be included? Should the offences that are included be addressed collectively or individually? At a more detailed level, decisions were required on the unit of analysis, the treatment of cumulative sentences, whether to include or exclude sentencing by visiting judges, how to deal with changes of sentence, and the inclusion or exclusion of cases depending on whether one of the chosen offences was the most serious offence in the case.

The choice of courts

New Zealand has a unitary court system headed by the Supreme Court accompanied by two other superior courts, the Court of Appeal and the High Court, and a series of inferior courts with limited statutory jurisdiction (Joseph, 2007). It is significant for sentencing research that New Zealand has a unitary system. It is free of the difficulties that can arise in federal systems with separate laws and court structures and varying jurisdictional boundaries between states and between state and federal systems that complicate and potentially confound research.

The sentencing of adult offenders in New Zealand takes place in the District Courts and the High Court. The High Court has exclusive first instance jurisdiction over the most serious offences, most notably murder, attempted murder and manslaughter (Brookbanks, 2007). Conversely, the District Courts have exclusive jurisdiction over the least serious cases. The High and District Courts have shared jurisdiction for the more serious cases spanning the gap between their respective exclusive jurisdictions. The District Courts dominate sentencing, with almost 99% of adult offenders sentenced in the District Courts. The 99% included a majority of the more serious cases where there was shared jurisdiction between the High and District Courts. For instance, in the three years 2004-2006, 78% of

offenders sentenced for sexual violation (rape), 94% of offenders sentenced for aggravated robbery, and 95% of offenders sentenced for the most serious physical assaults were sentenced in the District Courts (Sentencing Establishment Unit, 2008). These figures exemplify Joseph's (2007) description of the District Courts as the workhorses of the judicial system.

The dominance of the District Courts makes it the obvious choice for this research. Inclusion of High Court sentencing would confuse the analysis. The High Court is organised on a different geographical basis, with 18 courts across 3 circuits compared to 63 courts across 18 circuits in the District Courts. The sentencing experiences of High Court judges are very different from those of District Court judges. The High Court judges are exposed to a much narrower range of offences, with most of the offending at the high end of the seriousness spectrum and are only exposed to less serious matters in their appellate roles hearing appeals from the District Courts either in the High Court or as part of a divisional court in the Court of Appeal.

The offences

In determining the scope of offending to be covered the first consideration was whether to cover a range of offences in the same model or to have separate models for a selection of offences. It is common for studies to cover a broad range of offence types within a single model and to control for differences by the inclusion of offence type as a dummy variable (Crow & Gertz, 2008; Fearn, 2005; Johnson et al., 2008; Sharp, 2007; and Ulmer, 2005 are examples of studies covering a range of offence types). Doing so inevitably imposes limitations on the range of other variables that can be taken into account. This can be limiting if the relevant sentencing factors can vary widely between offences. In the United States the ability to use guideline scores may overcome this limitation to a substantial degree. There is no equivalent mechanism available in New Zealand that would enable a more broadly based study to be undertaken. Despite the availability of guideline scores some United States researchers chose to limit their studies to a single offence (Iles et al., 2011; Kautt, 2002; Pasko, 2002). Kautt (2002) expressed concern about the reliability issues that arise as a consequence of the need to include a greater number of independent variables to cover a range of offence types and chose to focus on drug trafficking offences.

In New Zealand it would be difficult to draw conclusions from or to defend a study that analysed a range of diverse offence types without developing a generic seriousness scale. The variation in the relevant sentencing factors between dishonesty, violent, drug,

sexual, and traffic offending is too great to be resolved by including each type as a dummy variable. The preference here is for a multi-offence study with separate models for each offence. The selection of offences from different subsets of general offending offered the possibility of gaining insight into the degree to which any effects are pervasive or whether they vary by offence.

Burglary, male assaults female, and aggravated drink driving were chosen as the best offences to include in this study because they had the best fit against the following considerations. The total number of offenders sentenced needed to be sufficient to support effective modelling. The rate of use of incarceration had to be sufficient to result in a large enough number of incarcerated offenders to support effective modelling of sentence length. The rate of use of incarceration could not be too high because it would otherwise preclude effective modelling of the decision to incarcerate. The choice of offences needed to avoid any peculiarities affecting the sentencing of the selected offences that could confound the development of effective models. The chosen offences should ideally represent different categories within the population of offence types.

The obvious omissions from study were any form of drug offending or sexual offending. With the exception of dealing in Class C drugs (mainly cases involving cannabis), none of the drug or sexual offences was sentenced in sufficient volume to support modelling. In addition a feature of the sentencing of sexual offences would, in any case, have made the inclusion of a sexual offence less desirable. Offenders convicted of serious sexual offences may be sentenced to preventive detention (Sentencing Act, 2002). Preventive detention is an indeterminate sentence designed to protect the community from the offender. Once offenders have served any minimum term prescribed by the judge they may be released but are subject to lifetime recall. This sentence may only be imposed by a High Court Judge. If a District Court Judge considers that the sentence may be appropriate for an offender they transfer the case to the High Court for sentencing (Robertson, 2007). The usefulness of these offences in a study of consistency within the District Courts is therefore reduced. It may result in varying proportions of the more serious instances of offenders appearing for sentence in the various circuits being excluded from the analysis.

While the Class C dealing volume was sufficient, the offence was rejected because of severe limitations on the term of imprisonment available to District Court judges if the charge was laid summarily or if the offender pleaded guilty at an early stage to an indictably laid charge. In these circumstances The Misuse of Drugs Act 1975 only allowed District Court judges sentencing offenders convicted of dealing in Class C drugs to

sentence up to a maximum of one year's imprisonment compared to the legislated maximum of eight years. If the judge was sentencing an offender convicted of cultivating prohibited plants they were restricted to a maximum of two years compared to the legislated maximum of seven years. In any instance where limits restricted the sentence available to a District Court judge, the judge could decline jurisdiction and transfer the case to the High Court where the only limit was the statutory maximum penalty (Robertson, 2007).

The numbers of offenders sentenced, incarcerated and the incarceration rates for the offences considered are set out in Tables 5.1 (property/dishonesty and drugs), 5.2 (violent and sex offences) and 5.3 (serious driving).

Table 5.1: Mean annual sentencing volumes for selected property/dishonesty and drug offences 2004-2006

Offence	Number sentenced	Number incarcerated	Incarceration rate
Burglary	3205	1411	0.44
Deception	2600	456	0.18
Unlawful taking of a motor vehicle	948	281	0.30
Theft (incl. shoplifting)	6940	437	0.06
Theft (excl. shoplifting)	3093	280	0.09
Shoplifting	3847	157	0.04
Receiving	1211	201	0.17
Arson	142	73	0.51
Intentional/wilful damage	3022	45	0.01
Class C Drugs - dealing offences*	435	206	0.47
Class B Drugs - dealing offences*	314	162	0.52

* Dealing includes importation, manufacture, sale, supply and possession for supply.

Source: The table was compiled from unit data obtained from the NZ Law Commission under the Official Information Act. The Law Commission originally obtained the data from the Ministry of Justice who extracted it from their administrative systems.

The higher number and rate of burglars incarcerated distinguished burglary from the other offences that could have been chosen. A potential problem for sentence length research addressing burglary is a five-year sentence limit on judges without a jury warrant sentencing a burglar charged indictably. In theory this could result in a bulge in sentences

at or near the five year point. However, this was purely a nominal issue because all 8 of the offenders charged indictably and sentenced to more than 4 ½ years and up to 5 years were sentenced by a jury warranted judge. Burglary was not subject to any other known idiosyncrasies that could confound the analysis.

Table 5.2: Mean annual sentencing volumes for selected violent and sexual offences 2004 - 2006

Offence	Number sentenced	Number incarcerated	Incarceration rate
Male assaults female	2709	377	0.14
Robbery	504	397	0.79
Grievous assaults	236	207	0.88
Injuring with intent to injure	301	189	0.63
Breach Domestic Violence Act	1086	171	0.16
Assault with a weapon	610	164	0.27
Assault with intent to injure	436	157	0.36
Threaten to kill or do GBH	427	102	0.24
Common assault (Crimes Act)*	2457	177	0.07
Common assault (Summary Offences Act)*	2111	46	0.02
Assault on a child	217	27	0.12
Sexual violation	174	168	0.97
Indecent assault or act (child or young person)	172	114	0.66
Indecent assault (adult)	108	40	0.37

* The two forms of common assault are differentiated by penalty; the Crimes Act has a maximum penalty of 12 months whereas the Summary Offences Act is 6 months.

Source: The table was compiled from unit data obtained from the NZ Law Commission under the Official Information Act. The Law Commission originally obtained the data from the Ministry of Justice who extracted it from their administrative systems.

More offenders were sentenced for male assaults female than any other sexual or violent offence and male assaults female was second to robbery for the number of offenders incarcerated. For research purposes male assaults female had two advantages over robbery. First, robbery was a composite set of offences with multiple maximum penalties (7, 10 and 14 years) whereas male assaults female was a singular offence. Secondly, the total volume for male assaults female was substantially larger, resulting in

much higher volumes per circuit. Only three circuits sentenced more than 50 robbery offenders per annum, whereas only one circuit sentenced fewer than 50 male assaults female offenders per annum. The other high volume offences such as the two forms of common assault had much lower numbers incarcerated (177 and 46 respectively compared to 377 for male assaults female).

Like all violent offences male assaults female poses a challenge because it sits in a hierarchy of overlapping offences. A less serious instance of male assaults female could be charged as common assault, whereas a more serious instance could be charged as assault with intent to injure. Variations in charging practices may contribute to differences in the mix of cases between circuits. However, any disadvantage created by the overlapping charging structure was considered to be outweighed by the usefulness of including a violent offence alongside the other chosen offence types.

The sentencing outcomes for male assaults female could be influenced by the offender being sentenced in a Family Violence Court. To control for sentencing in family violence courts a variable denoting court type was included in the modelling.

Table 5.3: Mean annual sentencing volumes for selected serious driving offences 2004-2006

Offence	Number sentenced	Number incarcerated	Incarceration rate
Adult 3rd or subsequent drink driving	3939	940	0.24
3rd or subsequent disqualified driving	2347	756	0.32
Serious driving causing death or injury	253	120	0.47
Careless driving causing death or injury	836	7	0.01

Source: The table was compiled from unit data obtained from the NZ Law Commission under the Official Information Act. The Law Commission originally obtained the data from the Ministry of Justice who extracted it from their administrative systems.

Both 3rd or subsequent drink driving (aggravated drink driving) and 3rd or subsequent disqualified driving were viable candidates for inclusion. Drink driving was chosen because the volumes were larger.

The time span

Three considerations informed the choice of the time period for the study. The period needed to be recent to avoid any concern that the results could be described as true of the past but not relevant now because of changes to the legislation or appellate authorities. The period needed to be long enough to result in sufficiently large numbers of offenders sentenced. The law and system settings needed to be stable through the time period.

The two years 2008-2009 were relatively recent. There had not been any change to the law or other settings that had the potential to change sentencing outcomes. The law and other system settings were stable through 2008-2009.

No guideline judgments were issued in relation to any of the three offences during the period. The only potentially significant change that occurred was the issuing of a short-lived generic guideline judgment on sentencing discounts for guilty pleas (*R v Hessel*, 2009) issued by the Court of Appeal in October 2009. In the absence of any studies of the impact of the judgment it is impossible to be definitive about its effect on sentences. However, there are two reasons for assuming it would have had little impact on sentencing outcomes in relation to this study. First, it only applied to one-eighth of the timeframe. Secondly, it was not expected to result in significant change. The Court did not believe that the guideline differed significantly from the practices being followed by many judges (*R v Hessel*, 2009), although it may have resulted in a reduction in the proportion of cases receiving the maximum discount because some judges may have altered their sentencing approaches in response to the Court taking a stricter approach to the definition of an early guilty plea. The guideline was replaced in 2010 by a revised guideline issued by the Supreme Court; the revised guideline (*Hessel v R*, 2010) did not apply to sentencing in 2008-2009.

The unit of analysis

The unit of analysis used in this research is the case. For the purposes of this research a case includes a single offender sentenced on one or more charges on the same day at the same court by the same judge. This concept of case was the only realistic choice because in New Zealand sentencing is governed by a totality principle requiring the judge to tailor the total sentence to fit the totality of the offending. The sentence on any single offence when there are multiple convictions may therefore be affected by the sentences for the other offences.

Concurrent and cumulative sentences

Defining the term of incarceration was complicated by the ability of the judge to impose concurrent sentences, cumulative sentences or a mix of both when sentencing on multiple charges. If all the sentences were concurrent the effective sentence was the longest concurrent sentence. If some or all of the sentences were cumulative the effective sentence was the longest cumulative sentence.

It was important to include both concurrent and cumulative sentences in the modelling of sentence length; exclusion of cumulative sentencing cases would have been likely to skew the results. Any skew would be exacerbated if there was variation in judicial practice in the use of concurrent and cumulative sentences between the circuits. The rate of use of cumulative sentences was substantial with 17.0% of 1,592 aggravated drink drivers, 20.8% of 2,048 burglars, and 36.8% of 558 male assault female offenders incarcerated on multiple offences receiving cumulative sentences. Importantly, there was also a substantial difference in the term of the effective sentence between the two forms. The average cumulative sentence for aggravated drink drivers was 13.6 months compared to 8.7 months for concurrent sentences; burglars 18.6 months compared to 15.3 months; and for male assaults female 10.1 months compared to 6.6 months. If cumulative sentences had not been included the datasets would have been substantially smaller, with the sentence range more limited and unduly weighted toward shorter sentences.

Categorisation of judges and courts and the implications for this research

The sentencing in any circuit was undertaken by a combination of resident judges (judges whose common room was in the circuit) and visiting judges from other circuits. With one significant exception sentencing in any circuit was primarily undertaken by judges who were resident in the circuit. The exception was Gisborne where there were no resident judges and all sentencing was by visiting judges. Excluding Gisborne, visiting judges undertook a minor but material proportion of sentencing; 17.1% of aggravated drink driving, 16.7% of burglary, and 13.8% of male assaults female sentences were imposed by visiting judges. The proportion of sentencing by visiting judges varied substantially between circuits; ranging from 6.4% to 47.5% for aggravated drink driving, from 5.4% to 45% for burglary, and from 1.8% to 49.7% for male assaults female. The inclusion of sentencing by visiting judges could have obscured the real sentencing outcomes attributable to the circuit and its resident judges. Sentencing by visiting judges was, therefore, excluded from the analysis, meaning that the Gisborne circuit was excluded.

The resident judges involved in sentencing in each circuit can be sorted into two categories, Family Court judges and criminal/general judges. The Family Court is a division of the District Courts but only judges with a specialist Family Court warrant can sit in the Family Court. Family Court judges generally do much less criminal work and sentencing than their colleagues. Anecdotally it is understood that Family Court judges spend about 75% of their time in the Family Court and 25% in other District Court work (this understanding was confirmed with staff at the Chief District Court Judge's Chambers). The sentences imposed by the two categories of judge could vary due to their different roles and experiences. To allow for this possibility judge type was included in the list of variables to be included in the modelling process.

In some circuits the judges have opted to run specialist Family Violence Courts. In 2008-2009 six circuits hosted specialist Family Violence Courts for all or part of the period (Ministry of Justice, 2010). The Family Violence Courts bring together like cases to be heard in the same session, with a view to remedying the problems giving rise to family violence while holding offenders to account (Ministry of Justice, 2010). The courts follow a 'therapeutic justice' model seeking speedy disposition to achieve accountability, safety and family repair. Offenders convicted of male assaults female offences may have been sentenced in these courts. The Family Violence Courts may have had a systematic influence on sentencing outcomes and consequently a variable distinguishing those courts hosting Family Violence Courts from those that do not was included in the list of variables to be included in the modelling process.

Changes of sentence – focus on the first or the final sentence

Some cases had a complex sentencing history with different sentences imposed on different dates due to re-sentencing or the outcome of an appeal. Re-sentencing can occur in several circumstances. The most common is a change because the offender breaches the original sentence and is re-sentenced. It can also occur when an offender is sentenced on new charges and the original sentence is incompatible with the new sentences; when an offender's circumstances have changed and the sentence is inappropriate; or when there is a successful appeal.

In this research only the sentences that were imposed on the earliest sentencing date were included. It was the sentence the judge considered was appropriate to the offending and the offender at the time. Re-sentencing is due to subsequent events unknown by the

original sentencing judge. Changes on appeal occur due to new information or because the appellate court found that the sentence was wrong in some aspect.

Primary and secondary offences

Having chosen the case as the unit of analysis consideration was required to be given to whether all cases that included a charge of aggravated drink driving, male assaults female or burglary should be included in the research. The cases could be divided into two groups, those where one of the three offences was the most serious or primary charge for sentence and those where some other charge was more serious.

The analysis was restricted to those cases where one of the three offences was the most serious offence. Inclusion of the cases where some other offence was more serious would have skewed the sentencing outcomes because the sentence on the more serious charge would have largely dictated the outcome. For instance, if an offender sentenced for aggravated robbery as the lead offence was also being sentenced for aggravated drink driving the judge's choices could be limited. If the aggravated robbery resulted in imprisonment the judge could either imprison or convict and discharge for the drink driving but could not impose a fine or any other intermediate sentence. In that circumstance the judge might impose a nominal sentence of imprisonment for the aggravated drink driving but that might not reflect what would have happened in the absence of the aggravated robbery. A series of tests primarily based on charge outcomes, sentence types and quantum, and the generic seriousness of the offence type were applied to ensure that cases with a burglary, male assaults female or aggravated drink driving charge as the most serious charge were retained in the final dataset for analysis. A more detailed summary of the tests is set out in Appendix C.

Summary of the scope

The models addressed the original sentencing by resident judges of offenders convicted for aggravated drink driving, burglary and male assaults female in the District Courts in 2008-2009 where that was the most serious offence if there were multiple convictions in the case. The unit of analysis was the case comprising all sentences imposed on a particular offender in the same court on the same day by the same judge. The models included both concurrent and cumulative sentences, with sentence length expressed as either the longest concurrent sentence or the longest chain of cumulative sentences.

The Data

The modelling made use of data from three sources. The logistic and linear regression models were created from case records assembled from administrative data maintained by the Ministry of Justice. The contextual circuit level variables for the hierarchical models were assembled from census data and from caseload and caseflow reports obtained from the Ministry of Justice. Judges' sentence notes (a transcript of their sentencing remarks) were expected to provide an additional source of information on sentencing factors for inclusion in the sentence length models. However, the dataset that was assembled was not representative of the circuits and could not be used in the final models. The effort in assembling and examining the data had some benefits; it informed the assessment of limitations of the current research. An expanded discussion of the sentence notes and the problems encountered can be found in Appendix D.

Data source one: Case level administrative data

The Ministry of Justice provided an electronic extract of raw data on all offenders sentenced for aggravated drink driving, male assaults female and burglary in 2008-2009. The data was drawn from the Justice Data Warehouse which is a comprehensive database of all persons charged with an offence and the disposition of the charges against them, including sentencing. The database takes a daily update of information from the Case Management System (CMS) managed by the Ministry of Justice. CMS is an administrative computer system supporting operational case management in the court registries and the courtroom. The conviction and sentencing data in CMS is subject to an auditing process providing assurance that the data entered is materially accurate.

The raw data provided by the Ministry of Justice was in three tables covering the offenders, the charges they had faced, and any sentences imposed on those charges. The Offenders Table included the sentencing date, offence type and core demographic data (sex, race and date of birth) on all offenders sentenced for any of the three offences in 2008-2009. The Charges Table contained all charges that those offenders had ever faced at any time and included the charge outcome type and date and other case processing data. The Sentences Table contained all sentences that had ever been imposed on any of the charges the offenders had been convicted of at any time, including the sentence type and amount and some case processing data.

Linking the tables

The tables of raw data were able to be linked together through a series of identifiers. The Ministry created a unique identifier for each individual offender specifically for this research, and attached it to each observation in all three tables. The tables of charges and sentences both included the “charge record number” (crn) - a unique identifier attached to each charge created in CMS when a charge is laid. The crn enabled each charge and the sentences imposed on the charge to be linked.

The raw extract included 36,636 offender observations for 26,117 individuals discharged without conviction, convicted and discharged or convicted and sentenced for one of the three offences. Some offenders were sentenced on multiple occasions over the two years. The Charges Table included 850,332 charges for the 26,117 offenders and the Sentences Table included 935,668 sentences imposed on those charges.

The data provided was at the lowest level of aggregation possible in order to facilitate its extraction by the Ministry of Justice, but it was not in a form readily susceptible to analysis. Before it could be used to model sentencing decisions, the raw data had to be transformed in accordance with the specification for the research.

Transforming the raw data into cases for sentencing

The transformation occurred in two parts. The first part was the refinement of the data to exclude some duplicated offender observations and to exclude any observations where an offence other than aggravated drink driving, male assaults female or burglary was the most serious offence. The original extract included duplicated offender observations when the same offender had been sentenced on the same day for multiple instances of one of the three offences and the offences had different offence codes. The duplicates were readily able to be recognised and deleted. After 8,021 duplicates were deleted, the number of potential cases was reduced to 28,615.

The next step was to determine whether the remaining aggravated drink driving, male assaults female and burglary observations were the most serious offences within the case. A detailed specification was drawn up with a prioritised series of tests employed to determine status. A summary of the tests is described in Appendix C. The application of the tests eliminated 2,297 observations because an offence other than aggravated drink driving, male assaults female or burglary was more serious, leaving 26,318 observations.

The second stage of the transformation was to expand each of the remaining observations into a single case record including the offender’s demographic profile, the

circuit, the judge, a summary of the case being sentenced, the sentence(s) on the case, and the offender's criminal history. When this process was completed a further tranche of observations/cases were removed, including: 4,877 cases sentenced by a visiting judge; 173 aggravated drink drive cases for offenders who did not have the two prior convictions necessary for the offence; 12 cases shown as having been sentenced in the High Court or Youth Court; 12 cases that did not have any sentences; and 56 cases that had sentences shown as cumulative on a sentence imposed on a separate case on an earlier date. The latter group of cases was excluded because it was impossible to determine whether the determination of sentence type or length was constrained by the sentencing of the earlier case. The finalised data was split into three datasets: aggravated drink driving (9,282 cases); male assaults female (6,166 cases); and burglary (5,740 cases).

Data source two: District Courts caseload, caseflow reports and sitting time reports

The second source of data was a set of reports on criminal summary and jury trial case volumes and judge sitting hours for the District Court circuits assembled by the District Courts Operating Unit of the Ministry of Justice.

Caseload and caseflow reports

The caseflow reports were detailed, with the information split into several categories based on how a case proceeded. There was separate information on undefended summary cases (where no plea was entered or the case was resolved by an initial plea of guilty); not guilty summary cases (where the defendant entered a not guilty plea but no defended hearing was set down); defended summary cases (where the defendant pleaded not guilty and a defended hearing date was set); and jury trial cases (where the defendant was committed for trial by jury).

The data was organised by court and included the balance of cases on hand at the beginning of the period, new cases filed, administrative actions that added to and deducted from the caseload, transfers of cases between courts, a range of case resolution categories and the balance of cases on hand at the end of the period. The data was able to be aggregated up to circuit level and was used to calculate or contribute to five contextual variables: the number of new cases per thousand adults; the number of new cases per judicial officer; the average active criminal cases per judicial officer; the number of cases disposed per judicial officer; and the proportion of cases disposed by a hearing of the charges.

Sitting time and judge strength

Judge strength in a circuit could have been measured by simply counting the number of judges in the circuit. However, this would have been a crude measure for three reasons. First, the judges did not just sit in the circuit they resided in, with judges also sitting as visiting judges in other circuits. The rate at which they did so was not uniform between the circuits. Second, the judges were not the only judicial officers to decide cases; a proportion of less serious cases were dealt with by justices of the peace or community magistrates. The proportionate contribution made by these judicial officers was not uniform across the circuits; particularly as community magistrates were only located in a small number of circuits. In Hamilton justices of the peace and community magistrates accounted for 35% of criminal summary sitting time, whereas in Manukau they accounted for only 6%. Third, the contribution of different classes of judge to the criminal jurisdiction varied. A simple count of judges takes in all judges irrespective of contribution; the contribution by judges with a Family Court warrant could vary between circuits.

The best available measures of effective criminal judge strength per circuit were the amount of criminal sitting time associated with all judges or all judicial officers. Sitting hours were converted to a full time equivalent count by dividing by two (to account for the two years) and then dividing by an estimate of expected sitting time per judge or per judicial officer per annum (768.6 hours). The estimate was formed after discussion with staff from the Chief District Court Judge's Chambers and assumed each judge was available to sit 183 days per annum and that they sat on average for 80% of the standard 5.25 sitting hours per day. The 183 days incorporated a standard deduction for planned and unplanned absences.

Data source three: 2006 census data

The third source was an extract of data sorted by District Court circuit from the 2006 census undertaken by Statistics New Zealand. It was made possible due to previous work undertaken by the Ministry of Justice to map the 40,954 census meshblocks to the 63 District Courts. Because each court had a one to one relationship with a circuit, it was a simple matter to link the meshblocks to the 18 circuits. The meshblocks are the smallest geographic units the census data is collected and processed in (each meshblock represents approximately 100 people).

The extract was limited to the adult population (those aged 17 or more). This definition of adult was adopted because it corresponded to the threshold between the youth and adult courts. There are some individuals under 17 who are sentenced in the adult courts but the number is very low. The data covered the sex, ethnicity, educational achievement, employment status and income of respondents. The data was used as the source of socio-demographic variables to be taken into account as part of the sentencing context at circuit level. It was also used in conjunction with various measures of caseload to create measures of offending and sentencing relative to population.

The Variables Modelled

The case level outcome variables

The substantive sentence is a combination of sentence type and quantum. Many researchers have focussed on these two core decisions and conceptualised sentencing as a two-step process with two outcome variables, one for the type of sentence and the other for quantum (Britt, 2000; Dixon, 1995; Kautt & Delone, 2006; Ulmer & Johnson, 2004). Research undertaken by Wheeler et al. (1982) is often cited as justification. Interviews conducted by Wheeler et al. (1982) conducted with judges revealed that the judges commonly decided first whether to imprison or not and then addressed the quantum. They also found that the factors predicting type and quantum may not be the same, indicating that they are separate decisions.

Bushway and Piehl (2001) challenged the use of the two-step process as an inadequate representation of sentencing under guidelines. They argued that the guidelines resolved the sentence type and that sentencing was a single step process with the judge only having to address quantum. The challenge is not relevant to New Zealand as the choice of sentence type is made by the judge. The idea that judges might be indifferent to sentence type is contrary to the statutory schema. Judges are required to “impose the least restrictive outcome that is appropriate in the circumstances” (section 8(g) of the Sentencing Act 2002). They must impose the least restrictive outcome in accordance with the hierarchy of sentence types set out in the Act. Restrictiveness is essentially a function of the nature of the sentence type rather than the quantum of the sentence.

Not all research has focussed just on type or quantum. Some researchers (Albonetti, 1997; Johnson, 2005; Johnson et al., 2008) addressing guideline sentencing systems have focussed on the decision to depart from the guideline. These decisions are

not relevant to this research because there are no similar departure decisions for the offences covered.

The sentence type decision

The dependent variable for the sentence type decision was specified as a dichotomous categorical variable with incarceration (imprisonment and home detention) coded 1 and non-incarceration (community detention, intensive supervision, community work, supervision, fines, reparation, orders to come up for sentence if called upon, and discharges) coded 0.

Most researchers have reduced the sentence type decision to a dichotomous choice between imprisonment or incarceration on one hand and non-custodial sentences on the other hand. It is commonly described as the “in-out decision”, reflecting whether or not the offender was placed in custody. The precise definition of the variables needs to be relevant to the jurisdiction being researched. In New Zealand, the key definitional issue is whether home detention should be treated as incarceration in the specification of the in-out decision. The decision could have been specified as the choice between imprisonment and any other sentence. Doing so would recognise that home detention is qualitatively different from imprisonment. The offender is restricted to a residential address with no physical barrier restraining them, but they are subject to electronic monitoring so that any movement beyond the address is picked up quickly.

There were three reasons for combining imprisonment and home detention into an incarceration grouping. First, home detention imposes substantial restrictions on the freedom of movement of the offender and in doing is qualitatively different to other non-imprisonment sentences and more similar to imprisonment. Second, under the law a judge should only impose home detention if he or she would otherwise have imposed a sentence of imprisonment of two years or less. Third, some offenders are imprisoned despite a judge preferring home detention. The distinction between home detention and imprisonment is blurred because for practical reasons home detention is not available in every case. It may not be available in the general location or for some specific residences due to cell phone coverage issues. In some cases the offender may not be able to offer a residence acceptable to the court or another resident may withhold consent for the offender to serve home detention at the address. Judges may have to impose imprisonment when they would prefer to impose home detention. There was no foolproof way in which such cases could be distinguished.

Two outcomes, discharges without conviction and conviction and discharge, which might not be thought of as sentences but which are sentence-like, have been included as non-custodial sentencing outcomes. Discharges are determined by a judge after a process that is similar to the sentencing process. The offender will have either pleaded or been found guilty. The judge is likely to have received submissions or heard argument from counsel and considers the offending and the offender in the same way as when imposing a substantive sentence. Both forms of discharge may result in the offender paying monies, undertaking some voluntary work or getting some help for a problem linked to their “offending” behaviour. These actions are akin to fines, reparation, community work or supervision sentences.

The sentence quantum decision

The length of incarceration sentences was adopted as the sentence quantum variable preferred for this research. The length of incarceration was specified as a continuous variable measured in months. It was preferred to the alternative of creating a single scale standardising all sentence types, via a series of adjustments to the different measures, for practical reasons. The measures used for each of the sentence types would need to have been reconciled and translated onto a single scale. The dollar value of fines, the months of supervision, the curfew term in community detention, the hours of community work and time in prison or on home detention would all need to have been equated. This would have been a somewhat arbitrary exercise, because the actual requirements of some sentence types can mean that two similar sounding sentences are quite different in detail. For example, a 12 month supervision sentence with regular reporting requirements is less onerous than such a sentence with the same reporting requirements but additional requirements to attend three programmes. Similarly, the specific hours of a community detention curfew can result in some offenders serving most of their sentence while asleep each night, but others might be detained across each weekend for the duration of the sentence. Thus, it was decided to focus on incarceration where there was little uncertainty in the measurement of the quantum.

Measuring the term of incarceration

Measurement of the term of incarceration required reconciliation of imprisonment and home detention sentences. The first decision required was to choose the most appropriate representation of imprisonment. In theory there was a choice between the nominal

sentence imposed by the judge and the actual time served. In New Zealand the effect of the nominal sentence depends on whether the offender is sentenced to a short sentence (defined as two years or less) or a long sentence (defined as a sentence of more than two years). Offenders serving short sentences (two years or less) must be released at half sentence and cannot be recalled. Offenders serving long sentences are subject to the possibility of release on parole at a minimum of one-third time served, but may be required to serve the full sentence. Release on parole is determined by the Parole Board based on the risk the offender poses to the community. An offender released on parole may be recalled. Actual time served on long sentences can vary from one-third to the full term on a case by case basis. The actual time served by those sentenced to a long sentence is determined by the Parole Board and is not a sentencing decision.

For the purposes of this research the best measure is the nominal sentence imposed by the judge. In New Zealand it is not the judge's responsibility to determine the time the offender should serve. The original sentencing decision is based on the seriousness of the offence and culpability of the offender for that offending. The nominal sentence reflects the quantum the judge considers appropriate.

An adjustment to home detention sentences was required to equate home detention with imprisonment. Home detention sentences are only available if the judge has decided that a short term of imprisonment would otherwise be imposed (Sentencing Act, 2002). However, home detention is not specified as a term of two years or less with the offender released having served half the term. Instead the term can be for any length from a minimum of three weeks to a maximum of twelve months with the offender serving the full term. To equate home detention to the nominal term of a short sentence of imprisonment the home detention sentence was doubled.

The decision to double the term of home detention was tested by capturing examples of sentencing decisions where the judge's sentencing remarks included reference to the term of imprisonment that would have been imposed and the term of home detention actually imposed. Many decisions included an acknowledgement by the judge that imprisonment was warranted, but they sentenced without any explicit reference to what the corresponding term of imprisonment would have been. Sixty-seven decisions were found that included the conversion mathematics. Two of the 67 were excluded as extreme outliers. In one case the judge made no adjustment and appeared not to realise that the rules were different for the two sentences. In the second case the judge acknowledged and quantified the appropriate imprisonment term but then struck the home detention term as

the time required to complete a residential drug treatment programme. In 43 of the remaining 65 cases the ratio of home detention to half the posited nominal term of imprisonment was 1:1. With the outliers excluded the mean of the ratio for the 65 cases was 1.01 (the mean term of home detention was 1% higher than the nominated term of imprisonment) with a standard deviation of 0.14. The result confirmed that the decision to double the term of home detention was reasonable.

The case level predictor variables

The predictor or independent variables cover a generic set of demographic variables, offence specific sets of offence related seriousness and case processing variables and criminal history variables. The determination of the variables was informed by the research literature, New Zealand reference texts on sentencing, appellate decisions, the draft sentencing guidelines and sentencing notes.

The variables used to develop the models are described in a series of tables in Appendix E. The description in the tables includes the coding for each variable and the expected effect of the variable on sentencing. The coding distinguishes between continuous and categorical variables. Some variables are listed as both continuous and with a categorical alternative. This was necessary because some of the continuous variables breached the assumption of linearity necessary for inclusion in regression models. The expected effects are primarily described in terms of whether the variable was expected to aggravate or mitigate the severity of the sentence but in some cases the effect is described as uncertain. Some variables were subsequently simplified for inclusion in the multi-level models. This was necessary because of the added complexity of those models. The changes made are detailed in Chapter 8.

The range of offender demographic variables was limited to the offender's age, gender and ethnicity. Information on employment status, occupation, wealth, marital status, family responsibilities, health status, education, or religious belief variables that have sometimes been included in research elsewhere was not available.

Circuit level contextual variables for multi-level modelling

Researchers have used a broad range of contextual measures at the circuit level, including population demographics, urbanisation, income, education, political preferences, employment levels, crime rates, caseloads and trial rates (Austin, 1981a; Britt, 2000; Dixon, 1995; Eisenstein et al., 1988; Johnson, 2005; Johnson, Ulmer & Kramer, 2008;

Kautt, 2002; Myers & Talarico, 1987; Steffensmeier et al., 1998). The factors modelled were determined from the research literature and shaped and limited by the available data.

Demographic and socio-economic contextual variables

Five demographic and socio-economic contextual variables from the census were used in the modelling. The five variables were: the adult population; median personal income; racial homogeneity; the proportion of adults with no educational qualifications; and the proportion of adults unemployed. The separation in time between the census data (2006) and the caseload and sentencing data (2008-2009) might have resulted in some degree of mismatch. This is unlikely to be significant in relation to population, income, ethnic homogeneity or educational levels. These measures were considered to be unlikely to have changed in such a way that the relationships between the circuits would be materially different. Any changes to unemployment rates might have varied between circuits, hence any findings in relation to unemployment need to be treated with greater caution.

Four of the five variables are self-explanatory, the exception being ethnic homogeneity. It is a variable based on the concept of ethnic heterogeneity introduced by Britt (2000). Ethnic heterogeneity was calculated as 1 minus the sum of the squares of the proportions each ethnic group was of the total population. Thus a population made up of a single ethnic group would score zero ($1 - 1$), and would be perfectly homogeneous. The higher the score the less ethnically homogeneous and more mixed or heterogeneous the circuit population is. The highest possible score occurs when all ethnic groups are of equal size. In the current research the highest possible score was 0.75 because there were four ethnic groupings (European, Māori, Pacific Peoples and All others).

An additional categorical variable distinguishing metropolitan and provincial circuits was created. Circuits were defined as metropolitan if more than 70% of the adult population was resident within the boundaries of city or suburban courts within the circuit.

Generic caseload and caseflow contextual variables

Seven generic caseload and caseflow variables were used in the modelling. There were two variables representing the judicial resource employed to address the caseload: the number of judicial officers undertaking criminal work (judges, justices of the peace and community magistrates) on a full-time equivalent basis; and the number of judges undertaking criminal work on a full-time equivalent basis. The crime rate was represented as the number of new cases per 1,000 adults. Caseload and caseflow as experienced by the

judicial officers was represented by three measures: the number of new criminal cases per full-time equivalent judicial officer; the average number of active cases per full-time equivalent judicial officer; and the number of cases disposed per full-time equivalent judicial officer. The hearing rate was represented as the proportion of disposals associated with the full or partial hearing of cases.

Offence specific caseload and caseflow contextual variables

Four variables were separately calculated for each of the three offences modelled. The four variables were: the number of offenders sentenced for the offence per 1,000 adults; the number of offenders sentenced for the offence per full time equivalent judge undertaking criminal work; the mean number of days from charge to sentence for the offence; and the proportion of offenders sentenced within 3 months of charge.

Chapter 6: Consistency between Circuits in the Use of Incarceration

This chapter reports descriptive statistics and the results of the in-out models addressing the likelihood of incarceration for each of the three offence types. The material is organised by offence with descriptive statistics followed by the sentence length linear regression models. The discussion of residual statistics has been consolidated into a single section to minimise duplication.

Inconsistency in the use of incarceration between circuits was demonstrated in the models for each of the offences. The position of the circuits in the spectrum from lenient to severe sentencing varied between the three offences, being significantly more severe for one offence did not mean significant differences existed for the other offences. Although similar numbers of circuits were significantly different to the reference circuit for aggravated drink driving (15 circuits) and burglary (13) fewer were different for male assaults female (7). The spread of relative risk also varied between the offences, there was wide variation for aggravated drink driving but substantial clustering for burglary and male assaults female. These differences validate the decision to proceed with separate models for each offence. All of the circuits that were statistically significantly different from the reference circuit were also substantively different from the reference circuit. Contrary to the proposition made by Priestley (2009) there is no evidence that compositional differences in caseload explain the variations between locations observed by the Law Commission.

Offence seriousness and criminal history variables were the principal influence on the likelihood of incarceration for all three offences. This finding is consistent with the findings from almost all sentencing research. Gender and ethnicity had an influence on sentencing decisions for burglary but not the other two offences.

The inconsistency between the circuits supports the theory that sentencing is a balance between formal and substantive rationalities. The influence of offence seriousness and criminal history variables suggests that formal rationalities are more significant and dictate the factors to be taken in to account. The finding of inconsistency combined with the dominant influence of offence seriousness and criminal history variables suggests that some of the influence of substantively rational sentencing concerns occurs via the interpretation and use of offence seriousness and criminal history factors. In the case of burglary the influence of substantively rational concerns is also seen in the effect of gender and ethnicity as sentencing factors.

Aggravated Drink Driving: Descriptive Statistics

In 2008-2009, 9,282 offenders were sentenced for aggravated drink driving by resident judges in the 17 circuits included in this study. Of these offenders 2,452 (26.4%) were incarcerated.

The demographic profile of offenders sentenced for aggravated drink driving

The demographic profile of the offenders sentenced is set out in Table 6.1. Male and Māori and Pacific offenders were over-represented compared to their respective shares of the population aged 17 or older in the 2006 census. Males were 86.4% of sentenced offenders compared to 48.0% of the population; Māori were 41.5% of sentenced offenders compared to 10.1% of the population; and Pacific offenders were 9.9% of sentenced offenders compared to 4.8% of the population.

Table 6.1: Descriptive statistics for the demographic variables for aggravated drink drivers sentenced in 2008-2009 (*n* = 9,282)

Continuous Variables	Mean	SD	Min	Max
Age at the time of offending	39.70	10.90	17	81
Categorical Variables and Coding	Coding	n	%	
Gender				
Male	0	8018	86.4	
Female	1	1264	13.6	
Ethnicity				
European	0	4031	43.4	
Māori	1	3850	41.5	
Pacific	2	919	9.9	
Other or unknown	3	482	5.2	
Age at the time of offending				
17 - 31	0	2320	25.0	
32 - 39	1	2347	25.3	
40 - 47	2	2409	25.9	
48 or older	3	2206	23.8	

Aggravated drink drivers were older than offenders in general. Almost 50% of aggravated drink drivers were 40 or older compared to 21% of offenders in general, and less than 1% of aggravated drink drivers were aged 17-19 compared to 20% of offenders in general.⁵ In part the difference is likely to be due to the time taken to acquire the prior

⁵ The comparison is not perfect, because the general age of offenders is measured at the time of sentencing whereas the age variable used in this research is the age of the offender at the time of offending. The data for all offenders was obtained on 6 March 2013 using the convicted offender calendar year statistics available from Statistics New Zealand (http://www.statistics.govt.nz/tools_and_services/tools/TableBuilder/justice-statistics.aspx)

convictions to qualify for the aggravated drink drive offence. It may also indicate that the mix of factors that contribute to offenders in general maturing out of criminal offending do not apply in the same way to drink driving.

Offenders sentenced by circuit

The volume of offenders sentenced per circuit is set out in Table 6.2. The volumes vary widely, ranging from 127 to 1,302 (M = 546, SD = 333).

Table 6.2: Descriptive statistics for the circuit variable for aggravated drink drivers sentenced in 2008-2009 ($n = 9,282$)

Categorical Variables and Coding	Coding	n	%
Circuit 0	0	818	8.8
Circuit 1	1	1117	12.0
Circuit 2	2	224	2.4
Circuit 3	3	674	7.3
Circuit 4	4	127	1.4
Circuit 5	5	1302	14.0
Circuit 6	6	559	6.0
Circuit 7	7	218	2.4
Circuit 8	8	252	2.7
Circuit 9	9	347	3.8
Circuit 10	10	427	4.6
Circuit 11	11	485	5.2
Circuit 12	12	791	8.5
Circuit 13	13	372	4.0
Circuit 14	14	194	2.1
Circuit 15	15	715	7.7
Circuit 16	16	660	7.1

Offence seriousness variables

The majority of offenders (60.5%) were sentenced for a single aggravated drink driving offence with no other accompanying convictions (see Table 6.3). On average aggravated drink drivers were close to twice (1.918 times) the legal limit. An overwhelming proportion of offenders (97.6%) were sentenced after being recorded as pleading guilty. Of the 3,677 offenders sentenced for more than one offence the most common accompanying offence was driving while disqualified (39.5%). Almost one in every twenty offenders (4.9%) was sentenced for multiple drink driving offences.

Table 6.3: Descriptive statistics for the offence seriousness variables for aggravated drink drivers sentenced in 2008-2009 ($n = 9,282$)

Continuous Variables	Mean	SD	Min	Max
Standardised alcohol level	1.918	0.528	1.0	4.5
Total charges sentenced	1.778	1.437	1	34
Categorical Variables and Coding	Coding	n	%	
Multiple drink drive charges				
No	0	8824	95.1	
Yes	1	458	4.9	
Concurrent disqualified driving charges				
No	0	7575	81.6	
Yes	1	1707	18.4	
Concurrent serious driving charges				
No	0	8771	94.5	
Yes	1	551	5.5	
Concurrent minor driving charges				
No	0	7942	85.6	
Yes	1	1340	14.4	
Concurrent breach of bail charges				
No	0	8869	95.6	
Yes	1	413	4.4	
Concurrent breach of sentence charges				
No	0	8877	95.6	
Yes	1	405	4.4	
Plea				
Guilty	0	9059	97.6	
Not guilty	1	223	2.4	
Total charges sentenced				
One	0	5615	60.5	
Two	1	2065	22.2	
Three or more	2	1602	17.3	

Case processing variables

The mean time from charge to sentencing was 15-16 weeks (108 days) but varied substantially from case to case (see Table 6.4). Almost half (49.9%) of the offenders were sentenced within three months of being charged, but almost a quarter (24.5%) were sentenced more than six months after being charged.

Table 6.4: Descriptive statistics for the case processing variables for aggravated drink drivers sentenced in 2008-2009 ($n = 9,282$)

Continuous Variables	Mean	SD	Min	Max
Elapsed time from charge to charge outcome (days)	108	87	1	366 ¹
Categorical Variables and Coding	Coding	n	%	
Elapsed time from charge to charge outcome				
≤ 6 weeks	0	2133	23.0	
> 6 weeks ≤ 3 months	1	2495	26.9	
> 3 months ≤ 6 months	2	2377	25.6	
> 6 months	3	2277	24.5	
Judge type				
Criminal	0	8213	88.5	
Family	1	1069	11.5	

¹ The maximum shown is not the true maximum. The true maximum was much greater (3,300 days) and probably the consequence of the offender absconding. The 345 cases that took more than one year were reset to a year and a day.

Criminal history variables

Offenders' criminal histories varied widely (see Table 6.5). The least prolific offenders had just the minimum two prior convictions for drink driving necessary to be an aggravated drink driver and no other convictions. By contrast the most prolific drink driver had 22 prior convictions for drink driving and the most prolific general offender had 368 prior convictions. More than half the offenders had more than two prior drink driving convictions. Most had a drink driving conviction within five years of the current sentencing. Many (37.9%) had previously been incarcerated, with 23.4% of all sentenced offenders previously incarcerated for drink driving.

Table 6.5: Descriptive statistics for the criminal history variables for aggravated drink drivers sentenced in 2008-2009 (*N* = 9,282)

Continuous Variables	Mean	SD	Min	Max
Prior convictions for drink driving	3.417	2.037	2	22
Years since last drink driving conviction	6.099	5.244	0.1	39
Prior convictions for any offence	19.527	23.243	2	368
Categorical Variables and Coding	Coding	n	%	
Prior convictions for drink driving				
Two	0	4092	44.1	
Three	1	2177	23.5	
Four	2	1199	12.9	
Five or six	3	1114	12.0	
Seven or more	4	700	7.5	
Drink driving convictions in the last 10 years				
None	0	1783	19.2	
One	1	2567	27.6	
Two	2	3292	35.5	
Three or more	3	1640	17.7	
Time since last drink driving conviction				
More than 5 years	0	4122	44.4	
5 years or less and more than 18 months	1	3687	39.7	
18 months or less	2	1473	15.9	
Prior convictions for any offence				
2 - 5	0	2222	23.9	
6 - 11	1	2379	25.6	
12 - 24	2	2376	25.6	
25 or more	3	2305	24.9	
Prior convictions for disqualified driving				
None	0	4182	45.0	
One	1	1930	20.8	
Two	2	974	10.5	
Three or four	3	989	10.7	
Five or more	4	1207	13.0	
Prior convictions for breach of sentence				
No or not in the past 5 years	0	7680	82.7	
Yes in the past 5 years	1	1602	17.3	
Prior incarceration				
Never	0	5763	62.1	
Yes but not for drink driving	1	1349	14.5	
Yes including for drink driving	2	2170	23.4	

Aggravated Drink Driving: In-out Model

The final model comprised 17 variables: 8 seriousness variables, 5 criminal history variables, 3 offender demographic variables and the circuit variable. A number of variables were included in the initial modelling but not retained because they were not statistically significant. Variables for breach of bail and minor driving charges had no significant effect on incarceration. A range of different formulations describing the offender's total prior convictions were fitted but were not significant. Likewise no formulation of the prior convictions for disqualified driving was found to be significant. The weight placed on the offender's prior convictions for drink driving and prior incarceration predominated. The inclusion of prior incarceration for an offence other than drink driving seemingly served as a broad indication of more serious offending for something other than drink driving in the past. The judge type variable distinguishing those District Court judges who primarily sat in the criminal jurisdiction from those who primarily sat in the family jurisdiction was not significant.

Variance inflation factors were measured for the variables included in the final model to detect multi-collinearity. None of the variables approached the threshold level of 10: the variance inflation factors for the variables ranged from 1.022 to 2.446.

The model includes five variables that were converted from continuous to categorical because the assumption of linearity between the continuous predictor variable and the log (logit) of the outcome variable was violated. The variables were: the age of the offender at the time of the offence; the number of days from charge to charge outcome being reached; the number of charges on the case; prior convictions for drink driving; and elapsed time since the prior conviction for drink driving.

The logistic regression parameter estimates for the offence seriousness, criminal history and demographic variables are set out in Table 6.6. These parameter estimates incorporate the influence of the circuits but the parameter estimates for the circuits are reported separately. The model correctly predicted the outcome for 85.7% of cases (92.7% for those not incarcerated and 66.5% for those incarcerated).

The offence seriousness and case processing variables had the expected effects with one exception. As expected the likelihood of incarceration increased as the alcohol level increased, if the offender faced more than one drink drive charge, if the drink driving charge was accompanied by disqualified driving, serious driving or breach of sentence convictions, as the number of charges sentenced increased, and as the number of days to reach the charge outcome increased. The exception was that offenders who were

sentenced after pleading not guilty and being found guilty following a defended hearing were less likely to be incarcerated. This is a surprising result. A possible explanation is that whatever led the offender to plead not guilty did not have the desired effect on conviction but may have contributed to a more lenient sentence. It is not substantively significant as not guilty pleas only affect 223 of the 9,282 cases included in the model.

Collectively the offence seriousness variables accounted for 34.0% of the improvement in prediction of incarceration compared to the baseline model when only the constant was included ($\Delta\chi^2 = 1644.533$, $df = 11$, $p < .001$).

The effects associated with the criminal history variables were all in the expected direction. The likelihood of incarceration increased with the number and recency of prior drink driving convictions. The differences between the categories of prior conviction were substantial. Each step in the scale was significantly higher than the previous step as well as the reference category. Elapsed time was also significant: the likelihood of incarceration increased as the number of convictions for drinking and driving in the previous ten years increased, and decreased as the number of years since the immediate prior conviction increased. A recent prior conviction for breach of sentence also increased the likelihood of incarceration. Prior incarceration increased the likelihood of incarceration, and prior incarceration for drink driving had a larger effect than prior incarceration for other offending.

Collectively the criminal history variables accounted for 39.6% of the improvement in prediction of incarceration compared to the baseline model when only the constant was included, and significantly improved the fit of the model ($\Delta\chi^2 = 1917.245$, $df = 12$, $p < .001$).

Table 6.6: Logistic regression parameter estimates for the offence seriousness, criminal history and offender demographic variables for the in-out model for aggravated drink drivers sentenced in 2008-2009 ($n = 9,282$)

Variable (reference categories in brackets)	B	S.E.	Sig.	Exp(B)	95% C.I. for Exp(B)	
					Lower	Upper
Constant	-8.666	.285	.000	.000		
Alcohol level	.529	.064	.000	1.697	1.498	1.921
Multiple drink drive (no)	1.932	.152	.000	6.906	5.123	9.308
Disqualified driving (no)	.347	.108	.001	1.414	1.144	1.749
Serious driving charge (no)	.380	.143	.008	1.462	1.105	1.936
Breach of sentence (no)	.746	.155	.000	2.109	1.558	2.856
Total convictions for sentence			.000			
Two convictions (1)	.239	.097	.014	1.270	1.050	1.535
Three or more convictions (1)	.742	.122	.000	2.100	1.652	2.670
Elapsed time - charge to charge outcome			.000			
> 6 weeks ≤ 3 months (≤ 6 weeks)	1.026	.119	.000	2.791	2.210	3.525
> 3 months ≤ 6 months (≤ 6 weeks)	1.168	.120	.000	3.216	2.540	4.070
> 6 months(≤ 6 weeks)	1.012	.127	.000	2.751	2.145	3.529
Plea (guilty)	-.512	.230	.026	.600	.382	.941
Prior convictions for drink drive			.000			
3 priors (2)	1.118	.106	.000	3.059	2.485	3.766
4 priors (2)	1.873	.125	.000	6.505	5.089	8.314
5 or 6 priors (2)	2.413	.138	.000	11.169	8.530	14.624
7 or more priors (2)	3.136	.170	.000	23.003	16.493	32.082
Drink drive convictions in the past 10 years			.000			
1 conviction (0)	1.213	.153	.000	3.365	2.491	4.545
2 convictions (0)	1.757	.159	.000	5.797	4.242	7.922
3 or more convictions (0)	1.967	.174	.000	7.148	5.079	10.059
Elapsed time since the last drink drive conviction			.000			
> 18 months ≤ 5 years (over 5 years)	.714	.086	.000	2.041	1.725	2.416
≤ 18 months (over 5 years)	1.052	.120	.000	2.864	2.265	3.622
Prior incarceration			.000			
Yes, not for drink drive (no)	.326	.097	.001	1.385	1.146	1.675
Yes, including for drink drive (no)	.577	.092	.000	1.781	1.488	2.132
Conviction for breach of sentence in last 5 years (no)	.503	.088	.000	1.653	1.393	1.963
Female (male)	-.091	.100	.362	.913	.751	1.110
Ethnicity			.805			
Māori (European)	.067	.076	.375	1.069	.922	1.240
Pacific (European)	.078	.137	.571	1.081	.826	1.415
Other or unknown (European)	-.013	.172	.940	.987	.704	1.384
Age at time of offending			.001			
Aged 32 - 39 (17 - 31)	-.225	.103	.028	.799	.653	.976
Aged 40 - 47 (17 - 31)	-.342	.114	.003	.711	.568	.889
Aged 48 or older (17 - 31)	-.493	.124	.000	.611	.479	.779
Block of 17 circuits (Table 7.7)			.000			

$\Delta\chi^2 = 4838.327$, $df = 46$, $p < .001$; R^2 Nagelkerke .593

Neither gender nor ethnicity had a statistically significant influence on the likelihood of incarceration. These results differ from much of the research conducted elsewhere, especially the lack of difference between female and male offenders. There was a statistically significant difference in the likelihood of incarceration by age. Younger offenders, those aged 17 – 31 who made up the reference category, were more likely to be incarcerated than any of the three older age categories. For aggravated drink driving the age effect may be related to the specific nature of the offence, younger offenders are generally more likely to have been convicted of the pre-requisite two drink drive convictions more quickly than older offenders. This may count against them at sentence. It is also possible that disquiet about “boy racers” and anti-drink drive publicity campaigns may have combined to cause younger offenders to be viewed as more culpable. Collectively, the demographic factors only made a small contribution (2.0%) to the improvement in prediction of incarceration compared to the baseline model when only the constant was included, but did significantly improve the fit of the overall model ($\Delta\chi^2=97.581$, $df=7$, $p<.001$).

Table 6.7 summarises the main effects associated with the circuits. Overall the addition of the circuit variables made a substantial contribution (24.4%) to the improvement in prediction of incarceration compared to the baseline model when only the constant was included, and significantly improved the fit of the overall model ($\Delta\chi^2=1178.960$, $df=16$, $p<.001$).

The odds of incarceration were significantly higher than the reference circuit in 14 of the 16 circuits. One circuit (15) was significantly lower and the result was non-significant for Circuit 13. The odds ratios for the 14 circuits with higher odds of incarceration varied substantially, from 1.73 (Circuit 9) to 58.73 (Circuit 8). The odds ratios across the circuits were not clustered together. Clustering would indicate that although almost all circuits varied from the reference circuit they were similar to one another. The spectrum of odds ratios and associated 95% confidence intervals is portrayed graphically in Figure 6.1.

Table 6.7: Logistic regression parameter estimates for the circuit dummy variables for the in-out model for aggravated drink drivers sentenced in 2008-2009 ($n = 9,282$)

Circuit (Reference Circuit 0)	B	S.E.	Sig.	Exp(B)	95% C.I. for Exp(B)	
					Lower	Upper
Circuit			.000			
Circuit 1	2.329	.168	.000	10.269	7.389	14.270
Circuit 2	2.752	.243	.000	15.671	9.731	25.236
Circuit 3	1.337	.186	.000	3.808	2.646	5.480
Circuit 4	2.980	.292	.000	19.686	11.115	34.867
Circuit 5	.811	.171	.000	2.250	1.609	3.147
Circuit 6	3.081	.188	.000	21.789	15.088	31.466
Circuit 7	2.675	.236	.000	14.519	9.139	23.066
Circuit 8	4.073	.231	.000	58.726	37.327	92.394
Circuit 9	.550	.244	.024	1.733	1.074	2.797
Circuit 10	2.436	.195	.000	11.433	7.802	16.753
Circuit 11	1.840	.194	.000	6.296	4.308	9.201
Circuit 12	3.428	.176	.000	30.819	21.821	43.528
Circuit 13	-.446	.278	.108	.640	.371	1.104
Circuit 14	2.135	.255	.000	8.458	5.135	13.932
Circuit 15	-.514	.233	.027	.598	.379	.944
Circuit 16	2.192	.184	.000	8.952	6.242	12.840

Figure 6.1: Odds ratios and 95% confidence intervals for incarceration of aggravated drink drivers by circuit (Circuit 0 is the reference)

The high odds ratio for Circuit 8 tends to suppress perception of the degree of difference between the circuits. The confidence intervals provide a guide to how the results would fall if a different circuit had been chosen as the reference circuit. For instance, had Circuit 8 been chosen 15 circuits would have had significantly lower odds ratios with no significant difference between Circuit 8 and Circuit 12. If the circuit at the centre of the spectrum before controls (Circuit 16) had been chosen, the circuits could have been formed into three groups: a lower severity grouping (15, 13, 0, 9, 5 and 3); a median grouping (11, 14, 16, 1, 10, 7, 2 and 4); and a higher severity grouping (6, 12, 8). These are nominal groupings as evidenced by the fact that there are overlapping confidence intervals between the circuits at the high end of one group and the low end of the next and within each grouping there are some pairs of circuits with no overlap between their respective confidence intervals.

As discussed in Chapter 4 the odds ratio is sometimes confused with relative risk. Odds ratios are not proportionate measures, whereas the relative risk is. The more intuitively easily understood relative risk comparison can be derived from the odds ratios and the unadjusted relative risk. The relative risks of incarceration by circuit are presented in Figure 6.2.

Figure 6.2: Relative risk of incarceration for aggravated drink driving by circuit

The relative risk puts the risk of incarceration into proportion. The higher odds of incarceration in Circuit 8 (odds 58 times greater than Circuit 0) translate to offenders being more than 7.4 times more likely to be incarcerated. The spectrum of differences portrayed by the relative risk (Figure 6.2) is more regular than the spectrum portrayed by the odds ratio (Figure 6.1).

Prediction 1 is supported for aggravated drink driving: after controlling for seriousness of offence, criminal history and offender demographic variables, similarly situated offenders were sentenced differently between circuits. The differences were not just statistically significant but were also substantial.

Male Assaults Female: Descriptive Statistics

In 2008-2009 6,166 offenders were sentenced for male assaults female by resident judges in the 17 circuits included in this study. Of these offenders 744 (12.1%) were incarcerated.

The demographic profile of offenders sentenced for male assaults female offending

The demographic profile of the offenders sentenced is set out in Table 6.8. Māori and Pacific offenders are over-represented compared to their respective shares of the male population aged 17 or older in the 2006 census. Māori were 45.8% of offenders sentenced compared to 10.6% of the population; and Pacific males were 15.9% of the offenders sentenced compared to 5.1% of the population.

Table 6.8: Descriptive statistics for the demographic variables for male assault female offenders sentenced in 2008-2009 (*n* = 6,166)

Continuous Variables	Mean	SD	Min	Max
Age at the time of offending	32.214	10.224	17	82
Categorical Variables and Coding	Coding	n	%	
Ethnicity				
European	0	1839	29.8	
Māori	1	2825	45.8	
Pacific	2	979	15.9	
Other or unknown	3	523	8.5	
Age at the time of offending				
17 - 23	0	1472	23.9	
24 - 30	1	1529	24.8	
31 - 38	2	1554	25.2	
39 or older	3	1611	26.1	

The age profile for male assaults female offenders differed from the general age of offenders at sentence in 2008-2009. Although the difference is less marked than it was for aggravated drink driving there are fewer young male assaults female offenders and a higher proportion of older offenders, especially offenders in their thirties. Only 8% of male assaults female offenders were aged 17-19 compared to 20% of offenders in general, whereas 31% of male assaults female offenders were aged in their thirties compared to 21% of offenders in general⁶.

Offenders sentenced by circuit

The volume of offenders sentenced per circuit is set out in Table 6.9. The volumes vary widely ranging from 78 to 863 (M = 363, SD = 229).

Table 6.9: Descriptive statistics for the circuit variable for male assaults female offenders sentenced in 2008-2009 (*n* = 6,166)

Categorical Variables and Coding	Coding	n	%
Circuit 0	0	701	11.4
Circuit 1	1	509	8.3
Circuit 2	2	181	3.0
Circuit 3	3	459	7.4
Circuit 4	4	98	1.6
Circuit 5	5	863	14.0
Circuit 6	6	365	5.9
Circuit 7	7	118	1.9
Circuit 8	8	175	2.8
Circuit 9	9	125	2.0
Circuit 10	10	231	3.8
Circuit 11	11	469	7.6
Circuit 12	12	496	8.0
Circuit 13	13	355	5.7
Circuit 14	14	78	1.3
Circuit 15	15	614	10.0
Circuit 16	16	329	5.3

⁶ The comparison is not perfect, because the general age of offenders is measured at the time of sentencing whereas the age variable used in this research is the age of the offender at the time of offending. The data for all offenders was obtained on 6 March 2013 using the convicted offender calendar year statistics available from Statistics New Zealand (http://www.statistics.govt.nz/tools_and_services/tools/TableBuilder/justice-statistics.aspx)

Offence seriousness variables

The statistics for offence seriousness variables are set out in Table 6.10.

Table 6.10: Descriptive statistics for the offence seriousness variables for male assaults female offending sentenced in 2008-2009 (*n* = 6.166)

Continuous Variables	Mean	SD	Min	Max
Total charges sentenced	1.853	1.546	1	19
Categorical Variables and Coding	Coding	n	%	
Multiple male assaults female charges				
No	0	5639	91.5	
Yes	1	527	8.5	
Concurrent threat to kill or do GBH				
No	0	6139	99.6	
Yes	1	27	0.4	
Concurrent breach of protection order				
No	0	5645	91.6	
Yes	1	521	8.4	
Concurrent other violence				
No	0	5339	86.6	
Yes	1	827	13.4	
Concurrent wilful or intentional damage				
No	0	5558	90.1	
Yes	1	608	9.9	
Charge type indicates use of weapon				
No	0	6070	98.4	
Yes	1	96	1.6	
Concurrent breach of bail charges				
No	0	5754	93.3	
Yes	1	412	6.7	
Concurrent breach of sentence charges				
No	0	5864	95.1	
Yes	1	302	4.9	
Plea				
Guilty	0	5723	92.8	
Not guilty	1	443	7.2	
Total charges sentenced				
One	0	3672	59.6	
Two	1	1303	21.1	
Three	2	553	9.0	
Four or more	3	638	10.3	

Most offenders (almost 60%) were sentenced for a single male assaults female offence with no accompanying offences. A further 21% had one accompanying charge. The 2,494 offenders who were sentenced for multiple offences mainly faced other violence, damage and breach offences and included 527 offenders (8.5% of the total) who were sentenced having been convicted of two or more charges of male assaults female. Based on whether the charge code for the offence included a reference to a weapon 1.6% of cases included a weapon. A high proportion of offenders (92.8%) pleaded guilty.

Case processing variables

The mean time from charge to sentencing was 16-17 weeks (116 days) but varied substantially between cases. Half the offenders were sentenced within three months of being charged, but a fifth (21.2%) were sentenced more than six months after being charged. Only a small proportion of charges were laid indictably (85 or 1.4%), a proxy for more serious offending.

Family Court judges sentenced 21.7% of offenders, which was approximately twice the rate for aggravated drink driving and burglary. Just over one-third of offenders were sentenced in a location with an active family violence court. Family Court judges undertook a higher proportion of sentencing in family violence courts (38.1%) than elsewhere (12.6%). The 12.6% rate from non-family violence courts was commensurate with aggravated drink driving (11.5%) and burglary (10.9%) suggesting that Family Court judges are significant actors in the family violence courts. The statistics for case processing variables are set out in Table 6.11.

There was a difference in elapsed time from charge to charge outcome between the family violence courts and ordinary criminal courts, with 58.3% of offenders in the ordinary criminal courts sentenced within three months of the charge being laid compared to 35.4% in the family violence courts. Conversely 14% of offenders in the ordinary criminal courts were sentenced more than six months after the charge was laid compared to 34% in the family violence court locations. This may be a consequence of a difference of process with sentencing delayed in the family violence courts to allow time for an offender to undertake agreed actions.

Table 6.11: Descriptive statistics for the case processing variables for male assaults female offending sentenced in 2008-2009 (*n* = 6,166)

Continuous Variables	Mean	SD	Min	Max
Elapsed time from charge to charge outcome (days)	116	97	1	366 ¹
Categorical Variables and Coding	Coding	n	%	
Elapsed time from charge to charge outcome				
≤ 6 weeks	0	1665	27.0	
> 6 weeks ≤ 3 months	1	1422	23.0	
> 3 months ≤ 6 months	2	1772	28.8	
> 6 months	3	1307	21.2	
Charge laid indictably				
No	0	6081	98.6	
Yes	1	85	1.4	
Judge type				
Criminal	0	4827	78.3	
Family	1	1339	21.7	
Case resolved in an active family violence court location				
No	0	3953	64.1	
Yes	1	2213	35.9	

¹ The maximum shown is not the true maximum; the true maximum was much greater (2,985 days) and probably the consequence of the offender absconding. The 246 cases that took more than one year were reset to a year and a day.

Criminal history variables

The statistics for the criminal history variables are set out in Table 6.12. The criminal histories of offenders sentenced for male assaults female varied widely. The average offender had 15 prior convictions for any offence but almost one in five offenders had no prior convictions. The most prolific offender had 293 prior convictions. Over 70% had no prior convictions for male assaults female, but 5.5% had three or more and the most prolific male assaults female offender sentenced had 11 prior convictions for the offence. Most offenders (68.6%) had not previously been incarcerated but 5.5% or 524 had been incarcerated on five or more separate occasions. The proportions of offenders with more relevant convictions in the recent past tended to be low. In the five years prior to the current sentencing, 18.2% of the offenders had been sentenced for male assaults female and 8.6% had been sentenced for a serious violent offence.

Table 6.12: Descriptive statistics for the criminal history variables for male assaults female offenders sentenced in 2008-2009 (*n* = 6,166)

Continuous Variables	Mean	SD	Min	Max
Prior convictions for male assaults female	0.510	1.049	0	11
Prior convictions for any offence	15.202	23.610	0	293
Categorical Variables and Coding	Coding	n	%	
Prior convictions for male assaults female				
None	0	4398	71.3	
One	1	1046	17.0	
Two	2	384	6.2	
Three or more	3	338	5.5	
Convicted for male assaults female last 5 years				
No	0	5043	81.8	
Yes	1	1123	18.2	
Convicted for serious violence last 5 years				
No	0	5635	91.4	
Yes	1	531	8.6	
Convicted for breach of protection order last 5 years				
No	0	5773	93.6	
Yes	1	393	6.4	
Convicted for breach of sentence last 5 years				
No	0	4889	79.3	
Yes	1	1277	20.7	
Prior convictions for any offence				
None	0	1134	18.4	
1 - 9	1	2483	40.3	
10 - 19	2	1047	17.0	
20 or more	3	1502	24.3	
Prior incarceration for any offence				
Never	0	4233	68.6	
Once	1	694	11.3	
2 - 4 times	2	715	11.6	
5 or more times	3	524	8.5	

Male Assaults Female: In-out Model

The final model comprised 17 variables: 7 seriousness variables, 7 criminal history variables, 2 offender demographic variables and the circuit variable. A number of variables were included in the initial modelling but not retained because they were not statistically significant. Offence seriousness variables for concurrent convictions for breach of a protection order, breach of bail, another violent offence, or wilful or intentional damage or for use of a weapon had no significant effect on incarceration and were omitted. The judge type variable distinguishing those District Court judges who primarily sat in the criminal jurisdiction from those who primarily sat in the family jurisdiction and a variable distinguishing whether the offender was sentenced in an active family violence court were not significant and were omitted.

Variance inflation factors were measured for the 17 variables included in the final model to detect multi-collinearity. None of the variables approached the threshold level of 10: the variance inflation factors for the 17 variables ranged from 1.018 to 2.607.

The model contains four variables that were converted from continuous to categorical because the assumption of linearity between the continuous predictor variable and the log (logit) of the outcome variable was violated. The converted variables were: the age of the offender; the number of charges on the case; prior convictions for male assaults female; and prior convictions for any offence.

The logistic regression parameter estimates for the offence seriousness, case processing, criminal history and demographic variables are set out in Table 6.13. These parameter estimates incorporate the influence of the circuits but the parameter estimates for the circuits are reported separately. The model correctly predicted the outcome for 91.3% of cases (97.4% for those not incarcerated and 47.0% for those incarcerated).

Table 6.13: Logistic regression parameter estimates for the offence seriousness, case processing, criminal history and offender demographic variables for the in-out model for male assaults female offenders sentenced in 2008-2009 ($n = 6,166$)

Variable (reference categories in brackets)	B	S.E.	Sig.	Exp(B)	95% C.I. for Exp(B)	
					Lower	Upper
Constant	-5.679	.371	.000	.003		
Multiple male assaults female (no)	.530	.152	.000	1.700	1.261	2.290
Threat to kill or do GBH (no)	1.882	.618	.002	6.563	1.956	22.018
Breach of sentence (no)	.579	.174	.001	1.784	1.269	2.509
Total convictions for sentence			.000			
Two convictions (1)	.783	.139	.000	2.187	1.666	2.872
Three convictions (1)	1.031	.172	.000	2.804	2.003	3.927
Four or more convictions (1)	2.064	.163	.000	7.876	5.720	10.845
Plea (guilty)	.758	.184	.000	2.133	1.486	3.061
Elapsed time - charge to charge outcome			.001			
> 6 weeks ≤ 3 months (≤ 6 weeks)	.495	.151	.001	1.640	1.221	2.204
> 3 months ≤ 6 months (≤ 6 weeks)	.346	.153	.024	1.414	1.048	1.907
> 6 months(≤ 6 weeks)	-.004	.177	.981	.996	.704	1.409
Charge laid indictably (no)	1.870	.306	.000	6.486	3.559	11.821
Prior convictions for male assaults female			.000			
1 prior (0)	.118	.175	.501	1.125	.798	1.586
2 priors (0)	.642	.211	.002	1.901	1.256	2.876
3 or more priors (0)	1.221	.217	.000	3.392	2.215	5.194
Conviction for male assaults female in last 5 years (no)	.609	.165	.000	1.838	1.331	2.539
Conviction for serious violence in last 5 years (no)	.908	.133	.000	2.480	1.913	3.216
Conviction for breach of protection order in last 5 years (no)	.515	.151	.001	1.674	1.244	2.251
Conviction for breach of sentence in last 5 years (no)	.723	.117	.000	2.060	1.637	2.593
Prior convictions for any offence			.000			
1-9 priors (0)	.640	.315	.042	1.897	1.024	3.515
10-19 priors (0)	1.308	.331	.000	3.698	1.932	7.078
20 or more priors (0)	1.229	.355	.001	3.417	1.705	6.847
Prior incarceration for any offence			.000			
Once (0)	.716	.166	.000	2.046	1.478	2.833
2-4 times (0)	.914	.185	.000	2.493	1.734	3.585
5 or more times (0)	1.694	.211	.000	5.443	3.596	8.239
Ethnicity			.332			
Māori (European)	.142	.132	.283	1.152	.889	1.492
Pacific (European)	.350	.198	.077	1.419	.963	2.091
Other or unknown (European)	.277	.304	.362	1.319	.728	2.392
Age at time of offending			.000			
Aged 24-30 (17-23)	-.422	.153	.006	.656	.486	.885
Aged 31-38 (17-23)	-.371	.163	.023	.690	.501	.951
Aged 39 or older (17-23)	-.892	.185	.000	.410	.285	.589
Block of 17 circuits (Table 7.14)			.000			

$\Delta\chi^2 = 1921.32$, $df = 46$, $p < .001$; R^2 Nagelkerke .514

The offence seriousness variables had the expected effects with one partial exception. The partial exception was that offenders whose cases were not sentenced until more than six months after charges were laid were not significantly more likely to be incarcerated than those who were sentenced within six weeks of charge, although offenders in two intermediate categories between 6 weeks and 6 months were more likely to be incarcerated. Although the family violence court variable was not significant in its own right it is possible that the unexpected result for those sentenced after 6 months is associated with the higher proportion of older cases sentenced in the family violence courts, with the cases taking longer to allow the offenders to complete agreed actions before sentencing occurred. Otherwise the likelihood of incarceration increased as expected if the offender was sentenced for multiple assaults by a male on a female, had a concurrent conviction for threat to kill or do GBH, had a concurrent conviction for breach of a sentence, faced more than one charge, pleaded not guilty, or faced a charge laid indictably.

Collectively the offence seriousness variables accounted for 47.6% of the improvement in prediction of incarceration compared to the baseline model when only the constant was included ($\Delta\chi^2=914.07$, $df=11$, $p<.001$).

The effects associated with the criminal history variables were as expected and were all statistically significant with one partial exception. The partial exception was that offenders with one prior conviction for male assaults female were not significantly more likely to be incarcerated than those with no prior convictions for the offence. The likelihood of incarceration increased as the total number of prior convictions increased, as the number of prior terms of incarceration increased, if the offender had more than one prior conviction for male assaults female, if the most recent of those convictions was in the last 5 years, if the offender had a recent conviction for serious violence, if the offender had a recent conviction for breaching a protection order, or if the offender had a recent conviction for breaching a sentence.

Collectively the criminal history variables accounted for 47.9% of the improvement in prediction of incarceration compared to the baseline model when only the constant was included, and significantly improved the fit of the model ($\Delta\chi^2=919.83$, $df=13$, $p<.001$).

Ethnicity had no statistically significant influence on the likelihood of incarceration. There was a statistically significant difference in the likelihood of incarceration by age. Younger offenders, those aged 17 – 23 who made up the reference category, were more likely to be incarcerated than any of the three older age categories.

Collectively the demographic factors only made a small contribution (1.5%) to the improvement in prediction of incarceration compared to the baseline model when only the constant was included, but did significantly improve the fit of the overall model ($\Delta\chi^2=29.33$, $df=6$, $p<.001$) .

Table 6.14 summarises the main effects associated with the circuits. Overall the addition of the circuit variables made a small contribution (3%) to the improvement in prediction of incarceration compared to the baseline model when only the constant was included, and significantly improved the fit of the overall model ($\Delta\chi^2=58.08$, $df=16$, $p<.001$).

Table 6.14: Logistic regression parameter estimates for the circuit dummy variables for the in-out model for male assaults female offenders sentenced in 2008-2009 ($n=6,166$)

Circuit (Reference Circuit 0)	B	S.E.	Sig.	Exp(B)	95% C.I. for Exp(B)	
					Lower	Upper
Circuit			.000			
Circuit 1	.836	.228	.000	2.308	1.475	3.610
Circuit 2	.761	.308	.013	2.140	1.170	3.914
Circuit 3	.368	.233	.114	1.444	.915	2.279
Circuit 4	.587	.401	.143	1.798	.819	3.944
Circuit 5	.382	.219	.081	1.465	.954	2.252
Circuit 6	.082	.264	.755	1.086	.647	1.822
Circuit 7	1.048	.358	.003	2.852	1.413	5.757
Circuit 8	.241	.334	.470	1.273	.662	2.448
Circuit 9	.351	.478	.463	1.420	.557	3.622
Circuit 10	.762	.272	.005	2.142	1.256	3.653
Circuit 11	-.131	.245	.593	.878	.543	1.417
Circuit 12	.618	.226	.006	1.855	1.190	2.891
Circuit 13	-.981	.359	.006	.375	.185	.758
Circuit 14	1.000	.402	.013	2.718	1.237	5.973
Circuit 15	.156	.269	.562	1.169	.690	1.979
Circuit 16	.348	.279	.212	1.416	.820	2.447

The odds of incarceration were significantly higher than the reference circuit in 6 of the 16 circuits. One circuit was significantly lower and the result was non-significant for nine circuits. The odds ratios for the six circuits with higher odds of incarceration varied from 1.854 (Circuit 12) to 2.852 (Circuit 7). The spectrum of odds ratios and associated 95% confidence intervals is portrayed graphically in Figure 6.3.

Figure 6.3: Odds ratios and confidence intervals for incarceration of male assaults female offenders sentenced in 2008-2009 by circuit (Circuit 0 is the reference)

The confidence intervals provide a guide to how the results would fall if a different circuit had been chosen as the reference circuit. For instance, had Circuit 7 been chosen 2 circuits would have had significantly lower odds ratios and there would not be any significant difference between Circuit 7 and the remaining 14 circuits. If the circuit at the centre of the spectrum before controls (Circuit 6) had been chosen no circuits would have been statistically significantly different from the reference circuit. However, a review of the confidence intervals would have identified that some circuits did differ significantly from one another just not from Circuit 6.

The relative risks of incarceration by circuit for male assaults female are presented in Figure 6.4. Figure 6.4 puts the risk of incarceration into proportion. The higher odds of incarceration in Circuit 7 (odds 2.85 times greater than Circuit 0) translate to offenders being more than 2.34 times more likely to be incarcerated. Unlike aggravated drink driving the spectrum of differences portrayed by the relative risk is similar to the spectrum portrayed by the odds ratio. This has occurred because odds ratios are a closer approximation of relative risk when the odds ratios are lower.

Figure 6.4: Relative risk of incarceration for male assaults female offenders sentenced in 2008-2009 by circuit

The odds ratios across the circuits were much more tightly clustered than for aggravated drink driving. There are two substantial clusters. First, the six circuits (1, 2, 7, 10, 12, and 14) that have significantly and substantially higher likelihoods of incarceration, and second, the balance of circuits, excluding circuit 13, that are not significantly different to the reference circuit. Circuit 13 stands on its own with a significantly and substantially lower likelihood of incarceration.

Prediction 1 is supported for male assaults female: after controlling for seriousness of offence, criminal history and offender demographic variables, similarly situated offenders were sentenced differently between circuits.

Burglary: Descriptive Statistics

In 2008-2009 5,740 offenders were sentenced for burglary by resident judges in the 17 circuits included in this study. Of these offenders 2,434 (42.4%) were incarcerated.

The demographic profile of offenders sentenced for burglary

The demographic profile of the offenders sentenced is set out in Table 6.15. Male and Māori and Pacific offenders were over-represented compared to their respective shares of the population aged 17 or older in the 2006 census. Males were 90.6% of sentenced offenders compared to 48.0% of the population census; Māori were 50.6% of sentenced offenders compared to 10.1% of the population; and Pacific offenders were 7.6% of sentenced offenders compared to 4.8% of the population.

Table 6.15: Descriptive statistics for the demographic variables for burglars (*n* = 5,740)

Continuous Variables	Mean	SD	Min	Max
Age at the time of offending	24.117	8.127	13	71
Categorical Variables and Coding	Coding	n	%	
Gender				
Male	0	5202	90.6	
Female	1	538	9.4	
Ethnicity				
European	0	2202	38.4	
Māori	1	2903	50.6	
Pacific	2	437	7.6	
Other or unknown	3	198	3.4	
Age at the time of offending				
13 - 19	0	2286	39.8	
20 - 29	1	2209	38.5	
30 - 39	2	868	15.1	
40 or older	3	377	6.6	

The age distribution for burglars was the opposite of the other two offences. Burglars were younger than offenders in general. Over 39% of the burglars were aged 17-19, almost twice the proportion for offenders in general (20%), and only 7% were over 40 compared to 21% of offenders in general⁷.

⁷ The comparison is not perfect, because the general age of offenders is measured at the time of sentencing whereas the age variable used in this research is the age of the offender at the time of offending. The data for all offenders was obtained on 6 March 2013 using the convicted offender calendar year statistics available from Statistics New Zealand (http://www.statistics.govt.nz/tools_and_services/tools/TableBuilder/justice-statistics.aspx)

Offenders sentenced by circuit

The volume of offenders sentenced per circuit is set out in Table 6.16. The volumes vary widely ranging from 120 to 747 (M = 338, SD = 196).

Table 6.16: Descriptive statistics for the circuit variable for burglars sentenced in 2008-2009 ($n = 5,740$)

Categorical Variables and Coding	Coding	n	%
Circuit 0	0	522	9.1
Circuit 1	1	747	13.0
Circuit 2	2	193	3.4
Circuit 3	3	522	9.1
Circuit 4	4	120	2.1
Circuit 5	5	694	12.1
Circuit 6	6	278	4.8
Circuit 7	7	143	2.5
Circuit 8	8	227	4.0
Circuit 9	9	165	2.9
Circuit 10	10	335	5.8
Circuit 11	11	281	4.9
Circuit 12	12	490	8.5
Circuit 13	13	208	3.6
Circuit 14	14	120	2.1
Circuit 15	15	426	7.4
Circuit 16	16	269	4.7

Offence seriousness variables

For ease of presentation the offence seriousness variables have been split between two tables. Table 6.17 describes the variables that were included in the final model and Table 6.18 describes variables that were included in earlier iterations but were not significant and not retained.

Table 6.17: Descriptive statistics for the offence seriousness variables for burglary included in the final model ($n = 5,740$)

Continuous Variables	Mean	SD	Min	Max
Total charges sentenced	4.302	5.164	1	113
Categorical Variables and Coding	Coding	n	%	
Burglary convictions				
One conviction	0	4157	72.4	
Two convictions	1	821	14.3	
Three or more convictions	2	762	13.3	
Concurrent serious receiving, unlawful taking, serious violence or serious driving conviction				
No	0	4222	73.6	
Yes	1	1518	26.4	
Concurrent breach of sentence charges				
No	0	4821	84	
Yes	1	919	16	
Approximate value of the lead burglary offence				
Less than \$500	0	2061	35.9	
\$500 to \$5,000	1	2589	45.1	
More than \$5,000	2	674	11.7	
Unknown	3	416	7.3	
Total charges sentenced				
One	0	1825	31.8	
Two	1	992	17.3	
Three or four	2	1145	19.9	
Five, six or seven	3	877	15.3	
Eight or more	4	901	15.7	
Order for reparation				
No	0	3132	54.6	
Yes	1	2608	45.4	

Table 6.18: Descriptive statistics for the offence seriousness variables for burglary not included in the final model ($n = 5,740$)

Categorical Variables and Coding	Coding	n	%
Burglary by day or night			
Day	0	2416	42.1
Night	1	2908	50.7
Unknown	2	416	7.2
Burglary offence dates			
One	0	4412	76.9
Two	1	762	13.3
Three or more	2	566	9.8
Total offence dates			
One	0	2366	41.2
Two	1	1089	19
Three or four	2	1118	19.5
Five or more	3	1167	20.3
Concurrent receiving conviction			
No	0	5226	91
Yes	1	514	9
Concurrent unlawful taking conviction			
No	0	5224	91
Yes	1	516	9
Concurrent other dishonesty conviction			
None	0	4159	72.5
One	1	793	13.8
Two or more	2	788	13.7
Concurrent violent conviction			
No	0	4870	84.8
Yes	1	870	15.2
Concurrent drug conviction			
No	0	5163	89.9
Yes	1	577	10.1
Concurrent serious driving conviction			
No	0	5087	88.6
Yes	1	653	11.4
Concurrent breach of bail convictions			
No	0	4876	84.9
Yes	1	864	15.1
Plea			
Guilty	0	5521	96.2
Not guilty	1	219	3.8
Reparation amount			
None	0	3433	59.8
< \$500	1	892	15.5
$\geq \$500 < \$1,000$	2	413	7.2
$\geq \$1,000 < \$2,500$	3	521	9.1
$\geq \$2,500$	4	481	8.4

Burglars were more likely than not to be sentenced for multiple offences, with 68.2% of burglars sentenced for two or more charges and an average of 4.3 charges per case. In most cases (72.3%) the offender was sentenced for a single burglary offence, although many also sentenced for other offences. More than one in four (26.4%) had an accompanying conviction for a serious dishonesty offence. A substantial proportion of offenders (45.4%) were ordered to pay reparation. An overwhelming proportion of offenders (96.2%) were sentenced having pleaded guilty.

Case processing variables

Time from charge to sentence, charge laid type and the judge type are described in Table 6.19. The mean time from charge to sentencing (120 days) varied substantially.

Table 6.19: Descriptive statistics for the case processing variables for burglary ($n = 5,740$)

Continuous Variables	Mean	SD	Min	Max
Elapsed time from charge to charge outcome (days)	120	96	1	366 ¹
Categorical Variables and Coding	Coding	n	%	
Elapsed time from charge to charge outcome				
≤ 6 weeks	0	1212	21.1	
> 6 weeks ≤ 3 months	1	1628	28.4	
> 3 months ≤ 6 months	2	1681	29.3	
> 6 months	3	1219	21.2	
Charge laid indictably				
No	0	5187	90.4	
Yes	1	553	9.6	
Judge type				
Criminal	0	5116	89.1	
Family	1	624	10.9	

¹ The maximum shown is not the true maximum; the true maximum was much greater (4,429 days) and probably the consequence of the offender absconding. The 267 cases that took more than one year were reset to a year and a day.

Almost half (49.5%) of the offenders were sentenced within three months of being charged, but over a fifth (21.2%) were sentenced more than six months after the charge was laid. Family Court judges sentenced 10.9% of burglars. The charge for the lead burglary offence was laid indictably (a proxy for more serious offending) in 9.6 % of cases.

Criminal history variables

Offenders' criminal histories varied widely (see Table 6.20).

Table 6.20: Descriptive statistics for the criminal history variables for burglars ($n = 5,740$)

Continuous Variables	Mean	SD	Min	Max
Prior convictions for burglary	3.112	9.132	0	382
Prior convictions for serious dishonesty	8.473	19.771	0	396
Prior convictions for any offence	22.363	34.421	0	471
Categorical Variables and Coding	Coding	n	%	
Prior convictions for burglary				
None	0	3134	54.6	
One	1	701	12.2	
Two, three or four	2	879	15.3	
Five or more	3	1026	17.9	
Convicted for burglary last 5 years				
No	0	3669	63.9	
Yes	1	2071	36.1	
Prior incarceration for burglary				
No	0	4082 ¹	71.1	
Once	1	674	11.7	
Twice or more	2	984	17.2	
Longest term of incarceration for burglary				
Never incarcerated for burglary	0	4095 ¹	71.3	
≤12 months	1	773	13.5	
>12 months ≤24 months	2	547	9.5	
> 24 months	3	325	5.7	
Prior convictions for serious dishonesty including burglary				
None	0	2227	38.8	
One or two	1	1004	17.5	
Three to nine	2	1177	20.5	
Ten or more	3	1332	23.2	
Prior convictions for serious violence				
No	0	4536	79.0	
Yes	1	1204	21.0	
Convicted for breach of sentence last 5 years				
No	0	3586	62.5	
Yes	1	2154	37.5	
Prior convictions for any offence				
None	0	915	15.9	
1 - 4	1	1159	20.2	
5 - 9	2	834	14.5	
10 - 19	3	897	15.6	
20 - 49	4	1145	20.0	
50 or more	5	790	13.8	

¹ The numbers for no prior incarceration and never incarcerated for burglary are not equal because the never incarcerated includes 13 observations where there was an incarceration sentence but the amount is unknown because of re-sentencing or a

successful appeal resulting in the original sentence term being set to zero in the Ministry of Justice Datawarehouse.

The least prolific offenders had no prior convictions for any offence, and the most prolific offender had 471 prior convictions. If criminal history is restricted to prior burglary convictions there was a similar range from 0 to 382 prior convictions for burglary. More than half the offenders sentenced (54.6%) had no prior convictions for burglary. However, only 15.9% of the offenders had no prior convictions for any offence. A high proportion of those with a prior conviction for burglary (2,071/2,606 or 79.5%) had at least one burglary conviction within five years of the current sentencing. A substantial minority of offenders (28.9%) had previously been incarcerated for burglary on at least one occasion and 41.5% of the offenders had previously been incarcerated for any offence on at least one occasion.

Burglary: In-out Model

The final model comprised 18 variables: 8 seriousness variables, 6 criminal history variables, 3 offender demographic variables and the circuit variable. A number of variables were included in the initial modelling but not retained because they were not statistically significant. Offence seriousness variables for concurrent convictions for receiving, unlawful taking, other dishonesty, a violent offence, a drug offence, serious driving, or a breach of bail had no significant effect on incarceration. The variables for concurrent convictions for receiving, unlawful taking and serious driving were consolidated into a concurrent serious offence variable along with serious violent and serious drug offences and the consolidated variable was included in the final model. Variables representing the plea, the value of reparation ordered and whether the burglary was committed in the daytime or night-time were not significant. The judge type variable distinguishing those District Court judges who primarily sat in the criminal jurisdiction from those who primarily sat in the family jurisdiction was not significant.

A variety of criminal history variables were also rejected. Criminal history is challenging because it can be represented in many ways by aggregating or disaggregating offence types or taking account of the age of past convictions or the seriousness of the sentences imposed. The criminal history variables included in earlier models and rejected were: prior convictions for burglary, prior incarceration for burglary, prior convictions for receiving, prior convictions for unlawful taking, prior convictions for other serious dishonesty (excluding burglary, receiving and unlawful taking), a composite variable for

serious dishonesty excluding burglary (receiving, unlawful taking and other serious dishonesty), prior incarceration for any offence, the longest term of incarceration for any offence, and prior convictions for breach of sentence. The rejection of prior convictions for burglary was, at first consideration, surprising. However, it became understandable when placed in the context of the variables that were included. Prior convictions for burglary were included in the variable covering all prior convictions for serious dishonesty. Indications of the offender's burglary history were also included in the variables measuring time since the most recent prior burglary conviction (if any) and the longest previous term of incarceration for burglary (if any).

Variance inflation factors were measured for the 18 variables included in the final model to detect any multi-collinearity. None of the variables approached the threshold level of 10: the variance inflation factors for the 18 variables ranged from 1.027 to 5.423.

The continuous variables were tested to determine if the assumption of linearity between the continuous predictor variable and the log (logit) of the outcome variable was met. The model contains eight variables that were converted from continuous to categorical because the assumption was violated. The converted variables were: the age of the offender at the time of the offence; the number of days from charge to charge outcome being reached; the number of burglary charges for sentence; the number of charges on the case; prior convictions for serious dishonesty; elapsed time since the prior conviction for burglary; the longest prior term of incarceration for burglary; and prior convictions for any offence.

The logistic regression parameter estimates for the offence seriousness, case processing, criminal history and demographic variables are set out in Table 6.21. These parameter estimates incorporate the influence of the circuits but the parameter estimates for the circuits are reported separately. The model correctly predicted the outcome for 82.9% of cases (87.0% for those not-incarcerated and 77.4% for those incarcerated).

Table 6.21: Logistic regression parameter estimates for the offence seriousness, case processing, criminal history and offender demographic variables for the in-out model for burglars sentenced in 2008-2009 ($n = 5,740$)

Variable (reference categories in brackets)	B	S.E.	Sig.	Exp(B)	95% C.I. for Exp(B)	
					Lower	Upper
Constant	5.567	0.253	.000	0.004		
Burglary convictions for sentence			.000			
Two (one)	.461	.117	.000	1.585	1.260	1.993
Three or more (one)	1.087	.140	.000	2.964	2.254	3.897
Other serious conviction (no)	.463	.102	.000	1.588	1.301	1.939
Breach of sentence conviction (no)	.528	.115	.000	1.695	1.352	2.126
Approximate value of the lead burglary offence			.000			
\$500 to \$5,000 (< \$500)	.449	.089	.000	1.566	1.317	1.863
More than \$5,000 (< \$500)	1.022	.137	.000	2.779	2.126	3.633
Unknown (< \$500)	.367	.155	.018	1.443	1.065	1.955
Total convictions for sentence			.000			
Two (one)	.422	.127	.001	1.525	1.189	1.957
Three or four (one)	.900	.126	.000	2.461	1.921	3.152
Five, six or seven (one)	1.409	.151	.000	4.092	3.045	5.499
Eight or more (one)	2.413	.171	.000	11.166	7.993	15.600
Reparation ordered (no)	-.513	.081	.000	.599	.510	.702
Charge laid indictably (no)	1.638	.160	.000	5.144	3.758	7.042
Elapsed time - charge to charge outcome			.023			
> 6 weeks ≤ 3 months (≤ 6 weeks)	.351	.116	.003	1.420	1.131	1.784
> 3 months ≤ 6 months (≤ 6 weeks)	.212	.117	.071	1.236	.982	1.556
> 6 months(≤ 6 weeks)	.161	.131	.217	1.175	.909	1.519
Prior convictions for serious dishonesty including burglary			.016			
One or two (none)	.296	.133	.026	1.345	1.037	1.744
Three to nine (none)	.447	.156	.004	1.563	1.151	2.122
Ten or more (none)	.613	.206	.003	1.846	1.233	2.764
Conviction for burglary last 5 years (no)	.838	.107	.000	2.313	1.876	2.850
Prior conviction for serious violence (no)	.595	.103	.000	1.813	1.480	2.221
Breach of sentence last 5 years (no)	.434	.094	.000	1.543	1.284	1.856
Prior convictions for any offence			.000			
1 - 4 (none)	.595	.172	.001	1.812	1.294	2.539
5 - 9 (none)	1.169	.191	.000	3.218	2.215	4.676
10 - 19 (none)	1.354	.212	.000	3.872	2.557	5.865
20 - 49 (none)	1.671	.241	.000	5.317	3.315	8.528
50 or more (none)	1.835	.291	.000	6.266	3.542	11.085
Longest prior term of incarceration for burglary			.000			
≤12 months (none)	.529	.138	.000	1.697	1.295	2.224
>12 months ≤24 months (none)	.732	.177	.000	2.079	1.469	2.941
> 24 months (none)	1.509	.251	.000	4.524	2.768	7.395
Female (male)	-.529	.144	.000	.589	.444	.782
Ethnicity			.015			
Māori (European)	.274	.089	.002	1.315	1.104	1.567
Pacific (European)	.323	.165	.049	1.382	1.001	1.908
Other or unknown (European)	.075	.234	.749	1.078	.681	1.706
Age at time of offending			.285			
Aged 20-29 (13-19)	.033	.102	.744	1.034	.846	1.263
Aged 30-39 (13-19)	.030	.147	.839	1.030	.773	1.374
Aged 40 or older (13-19)	-.282	.188	.135	.755	.522	1.091
Block of 17 circuits (Table 7.22)						

$\Delta\chi^2 = 3554.61$, $df = 53$ $p < .001$; R^2 Nagelkerke .620

The offence seriousness variables had the expected effects with one partial exception. The partial exception was that offenders in the two categories whose cases were not sentenced until more than three months after charges were laid were not significantly more likely to be incarcerated than those who were sentenced within six weeks of charge, whereas those sentenced between 6 weeks and 3 months were more likely to be incarcerated. This was unexpected but was explained by a higher proportion of older cases in the circuits with low odds ratios. The hierarchical generalised linear model for burglary reported in Chapter 9 confirmed there was a significant decrease in the likelihood of incarceration in the circuits where it took longer for cases to progress from charge to charge outcome (the results can be seen in Table 9.7).

Otherwise the likelihood of incarceration increased as expected if the offender was sentenced for multiple burglaries, had a concurrent conviction for a serious offence, had a concurrent conviction for breach of sentence, committed a burglary involving property with a higher nominal value, faced more than one charge, pleaded not guilty, or faced an indictably laid charge. As expected the likelihood of incarceration reduced if the offender was ordered to pay reparation.

Collectively the offence seriousness and case processing variables accounted for 47.3% of the improvement in prediction of incarceration compared to the baseline model when only the constant was included ($\Delta\chi^2=1683.02$, $df=16$, $p <.001$).

The effects associated with the criminal history variables were as expected. The likelihood of incarceration increased as the number of prior convictions for serious dishonesty increased; the total number of prior convictions for any offence increased; if the offender had a prior conviction for serious violence; if the most recent conviction for burglary was within five years of the current sentencing; as the length of any prior incarceration for burglary increased; or if the offender had a recent conviction for breaching a sentence.

Collectively the criminal history variables accounted for 44.5% of the improvement in prediction of incarceration compared to the baseline model when only the constant was included, and significantly improved the fit of the model ($\Delta\chi^2= 1583.22$, $df=14$, $p <.001$).

Gender and ethnicity both had statistically significant influence on the likelihood of incarceration. The odds of incarceration were significantly lower for females (odds ratio = 0.589) and when converted to relative risk equated to males being 1.4 times more likely to be incarcerated than similarly situated females. Both Māori (odds ratio = 1.315) and Pacific (odds ratio = 1.382) offenders were more likely to be incarcerated than Europeans.

When expressed as a relative risk both Māori and Pacific offenders were approximately 1.2 times more likely to be incarcerated than similar European offenders. There was no statistically significant difference in the likelihood of incarceration by age. This may be a consequence of the very high proportion of younger offenders, with 39% of offenders aged 19 or younger. Collectively the demographic factors only made a small contribution (1.2%) to the improvement in prediction of incarceration compared to the baseline model when only the constant was included, but did significantly improve the fit of the overall model ($\Delta\chi^2 = 42.44$, $df = 7$, $p < .001$).

Table 6.22 summarises the main effects associated with the circuits. Overall the addition of the circuit variables made a moderate contribution (6.9%) to the improvement in prediction of incarceration compared to the baseline model when only the constant was included, and significantly improved the fit of the overall model ($\Delta\chi^2 = 245.94$, $df = 16$, $p < .001$).

Table 6.22: Logistic regression parameter estimates for the circuit dummy variables for burglary offenders sentenced in 2008-2009 ($n = 5,740$)

Circuit (Reference Circuit 0)	B	S.E.	Sig.	Exp(B)	95% C.I. for Exp(B)	
					Lower	Upper
Circuit			.000			
Circuit 1	0.976	0.18	.000	2.654	1.865	3.776
Circuit 2	1.209	0.258	.000	3.351	2.021	5.556
Circuit 3	1.38	0.191	.000	3.974	2.735	5.773
Circuit 4	1.786	0.3	.000	5.963	3.313	10.734
Circuit 5	1.064	0.184	.000	2.899	2.021	4.158
Circuit 6	1.862	0.229	.000	6.436	4.112	10.073
Circuit 7	1.207	0.281	.000	3.342	1.928	5.796
Circuit 8	1.716	0.244	.000	5.56	3.447	8.968
Circuit 9	-0.06	0.283	.832	0.942	0.54	1.641
Circuit 10	1.761	0.215	.000	5.817	3.817	8.864
Circuit 11	1.526	0.22	.000	4.599	2.986	7.084
Circuit 12	1.666	0.195	.000	5.29	3.61	7.751
Circuit 13	0.362	0.259	.162	0.696	0.419	1.157
Circuit 14	1.4	0.308	.000	4.055	2.218	7.412
Circuit 15	0.018	0.207	.929	1.019	0.678	1.53
Circuit 16	1.528	0.232	.000	4.61	2.926	7.264

The odds of incarceration were significantly higher than the reference circuit in 13 of the 16 circuits. No circuit was significantly lower and the result was non-significant for three circuits. The odds ratios for the 13 circuits with higher odds of incarceration varied

from 2.654 (Circuit 1) to 6.436 (Circuit 6). The odds ratios fell into two distinct clusters. The spectrum of odds ratios and associated 95% confidence intervals is portrayed graphically in Figure 6.5. The circuits form into two clusters: the reference circuit plus the three circuits (9, 13, and 15) with odds ratios that are not significantly different to the reference circuit, and the remaining 13 circuits with significantly higher odds ratios.

The confidence intervals provide a guide to how the results would fall if a different circuit had been chosen as the reference circuit. For instance, if Circuit 6 had been chosen 5 circuits would have had significantly lower odds ratios and there would have been no significant difference between Circuit 6 and the other 11 circuits. If the circuit at the centre of the spectrum before controls (Circuit 12) had been chosen, 4 circuits would have had significantly lower odds ratios and there would have been no significant difference between Circuit 12 and the other 12 circuits. However, based on the confidence intervals it is apparent that some circuits differed significantly from one another, irrespective of the choice of the reference circuit or how they compared to that circuit.

Figure 6.5: Odds ratios for incarceration of burglary offenders sentenced in 2008-2009 by circuit (Circuit 0 is the reference)

The relative risks of incarceration for burglary by circuit in Figure 6.6 put the risk of incarceration into proportion. The higher odds of incarceration in Circuit 6 (odds 6.44 times greater than Circuit 0) translate to offenders being 2.49 times more likely to be incarcerated.

Figure 6.6: Relative risk of incarceration for burglary offenders sentenced in 2008-2009 by circuit

Prediction 1 is supported for burglary: after controlling for seriousness of offence, criminal history and offender demographic variables, similarly situated offenders were sentenced differently between the circuits. The difference in the likelihood of incarceration was substantial as well as statistically significant.

Residual Statistics, Influential Cases and Outliers

The discussion of residual statistics, influential cases and outliers has been combined into a single section for all three models because there is considerable overlap in the discussion between the three models. The residual statistics were reviewed for each model to identify influential cases and potential outliers. The values for Cook's distance were reviewed to identify any cases with an undue influence on the model and the standardised DFBeta measures were reviewed to identify any cases exerting an undue influence on any of the parameter estimates for the variables in the model. Standardised residuals were scrutinised to identify and examine potential outliers in the models with particular attention paid to residuals with an absolute value in excess of 3.

For all three models the values for Cook's distance and the standardised DFBeta measures were below 1. There were no cases exerting an undue influence on the model or on any of the parameter estimates for the variables in the models.

Examination of the standardised residuals raised concern about potential outliers in all three models. In all three models the proportion of standardised residuals with an absolute value greater than 1.96 (one standard deviation) was below 5% but more than 1% exceeded an absolute value of 2.58 (two standard deviations). The cases with an absolute value of 3 or more were scrutinised as potential outliers. The cases can be divided into two categories. First, cases where the model strongly predicted incarceration, with probabilities in excess of .9, but the offender was not incarcerated. Second, cases where the model strongly predicted a sentence other than incarceration, with probabilities less than .1, but the offender was incarcerated.

There were 142 cases (1.5% of total cases) with an absolute value of 3 or more in the model for aggravated drink driving. There were 37 cases of offenders who were not incarcerated despite the model strongly predicting they would be; and 105 cases of offenders who were incarcerated despite the model strongly predicting they would not be.

The male assaults female model contained 112 cases (1.8% of total cases) with an absolute value of 3 or more. There were 6 cases of offenders who were not incarcerated despite the model strongly predicting they would be; and 106 cases of offenders who were incarcerated despite the model strongly predicting they would not be.

The burglary model contained 99 cases (1.7% of total cases) with an absolute value of 3 or more. There were 41 cases of offenders who were not incarcerated despite the model strongly predicting they would be; and 58 cases of offenders who were incarcerated despite the model strongly predicting they would not be.

Cases are likely to have high standardised residuals because the sentence was largely influenced by one or more variables omitted from the models. There are a number of known variables that were omitted which may explain the outcomes for these cases. Many of these are generic variables applicable to all three models. Judges who did not incarcerate an offender who otherwise had a high modelled probability of incarceration may have been influenced by a judgement that the offender was committed to change, or that a sentence of incarceration would have a disproportionate impact causing severe hardship to either the offender or the offender's family.

Judges who incarcerated an offender who otherwise had a low modelled probability of incarceration may have been influenced by the fact that the offender was already a serving prisoner, the offender was being sentenced while in custody on a more serious charge, or the offender had already spent time remanded in custody. In all three situations judges may consider that there is no viable alternative sentence to incarceration.

The male assaults female and burglary cases were spread across the circuits with no apparent patterns. Given that Cook's distance and the standardised DFBeta measures did not indicate that any of the cases exerted an undue influence on the model or any of the model parameters the cases were retained in the model.

The aggravated drink driving cases were also spread across the circuits, but there were two notable concentrations. About two-thirds of the cases of no incarceration when it was strongly indicated occurred in circuits where sentencing was more severe. Likewise about two-thirds of the cases of incarceration when no incarceration was strongly indicated were in the circuits where sentencing was less severe. Excluding the cases with high standardised residuals from the model would have increased the degree of disparity. It is likely that for both groupings with high residuals the identity of the circuit was important. Assuming offenders incarcerated due to being sentenced prisoners, having spent time on remand or being in remand due to more serious charges yet to be resolved are relatively evenly spread across circuits they will constitute a higher proportion of incarcerated offenders in low incarceration circuits. The same proposition is equally likely to apply in the reverse for offenders not incarcerated due to future prospects of "going straight", they will account for a higher proportion of not-incarcerated offenders in high incarceration circuits. In these circumstances removing these cases could be criticised as "cooking the books" to increase the apparent level of disparity. Given that Cook's distance and the standardised DFBeta measures did not indicate that any of the cases exerted an undue influence on the model or any of the model parameters the cases were retained in the model.

Discussion of the Model Results

With the exception of a related article (Goodall & Durrant, 2013) which reported the results from a substantively similar version of the aggravated drink driving in-out model, albeit with a different reference circuit, these results are understood to be the first examining the influence of location on the use of incarceration in New Zealand employing statistical controls for a broad range of offence and offender variables.

Sentencing does vary by circuit

The model results strongly support the predicted effect of location: the circuit in which an offender was sentenced for aggravated drink driving, male assaults female or burglary significantly affected the likelihood of incarceration. After controlling for seriousness of offence, criminal history and offender demographic variables, similar offenders were sentenced differently between circuits.

For aggravated drink driving the odds of incarceration for almost all circuits (15/16) were statistically significantly different from the reference circuit. The odds ratios for the 14 circuits more likely to incarcerate ranged from 1.73 to 58.73, which when converted to relative risk means that depending on the circuit offenders were 1.6 to 7.4 times more likely to be incarcerated than offenders in the reference circuit.

For male assaults female the odds of incarceration were statistically significantly different for 7 of the 16 circuits relative to the reference circuit. The odds ratios for the 6 circuits more likely to incarcerate ranged from 1.85 to 2.85, which when converted to relative risk means that depending on the circuit offenders were 1.7 to 2.3 times more likely to be incarcerated than offenders in the reference circuit.

For burglary the odds of incarceration for 13 of the 16 circuits were statistically significantly different from the reference circuit. The odds ratios for the 13 circuits ranged from 2.65 to 6.44, which when converted to relative risk means that depending on the circuit offenders were 1.8 to 2.5 times more likely to be incarcerated than offenders in the reference circuit.

Contrary to the proposition made by Priestley (2009) there is no evidence that compositional differences in caseload explain the variations between locations observed by the Law Commission.

The disparity was substantial as well as statistically significant

The differences were substantial: they were not just statistically significant differences arising from a large number of cases, allowing relatively small substantive differences to attain statistical significance. The relative risks for the 14 circuits with a significant and greater risk of incarceration for aggravated drink driving ranged from 1.6 to 7.4 times greater with the risk more than five times greater in seven circuits. The relative risk of incarceration in the reference circuit was 1.6 times greater than circuit 15. The relative risks for the 6 circuits with a significant and greater risk of incarceration for male assaults female ranged from 1.9 to 2.9. The relative risk of incarceration in the reference circuit

was 1.6 times greater than circuit 13. The relative risks for the 13 circuits with a significant and greater risk of incarceration for burglary ranged from 1.8 to 2.5. The relative risks cannot be directly compared between the offences because the base rate of incarceration is different for each offence.

The substance of these differences can be measured in a more direct and practical way by estimating what the change in the number of offenders incarcerated would be if the offenders sentenced in the circuits with the highest odds ratios were transferred to the reference circuit. If the aggravated drink drivers from Circuit 8 were sentenced in Circuit 0, the difference in relative risk indicates that the number incarcerated would reduce from 140 to 19 (or 56 to 7.5 per 100). If the male assaults female offenders from Circuit 7 were sentenced in Circuit 0, the difference in relative risk indicates that the number incarcerated would reduce from 20 to 9 (or 17 to 7.6 per 100). If the burglars from Circuit 6 were sentenced in Circuit 0, the difference in relative risk indicates that the number incarcerated would reduce from 153 to 61 (or 55 to 22 per 100).

There are statistically significant and substantial differences in the rates of incarceration between circuits. The differences originally observed by the Law Commission (2006) cannot be explained by compositional differences in caseload between the circuits. These findings suggest that the current guidance is inadequate and fails to ensure that similar offenders who commit similar offences receive similar sentences.

Influence of seriousness of offence and criminal history variables (legal variables)

The models for all three offences indicate that offence seriousness and criminal history variables are the most significant influences on sentencing. Offence seriousness and criminal history variables made a 73.6% (aggravated drink driving), 95.5% (male assaults female) and 91.8% (burglary) contribution to the improvement in prediction of incarceration compared to the baseline model when only the constant was included. The results supported the assumption that, consistent with prior research, the seriousness of offence and criminal history variables would be the strongest influences on the decision to incarcerate. The importance of the seriousness of the offence and criminal history factors suggests the judges apply a common framework of factors when sentencing. However, the variation between circuits that remained after controlling for the seriousness and criminal history factors suggests there are differences in the threshold for incarceration between circuits for each of the three offences. There is little doubt that formally rational sentencing variables based on the Sentencing Act and appellate authorities dominate

sentencing decisions, especially for male assaults female and burglary. However, the results also suggest that the interpretation and use of the formally rational factors is influenced by substantively rational sentencing concerns and that the influence of these concerns varies between circuits. Based on the range and extent of differences in the sentencing of aggravated drink driving between circuits the influence of substantively rational concerns are greater for that offence. This may reflect differing views on the seriousness of the offence by judges, by the local community or by both. The multi-level random co-efficient models reported and discussed in Chapter 8 will shed more light on these propositions. The random co-efficient models will reveal whether there are differences in use for all the variables, just a selection or possibly for none of them.

Influence of gender and ethnicity (extra-legal) variables

Neither gender nor ethnicity had a statistically significant effect on the sentencing of aggravated drink driving or male assaults female. Both gender and ethnicity were significant influences on the sentencing of burglars. These differences underscore the usefulness of disaggregating the modelling by offence type.

The finding that gender was not statistically significant for aggravated drink driving adds an important qualification to the general findings in sentencing research in the United States (Ulmer, 2012) and in New Zealand (Jeffries et al., 2003; Triggs, 1999) which has almost universally found that females are sentenced more leniently. The finding that this does not apply to aggravated drink driving suggests the effect may be more nuanced. One possibility is that aggravated drink driving differs from the general run of offences, with the requirement that the offender must have two prior adult drink driving convictions possibly offsetting any sentencing leniency shown to females.

The finding that ethnicity had no significant effects on sentencing for aggravated drink driving and male assaults female but a significant effect for burglary adds to the mixed results in prior Australian and New Zealand research on indigeneity (Bond & Jefferies 2011; Snowball & Weatherburn, 2006; and Triggs, 1999 found no differences but Bond et al., 2011 and Jeffries and Bond, 2009 did find differences) but it is out of step with the overall weight of the United States research (Mitchell & Mackenzie, 2004; Ulmer, 2012). The effect of ethnicity on burglary adds a new perspective to Trigg's (1999) finding of no effect in New Zealand. It suggests that generalised research carries a risk of obscuring important differences via an averaging of effects.

Rank order of the circuits within and between offences

The position of each circuit on the spectrum from most lenient to most severe can be compared between the offences. Similar rankings would suggest that the differences are potentially associated with an overall sentencing severity policy or approach; conversely little similarity would suggest that policies or approaches are more offence specific. The order of circuits for each offence is portrayed in Table 6.23.

Six of the circuits were significantly more likely to incarcerate than the reference circuit for all three offences, but this did not include either of the two circuits most likely to incarcerate for aggravated drink driving (Circuit 8) or burglary (Circuit 6) with neither circuit significantly more likely to incarcerate male assaults female offenders. Two circuits (Circuits 13 & 15) were never significantly more likely to incarcerate than the reference circuit, with both of these circuits significantly less likely to incarcerate than the reference circuit for one offence each.

Table 6.23: Order of circuits by odds ratio from highest to lowest by offence type¹

Aggravated drink driving	Male assaults female	Burglary
8	7	6
12	14	4
6	1	10
4	10	8
2	2	12
7	12	16
10	4	11
1	5	14
16	3	3
14	9	2
11	16	7
3	8	5
5	15	1
9	6	15
0	0	0
13	11	9
15	13	13

¹ The circuits that were statistically significantly different to the reference circuit are shown in bold.

Spearman's rho (correlation co-efficient) was calculated for each of the three pairs of offences to determine whether there was any apparent link between the rank orders of the circuits for each offence. There was a link between the order of the circuits for

aggravated drink driving and burglary, $r_s = .79$, $p < .001$. However, there was no link between aggravated drink driving and male assaults female ($r_s = .40$, $p = .107$) or male assaults female and burglary ($r_s = .17$, $p = .523$).

The significant co-efficient for the pairing of aggravated drink driving and burglary points to the possibility that the sentencing of these offences is influenced by a general theory about the use of incarceration in each circuit. However, the result was not replicated for the pairing of male assaults female with either of these offences suggesting that the decision to incarcerate may be influenced more by a view of the relative seriousness of each offence type in each circuit rather than by some general theory about the use of incarceration. While it is not possible to draw any firm conclusions, a scan of the rank order of circuits across the offences does suggest the possibility that there was a general reluctance to incarcerate in the reference circuit and Circuits 13 and 15.

Summary

In sum, the logistic regression models support Prediction 1: similarly situated offenders were sentenced differently depending on which circuit they were sentenced in. The extent of inconsistency and the pattern varied between the offences. The strength of the findings also varied between the offences. The odds of incarceration for aggravated drink driving were spread across a continuum with many of the circuits varying from another. Conversely the burglary model found that the circuits fell into two broad groupings, more lenient and more severe circuits, with a substantial difference between the two. Although the circuit was significant the main determinants of sentence were offence seriousness and criminal history variables. Comparison of the rank order of the circuits by offence resulted in contradictory findings suggesting that sentencing in some circuits may be influenced by a general theory minimising the use of incarceration whereas in other circuits sentencing may be based on a view of the seriousness of the specific offence.

Chapter 7: Consistency between Circuits in the Length of Incarceration

This chapter reports descriptive statistics and the results of the sentence length models for each of the three offence types. The material is organised by offence with descriptive statistics and commentary followed by the sentence length linear regression models and discussion of the models. The discussion of residual statistics has been consolidated into a single section to minimise duplication.

Inconsistency in the length of incarceration between circuits was demonstrated for each of the offences. The position of the circuits on the spectrum from most lenient to most severe varied by offence. A significant difference for a circuit in relation to one offence did not predict significant differences for the other offences. Fewer circuits were significantly different for aggravated drink driving (7) and male assaults female (6) than burglary (14). All of the circuits that were statistically significantly different from the reference circuit for burglary and aggravated drink driving were substantially different from the reference circuit whereas only two of the six circuits for male assaults female were substantially different.

The number of circuits that were statistically significantly different on sentence length was similar to the number on the likelihood of incarceration for burglary (13 c.f. 14) and male assaults female (6 c.f. 7) but different for aggravated drink driving (7 c.f. 15). These distinctions validate the decision to model the sentence decision in two separate parts.

Offence seriousness and criminal history variables were found to be the principal influence on the likelihood of incarceration for all three offences. This finding is consistent with the findings of almost all sentencing research. In conjunction with the finding of inconsistency this points to differences in the interpretation and use of the offence seriousness and criminal history factors across circuits in the determination of sentence length as well as the choice of sentence type. If that is so, it supports the proposition that sentencing decisions are the result of a balance between formal rational sentencing factors and the influence of substantive rational sentencing concerns that varies between the circuits.

The review of residuals for the sentence length models for aggravated drink driving and male assaults female raised a significant issue. Almost all the observations reviewed as potential outliers had one of two characteristics: the incarceration sentence type was home detention (rather than imprisonment); or there were multiple offences and the

sentences were imposed cumulatively (rather than concurrently). This was a surprising discovery, raising the possibility that variables for these two characteristics should be included in the models. In the end they were not included, despite improving the performance of the models, because they are aspects of the outcome variable and their inclusion would have meant that the outcome was to some extent being used to predict the outcome. A description of the investigation and consideration of these characteristics is addressed in the consolidated discussion of the residuals with a more detailed discussion in Appendix F.

Aggravated Drink Driving: Descriptive Statistics

Of the 9,282 offenders sentenced by resident judges for aggravated drink driving, 2,452 (26.4%) were incarcerated. Sentence lengths were available for 2,312 (94.3%) of the incarcerated offenders. The mean sentence length for the 2,312 offenders for whom sentence length was available was 9.28 months, with a standard deviation of 5.10 months and a range from 0.5 months to 42 months.

The demographic profile of offenders incarcerated for aggravated drink driving

The demographic profile of incarcerated offenders is set out in Table 7.1.

Table 7.1: Demographic profile of offenders incarcerated for aggravated drink driving in 2008-2009 (*n* = 2,312)

Continuous Variable	Mean	SD	Min	Max
Age at the time of offending	37.87	10.55	17	75
Categorical Variables	N		%	
Gender				
Male	2023		87.5%	
Female	289		12.5%	
Ethnicity				
European	901		39.0%	
Māori	1184		51.2%	
Pacific	150		6.5%	
Other or unknown	77		3.3%	

Incarcerated aggravated drink drivers were: younger than those not incarcerated (37.9 c.f. 40.4 years); about as likely to be male than those not incarcerated (87.5% c.f. 86%); and more likely to be Māori than those not incarcerated (51.2% c.f. 38.1%).

Offenders incarcerated by circuit

The number of offenders incarcerated per circuit (Table 7.2) ranged from 27 to 357.

Table 7.2: Number of offenders incarcerated for aggravated drink driving in 2008-2009 by circuit ($n = 2,312$)

Categorical Variables	N	%
Circuit 0	95	4.1%
Circuit 1	311	13.5%
Circuit 2	80	3.5%
Circuit 3	134	5.8%
Circuit 4	46	2.0%
Circuit 5	186	8.0%
Circuit 6	249	10.8%
Circuit 7	79	3.4%
Circuit 8	138	6.0%
Circuit 9	46	2.0%
Circuit 10	138	6.0%
Circuit 11	136	5.9%
Circuit 12	357	15.4%
Circuit 13	27	1.2%
Circuit 14	60	2.6%
Circuit 15	43	1.9%
Circuit 16	187	8.1%

Offence seriousness and case processing variables

Most incarcerated offenders were sentenced for more than one offence (69.4%), which was substantially higher than non-incarcerated offenders (30.4%). The most common accompanying offence for offenders sentenced for more than one offence was driving while disqualified. On average incarcerated aggravated drink drivers were twice the legal limit, slightly higher than those not incarcerated (1.88 times). The proportion of incarcerated offenders sentenced after being recorded as pleading guilty was similar to those not incarcerated (98.2% c.f. 97.4%). Descriptive statistics for offence seriousness factors for incarcerated offenders are set out in Table 7.3.

Table 7.3: Descriptive statistics for offence seriousness variables linked to offenders incarcerated for aggravated drink driving in 2008-2009 ($n = 2,312$)

Continuous Variables	Mean	SD	Min	Max
Standardised alcohol level	2.03	0.54	1.0	4.4
Total charges sentenced)	2.58	2.08	1	34
Categorical Variables	N			%
Multiple drink drive charges				
No	2027			87.7%
Yes	285			12.3%
Concurrent disqualified driving charges				
No	1464			63.3%
Yes	848			36.7%
Concurrent serious driving charges				
No	2076			89.8%
Yes	236			10.2%
Concurrent minor driving charges				
No	1840			79.6%
Yes	472			20.4%
Concurrent breach of bail charges				
No	2107			91.1%
Yes	205			8.9%
Concurrent breach of sentence charges				
No	2069			89.5%
Yes	243			10.5%
Plea				
Guilty	2270			98.2%
Not guilty	42			1.8%

The mean time from charge to sentencing for incarcerated offenders was longer than for those not incarcerated (125 c.f. 100 days). The proportions of incarcerated offenders sentenced by criminal or family judges were similar to non-incarcerated offenders. Descriptive statistics for case processing factors for incarcerated offenders are set out in Table 7.4.

Table 7.4: Descriptive statistics for case processing variables for offenders incarcerated for aggravated drink driving in 2008-2009 ($n = 2,312$)

Continuous Variables	Mean	SD	Min	Max
Time from charge to charge outcome	125	85	1	366
Categorical Variables	N			%
Judge type				
Criminal	2038			88.1%
Family	274			11.9%

Criminal history variables

Offenders' criminal histories varied widely. The least prolific only had the minimum two prior convictions for drink driving necessary to be an aggravated drink driver and no other convictions, whereas the most prolific offender had 368 convictions in total for all offences. When the criminal history is restricted to drinking and driving the most prolific offender had 22 prior convictions. The ranges of the number of prior convictions are similar to those for offenders who were not incarcerated (2-21 for drink driving convictions and 2-327 for total convictions). However, incarcerated drink drivers had a higher mean number of prior drink driving convictions than those not incarcerated (4.70 c.f. 2.96). Likewise, incarcerated drink drivers had a substantially higher mean number of total prior convictions than those not incarcerated (30.25 c.f. 15.77). Over half of those not incarcerated had the minimum two prior convictions for drink driving, whereas only 18.6% of those incarcerated had only two prior convictions for drink driving. On average the immediate prior drink drive conviction for those incarcerated was 3.74 years ago, compared to 6.95 years for those not incarcerated. Incarcerated offenders were almost four times more likely to have been incarcerated before for drink driving than non-incarcerated offenders (50.7% c.f. 13.6%). The mean longest prior incarceration for drink driving for incarcerated offenders was longer than for offenders who were not incarcerated (6.63 months c.f. 4.36 months). Descriptive statistics for criminal history factors for incarcerated offenders are set out in Table 7.5.

Table 7.5: Descriptive statistics for criminal history variables linked to offenders incarcerated for aggravated drink driving in 2008-2009 (*n* = 2312)

Continuous Variables	Mean	SD	Min	Max
Prior convictions for drink driving	4.70	2.72	2	22
Years since last drink driving conviction	3.74	2.96	0.1	25
Longest prior incarceration for drink driving	6.63	4.83	0.5	24
Prior convictions for any offence	30.25	30.12	2	368
Categorical Variables	N		%	
Prior incarceration for drink driving	1173		50.7%	
Prior incarceration for any offence	1511		65.4%	

Aggravated Drink Driving: Sentence Length Model

The final model comprised 31 variables: 3 criminal history variables, 7 seriousness or case processing variables, 5 offender demographic variables and 16 circuit variables. Many of these were dummy variables (e.g., ethnicity and circuit). A range of variables included in the initial modelling were not retained because they were not statistically significant. The excluded variables comprised dummy variables for the number of prior drink driving convictions in the previous ten years; prior convictions for breach of sentence; prior incarceration for drink driving; whether the current case included a serious driving charge; whether the offender pleaded guilty; whether the current case included a minor driving charge; prior convictions for any offence; prior convictions for disqualified driving; and whether the offender was sentenced by a criminal or a family judge. The first five of the variables that were excluded from the model were significant in the in-out model thus demonstrating that different groupings of variables were significant for each decision and validating the decision to treat sentencing as a two part process separating type from quantum.

The variance inflation factors for the variables included in the final models were re-checked for multi-collinearity. None of the variables approached the threshold level of 10: the variance inflation factors ranged from 1.052 to 4.160.

The linear regression parameter estimates for the offence seriousness, criminal history and demographic variables are set out in Table 7.6. These parameter estimates incorporate the influence of the circuits but the parameter estimates for the circuits are reported separately.

The criminal history variables had the most substantial statistically significant influence on sentence length ($R^2 = .234$). As expected two of the variables had an aggravating effect and one had a mitigating effect. Each prior conviction for drinking and driving contributed approximately 2 weeks to the sentence (about 9.5 weeks at the average of 4.7 prior convictions). The longest prior term of incarceration for drink driving increased the sentence; it provided an additional measure of the seriousness of the prior offences. Each additional month added about 9 days to the sentence (about 8 weeks at the average longest prior term of 6.6 months). The impact of the prior drink driving convictions and incarceration was reduced if more time had elapsed since the immediate prior conviction; each additional year reduced the sentence by about 4 days (about 2 weeks at the average time since the last conviction of 3.7 years).

Table 7.6: Linear regression parameter estimates for the criminal history, offence seriousness, case processing and demographic variables for the sentence length model for aggravated drink drivers sentenced in 2008-2009 ($n = 2,312$)

Variable	B	S.E.	Sig.
(Constant)	1.769	.648	.006
Prior convictions for drink driving	.473	.047	.000
Years since the last drink drive conviction	-.131	.032	.000
Longest term of imprisonment for drink driving	.284	.024	.000
Alcohol levels standardised	1.135	.159	.000
Multiple drink drive charges	2.704	.292	.000
Case includes disqualified driving charge(s)	.503	.207	.015
Case includes breach(es) of bail	-1.498	.356	.000
Case includes breach(es) of sentence	-.692	.303	.023
Total number of charges for sentence	.501	.057	.000
Days from charge to charge outcome	.003	.001	.013
Female	-.176	.262	.501
Age when offended	-.010	.010	.323
Māori (cf European)	-.358	.191	.061
Pacific (cf European)	.022	.381	.953
Other (cf European)	-.468	.488	.338
Block of 16 circuit variables (Table 8.8)			

$R^2 = .234$ for Block 1 (Criminal history) $p < .001$; $\Delta R^2 = .109$ for Block 2 (Offence seriousness and case processing) $p < .001$, $\Delta R^2 = .002$ for Block 3 (Offender demographics) not significant; $\Delta R^2 = .041$ for Block 4 (Circuits) $p < .001$.

The seriousness variables had a statistically significant influence on sentence length ($\Delta R^2 = .109$). The seven seriousness variables were all aggravating factors expected to increase the sentence. This was only true for five of the variables. The length of the sentence increased by 34 days each time the ratio of the alcohol level to the legal limit increased by one unit. If the offender faced multiple drink driving convictions, sentence length increased by 81 days. If the offender faced a concurrent conviction for driving while disqualified, sentence length increased by 15 days. Sentence length increased as the total number of charges increased, with each additional charge adding 15 days to the sentence. Sentence length increased as the elapsed time from charge to charge outcome increased. Sentences were about 11 days longer for cases that took the average time (125 days) to reach sentencing.

The two seriousness factors that did not conform to expectations were the inclusion of breach of bail and breach of sentence offences. The model suggests that conviction for

a breach of bail reduced the sentence by 45 days and conviction for breach of sentence reduced the sentence by 21 days. A conviction for either offence must make the case more serious than the exact same case excluding either or both of these offences. The explanation for these unexpected effects must therefore lie in the construction of the model and the particular distribution of the modelled sentencing factors across the cases. The most likely explanation identified was that the correction occurs because these offences are associated with cases that have a disproportionately high average number of convictions for sentence but the breach offences are less serious than most other offences. A review of the data showed that both breach of bail and breach of sentence convictions did tend to be associated with cases that had a higher number of convictions for sentence. The 243 cases that included breach of sentence had an average of 4.9 convictions per case compared to 2.3 for those cases that did not include breach of sentence. The 205 cases that included breach of bail had an average of 5.8 convictions per case compared to 2.3 for those cases that did not include breach of bail. Because breaches of bail and breaches of sentence are often less serious than most other offences the model makes a negative correction to differentiate cases including breach of bail and/or breaches of sentence from similarly sized cases made up of more serious offences.

The combined contribution of the criminal history and offence seriousness variables to the overall model (a combined $R^2 = .342$ of the total model $R^2 = .385$) demonstrates that these variables are the principal influences on the length of incarceration.

The demographic variables did not have a statistically significant influence on sentence length ($\Delta R^2 = .002$). None of the variables achieved statistical significance.

Table 7.7 summarises the results for the circuit variables (a series of 16 dummy variables that contrast each of 16 circuits with the reference circuit (Circuit 0)). Overall the block of circuit variables had a statistically significant effect on the model ($\Delta R^2 = .041$). Sentence lengths were significantly different in 7 of the 16 circuits. The sentences ranged from about 6 weeks (Circuit 1) to almost 3 months longer (Circuit 16). In the context of this offence with a mean sentence of 9.1 months in the reference circuit, these are substantively significant differences.

Table 7.7: Linear regression parameter estimates for the circuit variables for the sentence length model for aggravated drink drivers sentenced in 2008-2009 ($n = 2,312$)

Variable	B	S.E.	Sig.
Circuit 1 v Circuit 0	1.510	.485	.002
Circuit 2 v Circuit 0	.951	.625	.128
Circuit 3 v Circuit 0	1.806	.547	.001
Circuit 4 v Circuit 0	1.069	.730	.143
Circuit 5 v Circuit 0	1.438	.521	.006
Circuit 6 v Circuit 0	-.802	.496	.106
Circuit 7 v Circuit 0	2.294	.621	.000
Circuit 8 v Circuit 0	.203	.550	.713
Circuit 9 v Circuit 0	1.163	.733	.113
Circuit 10 v Circuit 0	.330	.545	.544
Circuit 11 v Circuit 0	.952	.543	.080
Circuit 12 v Circuit 0	1.809	.473	.000
Circuit 13 v Circuit 0	-.921	.888	.300
Circuit 14 v Circuit 0	1.750	.670	.009
Circuit 15 v Circuit 0	-.065	.761	.932
Circuit 16 v Circuit 0	2.996	.515	.000

Prediction 2 is supported for aggravated drink driving: after controlling for criminal history, offence seriousness and offender demographic variables, similarly situated offenders were sentenced to different terms of incarceration in some circuits.

Male Assaults Female: Descriptive Statistics

Of the 6,166 offenders sentenced by resident judges for male assaults female, 744 (12.1%) were incarcerated. Incarceration sentence lengths were available for 723 or 97.2% of the 744 offenders incarcerated. The mean sentence length for the 723 offenders for whom sentence length was available was 7.44 months, with a standard deviation of 4.76 months and a range from 0.25 months to 35 months.

The demographic profile of offenders incarcerated for male assaults female

Incarcerated male assaults female offenders were: younger than those not incarcerated (30.9 c.f. 32.4 years); and more likely to be Māori than those not incarcerated (63.6% c.f. 43.4%). The demographic profile of incarcerated offenders is set out in Table 7.8.

Table 7.8: Demographic profile of offenders incarcerated for male assaults female in 2008-2009 (*n* = 723)

Continuous Variables	Mean	SD	Min	Max
Age at the time of offending	30.91	8.38	17	73
Categorical Variables	N		%	
Ethnicity				
European	168		23.2%	
Māori	460		63.6%	
Pacific	78		10.8%	
Other or unknown	17		2.4%	

Offenders incarcerated by circuit

The number of offenders incarcerated per circuit (Table 7.9) ranged from 9 to 82.

Table 7.9: Number of offenders incarcerated for male assaults female in 2008-2009 by circuit (*n* = 723)

Categorical Variables	N	%
Circuit 0	82	11.3%
Circuit 1	71	9.8%
Circuit 2	34	4.7%
Circuit 3	63	8.7%
Circuit 4	18	2.5%
Circuit 5	76	10.5%
Circuit 6	46	6.4%
Circuit 7	20	2.8%
Circuit 8	21	2.9%
Circuit 9	9	1.2%
Circuit 10	44	6.1%
Circuit 11	53	7.3%
Circuit 12	82	11.3%
Circuit 13	14	1.9%
Circuit 14	16	2.2%
Circuit 15	40	5.5%
Circuit 16	34	4.7%

Offence seriousness and case processing variables

Incarcerated offenders were substantially more likely to be sentenced for more than one offence than non-incarcerated offenders (77.3% c.f. 35.4%). The most common accompanying offences for offenders sentenced for more than one offence were another violent offence or a breach of a protection order. Although most offenders were sentenced for one male assaults female offence, the proportion of incarcerated offenders sentenced for two or more such offences was almost three times greater than those not incarcerated

(19.9% c.f. 7.0%). The proportion of incarcerated offenders sentenced after being recorded as pleading guilty was lower than those not incarcerated (90.5% c.f. 93.2%). Descriptive statistics for offence seriousness factors for incarcerated offenders are set out in Table 7.10.

Table 7.10: Descriptive statistics for offence seriousness variables linked to offenders incarcerated for male assaults female in 2008-2009 by circuit (*n* = 723)

Continuous Variables	Mean	SD	Min	Max
Total charges sentenced	3.50	2.71	1	19
Categorical Variables	N			%
Multiple male assaults female charges				
No	579			80.1%
Yes	144			19.9%
Concurrent breach of protection order charges				
No	542			75.0%
Yes	181			25.0%
Concurrent threat to kill do GBH charges				
No	703			97.2%
Yes	20			2.8%
Concurrent wilful or intentional damage charges				
No	600			83.0%
Yes	123			17.0%
Concurrent other violence charges				
No	526			72.8%
Yes	197			27.2%
Concurrent breach of bail charges				
No	630			87.1%
Yes	93			12.9%
Concurrent breach of sentence charges				
No	564			78.0%
Yes	159			22.0%
Weapon described in the charge				
No	712			98.5%
Yes	11			1.5%
Plea				
Guilty	654			90.5%
Not guilty	69			9.5%

The mean time from charge to sentencing for incarcerated offenders (114 days) was similar to those not incarcerated (116 days). The proportion of incarcerated offenders sentenced by judges with a Family Court warrant was lower than non-incarcerated offenders (16.3% c.f. 22.5%). A smaller proportion of incarcerated offenders were

sentenced in a District Court with an active family violence court than non-incarcerated offenders (24.9% c.f. 37.4%). Descriptive statistics for case processing factors for incarcerated offenders are set out in Table 7.11.

Table 7.11: Descriptive statistics for case processing variables linked to offenders incarcerated for male assaults female in 2008-2009 by circuit (*n* = 723)

Continuous Variables	Mean	SD	Min	Max
Time from charge to charge outcome	114	87	1	366
Categorical Variables	N		%	
Offender sentenced in a Family Violence Court				
No	543		75.1%	
Yes	180		24.9%	
Judge type				
Criminal	605		83.7%	
Family	118		16.3%	
Charge Laid Type				
Summary	687		95.0%	
Indictable	36		5.0%	

Criminal history variables

Offenders' criminal histories varied widely. The least prolific had no prior convictions for any offence. By contrast the most prolific offender for male assaults female had 11 prior convictions for that offence and the most prolific general offender had 231 prior convictions. These ranges were similar to those for offenders who were not incarcerated (0-8 for male assaults female convictions and 0-293 for total convictions). However, incarcerated male assaults female offenders had a higher mean number of prior convictions for male assaults female compared to those not incarcerated (1.49 c.f. 0.38). Likewise incarcerated male assaults female offenders had a substantially higher mean number of prior convictions for any offence than those not incarcerated (38.2 c.f. 12.1). Of those incarcerated 62.9% had a prior male assault female conviction compared to 24.0% of offenders who were not incarcerated. Incarcerated offenders were about seven times more likely to have been incarcerated before for male assaults female than non-incarcerated offenders (34.9% c.f. 4.9%). Incarcerated offenders were about three times more likely to have been incarcerated before for any offence than non-incarcerated offenders (76.5% c.f. 25.2%). The mean longest prior incarceration for any offence for incarcerated offenders was longer than for offenders who were not incarcerated (20.9 c.f. 16.6 months). Descriptive statistics for criminal history factors for incarcerated offenders are set out in Table 7.12.

Table 7.12: Descriptive statistics for criminal history variables linked to offenders incarcerated for male assaults female in 2008-2009 by circuit (*n* = 723)

Continuous Variables	Mean	SD	Min	Max
Prior convictions for male assaults female	1.49	1.71	0	11
Prior convictions for any offence	38.22	34.93	0	231
Longest prior incarceration for male asslts female	7.95	4.82	0.25	24
Longest prior incarceration for any offence	20.93	21.72	0.25	147
Categorical Variables	N		%	
Prior incarceration for male assaults female	252		34.9%	
Prior incarceration for any offence	553		76.5%	
Prior conviction for serious violence	376		52.0%	
In the previous five years the offender was:				
Sentenced for male assaults female	349		48.3%	
Sentenced for serious violence	222		30.7%	
Sentenced for breach of a protection order	175		24.2%	
Sentenced for breach of a sentence	422		58.4%	
Sentenced to incarceration	460		63.6%	

Male Assaults Female: Sentence Length Model

The final model comprised 27 variables: 4 seriousness or case processing variables; 3 criminal history variables; 4 offender demographic variables and 16 circuit variables. Many of these were dummy variables (e.g. ethnicity and circuit). A range of variables were included in the initial modelling but not retained because they were not statistically significant. The excluded variables comprised: whether the case included breach of a protection order; whether the case included other violence charges; whether the case included wilful or intentional damage charges; whether the case included breach of sentence charges; plea; the time from charge to charge outcome; the age of the most recent conviction for male assaults female; the longest prior term of imprisonment for male assaults female; total prior convictions for any offence; prior incarceration for any offence; the longest previous term of incarceration for any offence; and prior convictions for serious violence. Seven of the variables that were excluded from the model were significant in the in-out model thus demonstrating that different groupings of variables were significant for each decision and validating the decision to treat sentencing as a two part process separating type from quantum. The initial modelling included a variable representing whether or not the case included a charge or charges of threatening to kill or do GBH. The variable was statistically significant but the parameter estimate for it was unduly influenced by one case with a standardised DFBeta measure above one. When it was investigated it was found that it was the case with the longest sentence in the dataset (35

months). The influence of that one case on the parameter estimate for the variable was judged by running the model excluding the case; the regression co-efficient decreased from 2.674 to 1.664 and was not significant. The model was modified to exclude the case and the threat to kill or do grievous bodily harm variable was removed from the model. The change in the parameter estimate was a combination of the very high sentence for the case and low frequency of occurrence of charges of threatening to kill or do grievous bodily harm; such a charge was only present in 20 of the 723 cases.

The variance inflation factors for the variables included in the final model were re-checked for multi-collinearity. None of the variables approached the threshold level of 10: the variance inflation factors ranging from 1.056 to 1.846.

The linear regression parameter estimates for the offence seriousness, criminal history and demographic variables are displayed in Table 7.13. These parameter estimates incorporate the influence of the circuits but the parameter estimates for the circuits are reported separately.

The offence seriousness variables had the most substantial statistically significant influence on sentence length ($R^2 = .117$). All four variables were expected to increase the sentence but only three did. Sentence length increased by almost two months if the offender was being sentenced for multiple male assaults female convictions, by 14 days for each charge in the case, and by more than 2 months if the charge was laid indictably. An indictable charge indicated that the actual offending was more serious than the norm.

The seriousness factor that did not conform to expectations was sentencing for a breach of bail. The model suggested that conviction for a breach of bail reduced the sentence by 5 weeks. The negative effect was observed in the aggravated drink driving model. The same explanation applies. The breaches of bail were associated with cases that have a disproportionately high average number of convictions for sentence but breach of bail is less serious than most other offences. A review of the data showed that breach of bail convictions did tend to be associated with cases that had a higher number of convictions for sentence. The 93 cases that included breach of bail had an average of 6.4 convictions per case compared to 3.1 for those cases that did not include breach of bail. Although sometimes the circumstances of a breach can be serious, they are often not very serious. The model corrects for the lower seriousness of the mix of the charges in the case by reducing the sentence when a breach of bail offence is present.

Table 7.13: Linear regression parameter estimates for the criminal history, offence seriousness, case processing and demographic variables for the sentence length model for male assaults female offenders sentenced in 2008-2009 ($n = 722$)

Variable	B	S.E.	Sig.
(Constant)	3.162	.894	.000
Multiple male assaults female charges	1.989	.412	.000
Case includes breach(es) of bail	-1.172	.526	.026
Total number of charges for sentence	.478	.067	.000
Charge laid indictably (cf summarily)	2.387	.733	.001
Prior convictions for male assaults female	.269	.120	.025
Prior incarceration for male assaults female	1.387	.424	.001
Recent serious violence conviction (cf none or > 5 years ago)	1.122	.344	.001
Age when offended	-.016	.020	.408
Māori (cf European)	.399	.413	.334
Pacific (cf European)	.109	.308	.724
Other (cf European)	-.311	.357	.384
Block of 16 circuit variables (Table 8.17)			

$R^2 = .117$ for Block 1 (Offence seriousness and case processing) $p < .001$; $\Delta R^2 = .055$ for Block 2 (Criminal history) $p < .001$, $\Delta R^2 = .004$ for Block 3 (Offender demographics) not significant; $\Delta R^2 = .072$ for Block 4 (Circuits) $p < .001$.

The criminal history variables had a statistically significant influence on sentence length ($\Delta R^2 = .055$). All three variables had the expected effect. Each prior conviction for male assaults female contributed approximately 8 days to the sentence. Prior incarceration for male assaults female added about 6 weeks and a conviction for serious violence within the previous 5 years added about 5 weeks.

The combined contribution of the criminal history and offence seriousness variables to the overall model (a combined $R^2 = .172$ of the total model $R^2 = .247$) demonstrates that these variables are the principal influences on the length of incarceration.

The demographic variables did not have a statistically significant influence on sentence length ($\Delta R^2 = .004$). None of the variables achieved statistical significance.

Table 7.14 summarises the results for the circuit variables (a series of 16 dummy variables that contrast each of 16 circuits with the reference circuit (Circuit 0)).

Table 7.14: Linear regression parameter estimates for the circuit variables for the sentence length model for male assaults female offenders sentenced in 2008-2009 (n = 722)

Variable	B	S.E.	Sig.
Circuit 1 v Circuit 0	3.514	.689	.000
Circuit 2 v Circuit 0	1.119	.429	.009
Circuit 3 v Circuit 0	.834	.231	.000
Circuit 4 v Circuit 0	.308	.274	.262
Circuit 5 v Circuit 0	.346	.136	.011
Circuit 6 v Circuit 0	-.044	.129	.732
Circuit 7 v Circuit 0	.580	.149	.000
Circuit 8 v Circuit 0	.120	.128	.350
Circuit 9 v Circuit 0	.027	.163	.867
Circuit 10 v Circuit 0	.118	.078	.129
Circuit 11 v Circuit 0	.048	.067	.471
Circuit 12 v Circuit 0	-.034	.055	.534
Circuit 13 v Circuit 0	.091	.092	.325
Circuit 14 v Circuit 0	-.014	.081	.859
Circuit 15 v Circuit 0	.061	.054	.259
Circuit 16 v Circuit 0	.115	.053	.030

Sentence lengths were significantly different in 6 of the 16 circuits. The differences ranged from a minimal 3 days longer (Circuit 16) to 3.5 months longer (Circuit 1). Only two circuits sentenced to terms more than one month different from the reference circuit. In the context of this offence, with a mean sentence of 6.7 months in the reference circuit, these two circuits were substantively different. Overall the block of circuit variables had a statistically significant effect on the model ($\Delta R^2 = .072$).

Prediction 2 is supported for male assaults female: after controlling for criminal history, offence seriousness and offender demographic variables, similarly situated offenders were sentenced to different terms of incarceration in some circuits.

Burglary: Descriptive Statistics

Of the 5,740 offenders sentenced by resident judges for burglary, 2,434 (42.4%) were incarcerated. Sentence lengths were available for 2,338 (96.1%) of the incarcerated offenders. The mean sentence length for the 2,338 offenders for whom sentence length was available was 15.90 months, with a standard deviation of 11.24 months and a range from 0.25 months to 96 months.

The demographic profile of offenders incarcerated for burglary

Incarcerated burglars were: older than those not incarcerated (26.0 c.f. 22.8 years); more likely to be male than those not incarcerated (94.3% c.f. 88.1%); and more likely to be Māori than those not incarcerated (58.6% c.f. 44.8%). The demographic profile for the incarcerated offenders is set out in Table 7.15.

Table 7.15: Demographic profile of offenders incarcerated for burglary in 2008-2009 (*n* = 2,338)

Continuous Variables	Mean	SD	Min	Max
Age at the time of offending	26.04	8.12	15	71
Categorical Variables	N		%	
Gender				
Male	2204		94.3%	
Female	134		5.7%	
Ethnicity				
European	750		32.1%	
Māori	1370		58.6%	
Pacific	157		6.7%	
Other or unknown	61		2.6%	

Offenders incarcerated by circuit

The number of burglars sentenced per circuit (Table 7.16) ranged from 45 to 285.

Table 7.16: Number of offenders incarcerated for burglary in 2008-2009 by circuit (*n* = 2,338)

Categorical Variables	N	%
Circuit 0	150	6.4%
Circuit 1	285	12.2%
Circuit 2	79	3.4%
Circuit 3	243	10.4%
Circuit 4	45	1.9%
Circuit 5	279	11.9%
Circuit 6	149	6.4%
Circuit 7	57	2.4%
Circuit 8	80	3.4%
Circuit 9	46	2.0%
Circuit 10	146	6.2%
Circuit 11	133	5.7%
Circuit 12	203	8.7%
Circuit 13	67	2.9%
Circuit 14	57	2.4%
Circuit 15	197	8.4%
Circuit 16	122	5.2%

Offence seriousness and case processing variables

Incarcerated burglars were substantially more likely to be sentenced for more than one offence than those burglars not incarcerated (84.1% c.f. 56.5%). Incarcerated burglars were also more likely to have been sentenced for multiple burglaries than non-incarcerated burglars (40.1% c.f. 18.3%). The proportion of incarcerated burglars sentenced after being recorded as pleading guilty was slightly lower than those not incarcerated (95.2% c.f. 96.6%). Descriptive statistics for offence seriousness factors for incarcerated burglars are set out in Table 7.17.

Table 7.17: Descriptive statistics for offence seriousness variables linked to offenders incarcerated for burglary in 2008-2009 by circuit ($n = 2,338$)

Continuous Variables	Mean	SD	Min	Max
Total charges sentenced	6.30	6.65	1	113
Total burglary charges sentenced	2.35	3.47	1	63
Categorical Variables	N			%
Concurrent receiving charges				
No	1985			84.9%
Yes	353			15.1%
Concurrent unlawful taking charges				
No	2010			86.0%
Yes	328			14.0%
Concurrent serious driving charges				
No	1925			82.3%
Yes	413			17.7%
Concurrent breach of bail charges				
No	1860			79.6%
Yes	478			20.4%
Concurrent breach of sentence charges				
No	1683			72.0%
Yes	655			28.0%
Concurrent other serious offences (receiving, unlawful taking, violence, drugs & driving)				
No	1386			59.3%
Yes	952			40.7%
Approximate value of the lead burglary offence				
Less than \$500 or unstated	876			37.5%
\$500 to \$5,000	1096			46.9%
More than \$5,000	366			15.7%
Order for reparation				
No	1531			65.5%
Yes	807			34.5%
Burglary by day or night				
Day	1099			47.0%
Night	1069			45.7%
Unknown	170			7.3%
Burglary offence dates				
One	1518			64.9%
Two	415			17.8%
Three or more	405			17.3%
Plea				
Guilty	2237			95.7%
Not guilty	101			4.3%

The mean time from charge to sentencing for incarcerated burglars was longer than for those not incarcerated (133 c.f. 110 days). The proportion of incarcerated burglars sentenced by a judge with a Family Court warrant was lower than for burglars who were not incarcerated (9.1% c.f. 12.1%). A higher proportion of incarcerated burglars had been charged indictably than non-incarcerated burglars (18.9% c.f. 3.0%). Descriptive statistics for case processing factors for incarcerated burglars are set out in Table 7.18.

Table 7.18: Descriptive statistics for case processing variables linked to offenders incarcerated for burglary in 2008-2009 by circuit (n = 2,338)

Continuous Variables	Mean	SD	Min	Max
Time from charge to charge outcome	133	98	1	366
Categorical Variables	N		%	
Judge type				
Criminal		2125		90.9%
Family		213		9.1%
Charge Laid Type				
Summary		1896		81.1%
Indictable		442		18.9%

Criminal history variables

Offenders' criminal histories varied widely. The least prolific had no prior convictions for any offence. By contrast the most prolific burglar had 382 prior convictions for burglary and the most prolific general offender had 471 prior convictions. Incarcerated burglars had a substantially higher mean number of prior convictions for burglary than those not incarcerated (5.9 c.f. 1.1). Likewise, incarcerated burglars had a substantially higher mean number of prior convictions for any offence than those not incarcerated (37.4 c.f. 11.5). The majority (71.6%) of those incarcerated for burglary had at least one prior burglary conviction compared to 26.1% of burglars who were not incarcerated. Incarcerated burglars were more than five times more likely than non-incarcerated burglars to have been incarcerated before for burglary (54.5% c.f. 10.3%). Incarcerated burglars were about three and a half times more likely to have been incarcerated before for any offence than non-incarcerated burglars (71.8% c.f. 19.8%). Incarcerated burglars who had previously been incarcerated had higher average longest prior terms than those not incarcerated (21.7 c.f. 17.2 months). Descriptive statistics for the criminal history factors for incarcerated burglars are set out in Table 7.19.

Table 7.19: Descriptive statistics for criminal history variables linked to offenders incarcerated for burglary in 2008-2009 by circuit (*n* = 2,338)

Continuous Variables	Mean	SD	Min	Max
Prior convictions for burglary	5.92	12.95	0	382
Prior convictions for serious dishonesty excl burglary	9.32	19.04	0	382
Prior convictions for serious dishonesty incl burglary	15.24	25.9	0	396
Prior convictions for any offence	37.36	41.72	0	471
Longest prior incarceration for burglary (<i>n</i> =1274)	18.3	12.58	1	72
Longest prior incarceration for any offence (<i>n</i> =1678)	21.94	19.18	0.25	156
Categorical Variables	N		%	
Prior incarceration for burglary				
Never	1064		45.5%	
Once	497		21.3%	
Two or more occasions	777		33.2%	
Prior incarceration for any offence				
Never	660		28.2%	
Once	687		29.4%	
Two or more occasions	991		42.4%	

Burglary: Sentence Length Model

The final model comprised 50 variables: 15 seriousness or case processing variables, 14 criminal history variables, 5 offender demographic variables and 16 circuit variables. Many of these were dummy variables (e.g. ethnicity and circuit). A range of variables were included in the initial modelling but not retained because they were not statistically significant. The seriousness variables excluded were: whether the case included unlawful taking (theft of a vehicle); whether the case included dishonesty offending other than burglary, receiving, or unlawful taking; the time since the last conviction for burglary; and whether the burglary occurred during the day or at night. The criminal history variables that were omitted were: total prior convictions for any offence; prior incarceration for any offence; and the longest previous term of incarceration for any offence. The extent of criminal history information available was very wide and could be used in different ways. For instance, it was possible to deal with different types of serious dishonesty separately or to combine them, and to include or exclude burglary convictions from the mix of serious dishonesty offences. Other serious convictions for drug dealing or violence or other offending could also be disaggregated or aggregated together. Ultimately a set of dummy variables for prior convictions for burglary, a set of dummy variables for prior convictions for other serious dishonesty offending, and a variable for all other serious offending

(excluding dishonesty) were included in the model. Variables representing first offenders and prior incarceration for burglary were included in initial models but excluded because they had no significant effect. Any effect expected due to an offender having no prior convictions at all was potentially captured by the other prior conviction variables. Prior incarceration for burglary was indirectly incorporated via the variable for longest prior incarceration for burglary.

The variance inflation factors for the variables included in the final model were re-checked for evidence of multi-collinearity. None of the variables approached the threshold level of 10: with the variance inflation factors ranging from 1.059 to 2.978.

The linear regression parameter estimates for the offence seriousness, criminal history and demographic variables are set out in Table 7.20. These parameter estimates incorporate the influence of the circuits but the parameter estimates for the circuits are reported separately.

The offence seriousness variables had the most substantial statistically significant influence on sentence length ($R^2 = .343$). With two exceptions the variables were expected to increase sentence lengths. The exceptions were cases that included an order for reparation and whether the offender was sentenced by a criminal or a family judge. Orders for reparation are taken into account in determining the severity of sentence, and as expected, a reparation order reduced the sentence. Judge type was not expected to affect sentence length but was included as a potential explanatory variable. It was found to be significant: offenders sentenced by Family Court judges received sentences about 6 weeks shorter than those sentenced by criminal judges.

The remaining 13 variables all represented increasing levels of seriousness and were expected to increase sentence length. This expectation was true for 12 of the 13 variables. The 12 variables and their approximate effects on sentence length are described in Table 7.21.

Table 7.20: Linear regression parameter estimates for the criminal history, offence seriousness, case processing and demographic variables for the sentence length model for burglary offenders sentenced in 2008-2009 ($n = 2,338$)

Variable	B	S.E.	Sig.
(Constant)	-1.687	1.087	.121
Number of burglary charges for sentence	.713	.074	.000
\$500 to \$5,000 cf Less than \$500 or unknown	1.236	.366	.001
More than \$5,000 cf Less than \$500 or unknown	3.541	.502	.000
Count of offence dates for the burglaries	2.314	.259	.000
Case includes receiving	2.440	.490	.000
Case includes breach of bail	-1.633	.449	.000
Case includes breach of sentence	.888	.388	.022
Case includes violent offence	1.887	.416	.000
Case includes serious driving offence	1.136	.463	.014
Total number of charges for sentence	.328	.039	.000
Charge laid indictably or summarily	5.738	.463	.000
Plea	1.951	.822	.018
Days from charge to charge outcome	.005	.002	.005
Whether a reparation order was made	-.831	.370	.025
Categorisation of judges by warrant	-1.391	.594	.019
One cf no prior convictions for burglary	.497	.546	.364
Two cf no prior convictions for burglary	.837	.640	.191
3, 4 or 5 cf no prior convictions for burglary	2.569	.576	.000
6 to 9 cf no prior convictions for burglary	3.571	.728	.000
10 to 19 cf no prior convictions for burglary	5.355	.849	.000
20 or more cf no prior convictions for burglary	6.374	1.062	.000
Longest prior term of incarceration for burglary	.128	.022	.000
One cf no prior convictions for serious dishonesty excl burglary	.399	.601	.507
Two cf no prior convictions for serious dishonesty excl burglary	1.036	.646	.109
3, 4 or 5 cf no prior convictions for serious dishonesty excl burglary	2.693	.578	.000
6 to 9 cf no prior convictions for serious dishonesty excl burglary	2.672	.676	.000
10 to 19 cf no prior convictions for serious dishonesty excl burglary	2.240	.665	.001
20 or more cf no prior convictions for serious dishonesty excl burglary	2.957	.750	.000
Criminal history for serious violent, drugs or driving offences	.100	.034	.003
Female	-1.137	.723	.116
Age at the time of offence	-.014	.027	.604
Māori cf European	.359	.390	.358
Pacific cf European	.504	.743	.498
Other or unknown cf European	-.106	1.064	.921

$R^2 = .343$ for Block 1 (Offence seriousness and case processing) $p < .001$; $\Delta R^2 = .140$ for Block 2 (Criminal history) $p < .001$, $\Delta R^2 = .001$ for Block 3 (Offender demographics) not significant; $\Delta R^2 = .039$ for Block 4 (Circuits) $p < .001$.

Table 7.21: Approximate effects of offence seriousness and case processing variables for burglary found to increase sentence length

Variable	Approximate impact on sentence
Number of burglary charges for sentence	An additional 3 weeks for each burglary charge
\$500 to \$5,000 cf Less than \$500 or unknown	An additional 5 weeks if the main charge indicated an offence where the property taken was more than \$500 but less than \$5,000.
More than \$5,000 cf Less than \$500 or unknown	An additional 3.5 months if the main charge indicated an offence where the property taken was more than \$5,000.
Count of offence dates for the burglaries	An additional 10 weeks for each distinct date for the burglary offences being sentenced.
Case includes receiving	An additional 10 weeks if the case also included a receiving offence.
Case includes breach of sentence	An additional 3.5 weeks if the case also included a sentence breach offence.
Case includes violent offence	An additional 8 weeks if the case also included a violent offence.
Case includes serious driving offence	An additional month if the case also included a serious driving offence.
Total number of charges for sentence	An additional 9 days for each offence for sentence.
Charge laid indictably or summarily	An addition of almost 6 months if the main burglary large was laid indictably rather than summarily.
Plea	An addition of almost 2 months if the offender maintained a not guilty plea throughout.
Days from charge to charge outcome	An addition of time depending on how long the case took from charge to charge outcome. Compared to an offender sentenced 30 days after being charged an offender sentenced after 3 months would receive an additional 9 days, an offender sentenced after 6 months would receive an additional 22 days and an offender sentenced after one year would receive an additional 50 days.

The seriousness factor that did not conform to expectations was sentencing for a breach of bail. The model suggested that conviction for a breach of bail reduced the sentence by 7 weeks. The negative effect was also observed in the aggravated drink driving model. The same explanation applies. The breaches of bail were associated with cases that have a disproportionately high average number of convictions for sentence. A review of the data showed that breach of bail convictions did tend to be associated with cases that had a higher number of convictions for sentence. The 478 cases that included

breach of bail had an average of 9.6 convictions per case compared to 5.4 for those cases that did not include breach of bail. Because breach of bail is often less serious than most other offences the model makes a negative correction to differentiate cases including breach of bail from similarly sized cases made up of more serious offences.

The criminal history variables had a statistically significant influence on sentence length ($\Delta R^2 = .140$). All 14 variables had the expected effect. The number of variables was expanded by the inclusion of two sets of six dummy variables for prior convictions for burglary and prior convictions for other serious dishonesty offences (receiving, unlawful taking, fraud, deception and the most serious thefts). The effect on sentence length of having 1 or 2 prior burglary convictions was not significantly different from the effect of having no prior burglary convictions. However, offenders with 3, 4 or 5 prior convictions could expect an additional 2.5 months, with the additional amount rising to almost 6.5 months for those with 20 or more prior convictions for burglary. Likewise, although the effect on sentence length of 1 or 2 prior convictions for other serious dishonesty offences was not significantly different from no prior other serious dishonesty convictions, offenders with 3, 4 or 5 prior convictions could expect an additional 2.5 months. The effect of more than 5 previous convictions for other serious dishonesty offences was only marginal with an additional 3 months for those with 20 or more prior convictions for other serious dishonesty compared to those with no such prior convictions. The increases for those with 6 to 9 or 10 to 19 prior convictions was little different to those with 3, 4 or 5 prior convictions. If the offender had a wider criminal past encompassing serious violent, drug or driving offences, sentence length increased by 3 days for each prior conviction. The significance of the offender's criminal history for burglary increased if the offender had previously been incarcerated for burglary, with sentence length increasing by almost 4 days for each month of the longest prior term of incarceration for burglary, with offenders whose previous longest term matched the average previous longest term (almost 16 months) receiving an additional 2 months.

The combined contribution of the criminal history and offence seriousness variables to the overall model (a combined $R^2 = .482$ of the total model $R^2 = .523$) demonstrates that these variables are the principal influences on the length of incarceration.

The demographic variables did not have a statistically significant influence on sentence length ($\Delta R^2 = .001$). None of the variables achieved statistical significance, in contrast to the results from the in-out model. The more severe sentencing of males and

Māori and Pacific offenders reported in the in-out model was not replicated in the sentence length model.

Table 7.22 summarises the results for the circuit variables (a series of 16 dummy variables that contrast each of the 16 circuits with the reference circuit (Circuit 0)). Overall the block of circuit variables had a statistically significant effect on the model ($\Delta R^2 = .039$).

Table 7.22: Linear regression parameter estimates for the circuit variables for the sentence length model for burglary offenders sentenced in 2008-2009 ($n = 2,338$)

Variable	B	S.E.	Sig.
Circuit 1 v Circuit 0	6.315	.812	.000
Circuit 2 v Circuit 0	5.038	1.125	.000
Circuit 3 v Circuit 0	4.197	.832	.000
Circuit 4 v Circuit 0	5.132	1.362	.000
Circuit 5 v Circuit 0	4.768	.819	.000
Circuit 6 v Circuit 0	5.146	.921	.000
Circuit 7 v Circuit 0	3.594	1.255	.004
Circuit 8 v Circuit 0	3.118	1.109	.005
Circuit 9 v Circuit 0	-1.803	1.352	.183
Circuit 10 v Circuit 0	4.691	.929	.000
Circuit 11 v Circuit 0	5.788	.954	.000
Circuit 12 v Circuit 0	5.535	.871	.000
Circuit 13 v Circuit 0	-2.663	1.175	.024
Circuit 14 v Circuit 0	5.841	1.239	.000
Circuit 15 v Circuit 0	.056	.878	.949
Circuit 16 v Circuit 0	4.553	.972	.000

Sentence lengths were significantly different in 14 of the 16 circuits. For the 13 circuits with longer sentences the differences ranged from 3 months (Circuit 8) to more than 6 months longer (Circuit 1). Offenders in Circuit 13 were sentenced to terms more than two and a half months shorter than Circuit 0. In the context of this offence with a mean sentence length of almost 14.6 months in the reference circuit, the differences for all 14 circuits are substantively significant.

Prediction 2 is supported for burglary: after controlling for criminal history, offence seriousness and offender demographic variables, similarly situated offenders were sentenced to different terms of incarceration in some circuits.

Residual Statistics, Influential Cases and Outliers

Residual statistics were reviewed to identify influential cases and potential outliers. The values for Cook's distance were substantially below 1 in all three models, indicating that there were no cases with an undue influence on the model. The standardised DFBeta measures were all less than 1 for aggravated drink driving and burglary, indicating that none of the cases had an undue influence on the parameter estimates for any of the variables included in the model. However, as reported earlier, one case in the male assaults female model had a standardised DFBeta value exceeding 1 for the threatening to kill or do grievous bodily harm variable. This case had an undue influence on the parameter estimate. The regression co-efficient for the variable was non-significant when the case was removed from the model and as a result both the case and the variable were excluded from the model.

Examination of leverage values and standardised residuals raised concern about potential outliers in all three models. Cases with leverage values that varied from the average leverage value by more than a factor of 2 were scrutinised.

Similar patterns were observed for both aggravated drink driving and male assaults female. The cases with high leverage values were unevenly distributed across the circuits. For aggravated drink driving 78% (102/131) of the cases were from four circuits (4, 9, 13 & 15). The four circuits had the lowest absolute number of incarcerated offenders across all circuits. None of the four circuits was statistically significantly different from the reference circuit. A similar result was found for male assaults female with 57% (21/37) of the cases from two circuits (9 & 13). Neither circuit was statistically significantly different from the reference circuit. As none of the cases for either offence had an undue influence on the model (all values for Cook's distance were less than 1) no changes were made.

Burglary was different; the 34 cases with high leverage values were spread across the circuits. The high leverage values were found to be associated with cases that tended to have more extreme values for important sentencing factors. Most of the 34 were associated with three important sentencing factors (very high numbers of burglary convictions for sentence; high total convictions for sentence; or extensive criminal histories). As none of the cases for either offence had an undue influence on the model (all values for Cook's distance were less than 1) no changes were made.

The standardised residuals for aggravated drink driving and male assaults female are dealt with together because the same issues are raised for both. For aggravated drink driving the proportion of cases with an absolute value greater than 1.96 was marginally

above 5% and the number with an absolute value greater than 2.58 exceeded 1% (37 versus 23). Twenty-four cases with a standardised residual value exceeding an absolute value of 3 were scrutinised. For male assaults female 5% of cases had standardised residuals exceeding 1.96 and 2.2% had an absolute value exceeding 2.58. There were seven cases that exceeded an absolute value of 3.

All bar one of the cases for aggravated drink driving and all 7 of the cases for male assaults female had actual sentences substantially higher than the predicted sentence. The cases for both offences were spread across the circuits with no apparent concentrations. Scrutiny of the offence seriousness and criminal history variables for both offences confirmed that these were cases likely to result in high actual sentences. All bar one of the 30 cases with high actual sentences between the offences were found to have either one of two common characteristics. For 12 of the 29 cases the sentence type was home detention. For the remaining 17 cases the sentence was for cumulative terms of imprisonment.

The imposition of cumulative sentences and the choice of home detention over imprisonment are significant characteristics; they are part of the sentencing decision and would not normally be thought of as sentencing factors. Cumulative versus concurrent sentencing is a choice of how to express the sentence when there are multiple charges. The choice of home detention over imprisonment is the choice of sentence type within the incarceration grouping.

The potential responses included making no changes or adding variables based on these characteristics. To assist in the consideration of the way forward, a detailed examination of the two characteristics was undertaken, including formulating new variables and observing their effects by adding them to the models. The models were improved by the inclusion of both a cumulative sentence variable (0 = no cumulative sentence, 1 = cumulative sentence) and a home detention variable (0 = imprisonment, 1 = home detention). An examination of the effects on the model and the possible explanations for the impact of the two variables led to the conclusion that their inclusion would amount to using an element of the outcome to predict the outcome and would result in mis-estimation of the parameter estimates for the circuit variables. A detailed description of the examination of the two variables is contained in Appendix F.

The models for aggravated drink driving and male assaults female were retained without adjustment. Although there were a range of individual cases with high leverage values or with high standardised residuals, none of these observations had Cook's distance

or standardised DFBeta values that indicated undue influence on the overall model or individual parameter estimates and no observations were removed.

For burglary the number of cases with standardised residuals with an absolute value greater than 1.96 was above 5%. There were 24 cases that exceeded an absolute value of 3. Twenty-three had actual sentences substantially higher than the predicted sentence and one was substantially lower. The 24 cases were scrutinised as potential outliers. Informed by the findings from similar scrutiny of residuals for aggravated drink driving and male assaults female, the initial focus was on the use of home detention and cumulative sentences. However, there was no similar pattern for burglary. Wider scrutiny did not reveal any other common features. The high residuals were all from different circuits, and scrutiny of the seriousness and criminal history variables confirmed that these were cases likely to result in high actual sentences. After examination of the residuals the burglary model was retained in its original form without adjustment. Although there were a range of individual cases with high leverage values or with high standardised residuals, none of these cases had Cook's distance or standardised DFBeta values that indicated undue influence on the overall model or individual parameter estimates and no cases were removed.

Discussion of the Model Results

Sentencing varies by circuit – statistically and substantively

The model results strongly support the predicted effect of location: the circuit in which an offender was sentenced for aggravated drink driving, male assaults female or burglary significantly affected the length of sentence. After controlling for seriousness of offence, criminal history and offender demographic variables similar offenders were sentenced differently between circuits.

The extent of the difference varied between the three offences. The number of circuits significantly different to the reference circuit was greatest for burglary, with 14 of 16 circuits statistically significantly different from Circuit 0. Thirteen of the 14 circuits imposed longer terms of incarceration and one (Circuit 13) imposed shorter terms. For aggravated drink driving seven circuits imposed significantly longer terms and none imposed shorter terms. For male assaults female six circuits imposed significantly longer terms and none imposed shorter terms.

The circuits also varied in the proportionate size of the variations for the circuits significantly different to the reference circuit. For burglary the differences ranged from

18.2% (Circuit 13 where sentences were 2.5 months shorter) to 43.2% (Circuit 1 where sentences were more than 6 months longer). For burglary all the differences for the circuits that were statistically significantly different were substantively significant when compared to the mean sentence (14.6 months) in the reference circuit. For aggravated drink driving the differences ranged from 15.8% (Circuit 5 where sentences were 6 weeks longer) to 32.9% (Circuit 16 where sentences were almost 3 months longer). For aggravated drink driving all the differences for the circuits that were statistically significantly different were substantively significant when compared to the mean sentence (9.1 months) in the reference circuit. The differences were more extreme for male assaults female, ranging from about 1.6% (Circuit 16 where sentences were 3 days longer) to 52.6% (Circuit 1 where sentences were 3.5 months longer). The differences were only substantively significant (a month or more) in two of the six circuits. In the other four circuits sentences ranged from 3 to 25 days longer than the mean sentence (6.7 months) in the reference circuit.

Influence of the seriousness of offence and criminal history variables

The models for all three offences indicate that offence seriousness and criminal history variables are the most significant influences on sentencing. Variables forming these two groupings contributed 88.9% (aggravated drink driving), 69.4% (male assaults female) and 92.4% (burglary) of the R^2 for the three models. The results supported the assumption that the seriousness of offence and criminal history variables would be the strongest influences on sentence length. The importance of the seriousness of the offence and criminal history factors suggests the judges apply a common framework of factors when sentencing. However, the variation between circuits that remained after controlling for the seriousness and criminal history factors suggests that the judges were not applying the factors in the same way.

Influence of gender and ethnicity

The demographic variables of gender, ethnicity and age were not statistically significant in any of the sentence length models. The finding that gender was not statistically significant ran counter to the general findings in sentencing research in the United States (Ulmer, 2012) and in New Zealand (Jeffries et al., 2003; Triggs, 1999).

The finding of no significant effects for ethnicity adds to the mixed results in prior Australian and New Zealand research on indigeneity (Bond & Jefferies 2011; Snowball &

Weatherburn, 2006; and Triggs, 1999 found no differences but Bond et al., 2011 and Jeffries & Bond, 2009 did find differences) but it is out of step with the overall weight of the United States research (Mitchell & Mackenzie, 2004 & Ulmer, 2012). These results may be due to the more detailed offence specific models employed in this research. The extension of this approach to other offences does not mean that no effects due to gender or ethnicity will be found. It is possible that more detailed models for other offences may reveal effects that would otherwise have remained hidden (as the ethnicity effect for the burglary in-out model demonstrates).

Rank order of the circuits within and between offences and sentencing decisions

The model results for the circuits can be combined into a single table showing the rank order of the circuits from highest to lowest co-efficient (from most severe to most lenient). The rank order of the circuits can also be compared between the two sentence decisions for each offence.

The consideration of the rank order begins with rank order of circuits for the offences for sentence length, the rank orders are displayed in Table 7.23. Five circuits (1, 3, 5, 7, & 16) imposed longer terms of incarceration than the reference circuit for all three offences. Three circuits (9, 13 & 15) never incarcerated for longer terms than the reference circuit, although Circuit 13 did impose significantly shorter terms of imprisonment for burglary. Spearman's rho (correlation co-efficient) was calculated for each of the three pairs of offences to determine whether there was any apparent link between the rank orders of the circuits for each offence. There was no correlation between the rank orders for any of the pairs: aggravated drink driving and burglary ($r_s = .34, p = .178$); aggravated drink driving and male assaults female ($r_s = .34, p = .184$); and male assaults female and burglary ($r_s = .02, p = .948$). These observations suggest that the decision on the length of incarceration is influenced more by a view of the relative seriousness of each offence type than by some general theory about sentence length. This also means that the position of each circuit in the hierarchy for one offence is not a predictor of its position for another offence.

Table 7.23: Order of circuits by co-efficient from largest to smallest by offence type¹

Aggravated drink driving	Male assaults female	Burglary
16	1	1
7	2	14
12	3	11
3	7	12
14	5	6
1	16	4
5	4	2
9	10	5
4	8	10
11	13	16
2	15	3
10	11	7
8	9	8
0	0	15
15	14	0
6	12	9
13	6	13

¹ The circuits that were statistically significantly different to the reference circuit are shown in bold.

The order of circuits by sentence length was also contrasted with the order by the likelihood of incarceration (odds ratio) for each offence. The order of the circuits across the six models is portrayed in Table 7.24.

Two circuits (1 & 7) were both more likely to incarcerate and to impose longer sentences across all three offences. Two circuits (13 & 15) were never more likely to incarcerate and never imposed longer sentences across all three offences.

The contrast of the two decisions for each offence type reveals a different pattern for each offence type. For aggravated drink driving all seven circuits that imposed longer sentences were also significantly more likely to incarcerate, thus leaving seven circuits that were more likely to incarcerate but which did not impose longer sentences. For male assaults female three of the six circuits that imposed longer sentences were more likely to incarcerate. Unlike aggravated drink driving there were circuits (Circuits 3, 5 & 16) that were not more likely to incarcerate but which did impose longer sentences. The rank order for the likelihood of incarceration was not a good predictor of the rank order for length of sentence or vice versa for either offence, as demonstrated by the calculation of Spearman's

rho: aggravated drink driving ($r_s = .20$, $p = .439$); and male assaults female ($r_s = .42$, $p = .094$).

Table 7.24: Order of circuits by odds ratio from highest to lowest and by co-efficient (sentence length) from longest to shortest by offence type¹

Aggravated drink driving		Male assaults female		Burglary	
OR	Co-efficient	OR	Co-efficient	OR	Co-efficient
8	16	7	1	6	1
12	7	14	2	4	14
6	12	1	3	10	11
4	3	10	7	8	12
2	14	2	5	12	6
7	1	12	16	16	4
10	5	4	4	11	2
1	9	5	10	14	5
16	4	3	8	3	10
14	11	9	13	2	16
11	2	16	15	7	3
3	10	8	11	5	7
5	8	15	9	1	8
9	0	6	0	15	15
0	15	0	14	0	0
13	6	11	12	9	9
15	13	13	6	13	13

¹ The circuits that were statistically significantly different to the reference circuit are shown in bold.

For burglary all 13 of the circuits that imposed longer sentences were more likely to incarcerate. There was a link between the order of the circuits for the two decisions for burglary, $r_s = .52$, $p < .05$. The burglary result was due to the exact match for the four circuits that did not impose longer sentences and were not more likely to incarcerate. When the analysis was re-run for the 13 circuits that were all significantly more likely to incarcerate and impose longer sentences there was no correlation ($r_s = -.08$, $p = .803$).

Overall there is little evidence that the position of a circuit in the rank order for one offence or one decision is predictive of its position for another offence or decision. The contrast of the rank orders for aggravated drink driving and burglary for the use of incarceration (see Chapter 6, p.149-50) is the exception. These findings suggest that in general sentencing approaches are tailored to the offences rather than being based on some general theory. The findings suggest that in addition to inconsistency of sentencing

decisions between circuits for the same offence there is little congruence within and between circuits on the general seriousness of the three offences relative to one another.

In one sense the results for the rank ordering of the two decisions for each offence are puzzling. If after controlling for sentencing variables some circuits are more likely to incarcerate than others it would be reasonable to assume that they would be more likely to impose shorter sentences. They would impose shorter sentences because they are incarcerating less serious cases. If that were true then Spearman's rho would be negative. In fact, as is readily observed in the tables, circuits that more likely to incarcerate are more likely to impose longer terms.

Summary

In sum, the linear regression models for all three offences support Prediction 2: similarly situated offenders were sentenced differently depending on which circuit they were sentenced in. The extent of inconsistency and the pattern varied between the offences. The strength of the findings varied between the offences. Although the circuit was significant, the main determinants of the length of sentence were offence seriousness and criminal history variables. Gender, ethnicity and age did not have any significant effect on sentence length. Comparison of the rank order of the circuits by offence suggested that the order was more likely to be due to views about the appropriate sentence lengths related to the particular offence than any general theory about sentence lengths.

Chapter 8: The Interpretation and Use of Sentencing Variables

Subject to evidence of inconsistency between circuits and confirmation that offence seriousness and criminal history variables were the principal determinants of sentence it was predicted that there would be differences in the interpretation and use of sentencing variables between circuits. The models presented in the two preceding chapters demonstrate that both of conditions were met. To test the predictions case level random co-efficient models were assembled.

Significant random effects were demonstrated for a selection of offence seriousness and criminal history variables in the determination of incarceration and the determination of sentence length for all three offences. However, the results for male assaults female were much weaker. The weaker result is not surprising given that the extent of variation between circuits was more restricted for this offence. It may also reflect limitations due to omitted variables, the comparatively low rate of incarceration and the narrow scope of the offence within the spectrum of physical assaults. Overall, Predictions 3 and 4 were supported: the interpretation and use of some or all of the seriousness of offence, criminal history or extra-legal variables differed between circuits for both the use of incarceration and the determination of sentence length.

Development of the Models

The first iterations of the multi-level models were based on the logistic and linear regression models reported in the preceding chapters. Ordinarily multi-level models would be built up based on the sequential addition of variables according to their empirical or theoretical relevance (Raudenbush & Bryk, 2002). However, for this research the development of the single level models made that process redundant. Comparisons of the co-efficients reported in the single level models to case level fixed effects multi-level models using the same variables confirmed the validity of using the same variables.

In accord with the advice of Raudenbush & Bryk (2002) no attempt was made to run a model saturated with all the individual variables that had significant fixed effects and then refine it backwards to delete any variables that did not have significant random effects. To be effective the data had to be capable of sustaining stable parameter estimates in each separate unit (circuit). The approach followed was to isolate those variables that might have random effects in a multivariate model by identifying those that had significant bivariate random effects (a significant effect when the variable was allowed to vary

randomly while all other case level variables were fixed). The variables with significant random effects became the pool of variables for the development of a multivariate random co-efficients model. The multivariate random co-efficient models were built up by the sequential inclusion of the variables based on the strength of their bivariate random effects.

A problem was encountered with four of the six models. The use of the variables as specified for the single level models was successful when the models were limited to fixed effects but were too complex when the models were extended to random effects. The models had to be adjusted by simplifying some variables. With one exception simplification was achieved by reducing the number of categories for some categorical variables. The exception was the combining of two continuous variables into a new variable for the burglary sentence length model. The simplifications were based on a review of the results of the earlier models and sought to minimise the loss of predictive power by combining categories with minimal differences in their co-efficients. The specific changes made are listed in conjunction with the reporting of each of the models.

Once the variables were simplified the modelling approach was repeated. The first step was a model treating the case level variables as fixed and allowing the intercept to vary randomly. The model was then varied to identify those variables that had a bivariate random effect. Finally an iterative approach was followed to develop a model incorporating those variables that continued to have a significant random effect in a multivariate random co-efficient model.

All continuous variables were grand mean centred to improve the interpretation of the model (Raudenbush & Bryk, 2002). Grand mean centring means that the variable for each case was centred on the mean calculated across all the cases in the dataset. The alternative was to use group mean centring, with the variable for each case centred on the mean for the circuit (group) the case was sentenced in. In the context of this research the grand mean was preferred to the group (circuit) mean because the use of the group mean could result in cases that were otherwise identical, except for circuit, differing in relation to seriousness and criminal history variables whenever the group (circuit) means varied from the grand mean. For instance, two cases for offenders with five previous convictions who were sentenced in different circuits could have transformed variables showing one as three above the mean (for a circuit mean of two) and the other as two above the mean (for a circuit mean of three), whereas a grand mean of 2.5 places both cases 2.5 above the mean. The use of the group mean would have resulted in a mis-estimation of the influence of the variable and the effect of circuits on sentencing. Use of the grand mean allows the

intercept to be interpreted as representing an offender at the mean for continuous variables and in the reference category for categorical variables.

Aggravated Drink Driving In-out Model

Six variables were simplified for the aggravated drink driving in-out hierarchical generalised linear model. The changes are listed in Table 8.1.

Table 8.1: Aggravated drink driving in-out model simplification of case level variables

Variable	Original categories	Simplified categories
Age	0 = 17 - 31 1 = 32 - 39 2 = 40 - 47 3 = 48 or older	0 = 17 - 31 1 = 32 or older
Total charges sentenced	0 = 1 charge for sentence 1 = 2 charges for sentence 2 = 3 or more charges for sentence	0 = 1 or 2 charges for sentence 1 = 3 or more charges for sentence
Elapsed time from charge to charge outcome	0 = ≤ 6 weeks 1 = > 6 weeks ≤ 3 months 2 = > 3 months ≤ 6 months 3 = > 6 months	0 = ≤ 6 weeks 1 = > 6 weeks
Drink driving convictions in the last ten years	0 = none 1 = one 2 = two 3 = three or more	0 = none 1 = one 2 = two or more
Time since last drink driving conviction	0 = more than 5 years 1 = 5 years or less and more than 18 months 2 = 18 months or less	0 = more than 5 years 1 = less than or equal to 5 years
Prior incarceration	0 = never 1 = yes but not for drink driving 2 = yes including for drink driving	0 = no 1 = yes

The finalised model is reported in Tables 8.2 (the fixed effects) and 8.3 (the random effects).

Table 8.2: Case level random co-efficient model for aggravated drink driving in-out decision fixed effects component (for random effects see Table 8.3) ($n = 9,282$)

Variables (reference categories in brackets)	Co-efficient	S.E.	<i>t</i>-ratio	<i>p</i>-value	Odds Ratio
Intercept ^a	-6.929	0.376	-18.439	<.001	0.001
Alcohol level ^{a b}	0.541	0.092	5.895	<.001	1.718
Not guilty plea (guilty)	-0.520	0.229	-2.273	.023	0.595
Multiple drink drive charges (no) ^a	2.515	0.305	8.245	<.001	12.372
Disqualified driving (no)	0.445	0.091	4.89	<.001	1.561
Serious driving charge (no)	0.492	0.142	3.473	<.001	1.636
Breach of sentence (no)	0.835	0.155	5.399	<.001	2.305
Three or more convictions for sentence (no)	0.559	0.098	5.68	<.001	1.748
Charge to outcome > 6 weeks (< 6 weeks)	1.099	0.113	9.732	<.001	3.002
Drink drive conviction in the last five years (no) ^a	0.574	0.084	6.871	<.001	1.775
Prior convictions for drinking and driving					
3 priors (2) ^a	1.318	0.184	7.173	<.001	3.736
4 priors (2) ^a	2.052	0.189	10.868	<.001	7.785
5 or 6 priors (2) ^a	2.649	0.239	11.066	<.001	14.142
7 or more priors (2) ^a	3.459	0.280	12.357	<.001	31.799
Drink drive convictions in the past 10 years					
1 (0)	1.339	0.154	8.691	<.001	3.814
2 or more (0)	1.972	0.154	12.837	<.001	7.188
Previously imprisoned (no)	0.512	0.076	6.713	<.001	1.669
Conviction for breach of sentence in the last five years (no)	0.537	0.087	6.162	<.001	1.711
Aged 32 or more (aged 17 - 31) ^a	-0.395	0.151	-2.618	.019	0.674

^a There were 16 degrees of freedom for these parameters and 9119 for all other parameters.

^b The variable was grand mean centred.

Table 8.3: Case level random co-efficient model for aggravated drink driving in-out decision random effects component (for fixed effects see Table 8.2) ($n = 9,282$)

Random Effect	Standard Deviation	Variance Component	<i>d.f.</i>	χ^2	<i>p</i> -value
Intercept	1.184	1.401	16	56.150	<.001
Alcohol level	0.257	0.066	16	28.843	.025
Multiple drink drive charges (no)	1.059	1.121	16	50.601	<.001
Drink drive conviction in the last five years (no)	0.230	0.053	16	32.728	.008
Prior convictions for drinking and driving					
3 priors (2)	0.628	0.395	16	34.373	.005
4 priors (2)	0.589	0.347	16	29.363	.021
5 or 6 priors (2)	0.821	0.674	16	38.534	.002
7 or more priors (2)	0.945	0.892	16	36.383	.003
Aged 32 or more (aged 17 - 31)	0.464	0.216	16	32.880	.008

There were significant random effects for a series of offence seriousness and criminal history variables. The variation in the likelihood of incarceration between circuits is partially explained by differences in the custodial threshold centred around: the alcohol level; whether the offender was being sentenced for multiple counts of drinking and driving; the timing of the most recent conviction; the number of prior convictions; and age. With the exception of age these variables are all directly representative of the seriousness of the current offending or the offenders' past records for drink driving. However, it is possible that age is indirectly related to criminal history and adds a nuance to the consideration of the timing of past convictions for drink driving. Older offenders may be less likely to be incarcerated because their prior convictions were spread over a longer period and they were judged to be less culpable than younger offenders who had acquired the same record more quickly. The influence of age might also reflect heightened concern about those younger offenders identified as or perceived to be "boy racers". It can be inferred from the results that differences in views about the significance of prior offending for incarceration are more important to understanding the variation in the custodial threshold than the offence seriousness factors. While multiple drink drive charges have a substantial random effect the variable is only in play in 458 of the 9,282 cases. The influence of other seriousness factors such as the presence of disqualified driving, other serious driving and breach of sentence convictions is similar across the circuits

demonstrating that variation is not located in differing interpretations of the significance of these variables. The consequences of repeated offending and concern about the risk of re-offending appear to vary between the circuits. The variations between circuits for aggravated drink driving are associated with how seriously drink driving type is viewed by the judges and the wider community in each circuit.

Prediction 3 is supported: the interpretation and use of some of the seriousness of offence and criminal history variables differed between circuits for the use of incarceration for aggravated drink driving.

Aggravated Drink Driving Sentence Length Model

No variables were simplified for this model. The finalised model is reported in Tables 8.4 (the fixed effects) and 8.5 (the random effects). There were significant random effects for a series of offence seriousness and criminal history variables. The variation in the length of sentence between circuits is partially explained by differences in the custodial threshold centred around: whether the offender was sentenced for multiple drink driving convictions; a greater number of charges for sentence; more prior convictions for drinking and driving; and longer prior terms of incarceration for drinking and driving. While there is some overlap with the in-out decision there is also some notable differences. In particular the alcohol level does not form part of the set of variables contributing to the explanation of variation in the determination of sentence length. Conversely the number of charges for sentence is treated differently between the circuits in the determination of sentence length but was not a distinguishing factor for the custodial threshold. The timing of prior convictions for drinking and driving was not a distinguishing factor in the determination of the custodial threshold but the longest prior term of incarceration was. This suggests that the severity of prior offending matters more in some circuits than others for the determination of sentence length. There is a more even balance between the seriousness of current offending and criminal history factors as determinants of the differences in approach between circuits than there was for the use of incarceration.

Table 8.4: Case level random co-efficient model for aggravated drink driving sentence length decision fixed effects component (for random effects see Table 8.5) ($n = 2,312$)

Variable	Co-efficient	S.E.	<i>t</i> -ratio	Approx <i>d.f.</i>	<i>p</i> -value
Intercept	9.030	0.277	32.642	16	<.001
Number of days from charge to charge outcome ^a	0.003	0.001	2.555	2221	.011
Alcohol level ^a	1.103	0.153	7.193	2221	<.001
Multiple drink drive charges (no)	2.514	0.403	6.232	16	<.001
Disqualified driving (no)	0.452	0.204	2.222	2221	.026
Breach of bail (no)	-1.516	0.351	-4.315	2221	<.001
Breach of sentence (no)	-0.846	0.298	-2.835	2221	.005
Total number of charges for sentence ^a	0.546	0.113	4.824	16	<.001
Prior convictions for drink driving ^a	0.456	0.072	6.325	16	<.001
Years since the last drink drive conviction ^a	-0.126	0.031	-4.135	2221	<.001
Longest term of imprisonment for drink driving ^a	0.330	0.040	8.18	16	<.001

^a This variable was grand mean centred.

Table 8.5: Case level random co-efficient model for aggravated drink driving sentence length decision random effects component (for fixed effects see Table 8.4) ($n = 2,312$)

Variable	Standard deviation	Variance Component	<i>d.f.</i>	χ^2	<i>p</i> -value
Intercept	0.964	0.929	16	145.278	<.001
Multiple drink drive charges	1.085	1.176	16	28.643	.026
Total number of charges for sentence	0.391	0.153	16	70.144	<.001
Prior convictions for drink driving	0.242	0.058	16	46.790	<.001
Longest term of imprisonment for drink driving	0.125	0.016	16	46.648	<.001
level-1, <i>r</i>	3.911	15.292			

Prediction 4 is supported: the interpretation and use of some of the seriousness of offence and criminal history variables differed between circuits for the determination of incarceration sentence lengths for aggravated drink driving.

Male Assaults Female In-Out Model

Five variables were simplified for the male assaults female in-out hierarchical generalised linear model. The changes are listed in Table 8.6. The finalised model is reported in Tables 8.7 (the fixed effects) and 8.8 (the random effects).

Table 8.6: Male assaults female in-out model simplification of case level variables

Variable	Original categories	Simplified categories
Age	0 = 17 - 23 1 = 24 - 30 2 = 31 - 38 3 = 39 or older	0 = 17 - 23 1 = 24 or older
Total charges sentenced	0 = 1 charge for sentence 1 = 2 charges for sentence 2 = 3 charges for sentence 3 = 4 or more charges for sentence	0 = 1 charge for sentence 1 = 2 or more charges for sentence
Prior convictions for any offence	0 = none 1 = 1 - 9 2 = 10 - 19 3 = 20 or more	0 = none 1 = 1 - 9 2 = 10 or more
Prior convictions for male assaults female	0 = none 1 = one 2 = two 3 = three or more	0 = none or 1 1 = two or more
Prior incarceration for any offence	0 = never 1 = once 2 = 2 - 4 times 3 = 5 or more times	0 = no 1 = yes

Table 8.7: Case level random co-efficient model for the male assaults female in-out decision fixed effects component (for random effects see Table 8.8) ($n = 6,166$)

Variables (reference categories in brackets)	Co-efficient	S.E.	<i>t</i>-ratio	<i>p</i>-value	Odds Ratio
Intercept ^a	-5.202	0.342	-15.213	<.001	0.006
Charge laid indictably (no)	1.866	0.297	6.285	<.001	6.466
Not guilty plea (guilty)	0.544	0.176	3.095	.002	1.723
Multiple male assaults female (no)	0.702	0.144	4.865	<.001	2.017
Threat to kill or do GBH (no)	2.267	0.584	3.881	<.001	9.655
Breach of sentence (no)	1.026	0.159	6.446	<.001	2.790
Two or more charges for sentence (one)	1.113	0.120	9.297	<.001	3.044
Elapsed time - charge to charge outcome					
> 6 weeks ≤ 3 months	0.550	0.144	3.808	<.001	1.733
> 3 months ≤ 6 months	0.420	0.145	2.895	.004	1.523
> 6 months	0.099	0.168	0.589	.556	1.104
Two or more prior convictions for male assaults female (0 or 1) ^a	0.815	0.162	5.047	<.001	2.260
Conviction for male assaults female in the last 5 years (no)	0.609	0.119	5.134	<.001	1.839
Conviction for breach of protection order in the last 5 years (no)	0.607	0.143	4.244	<.001	1.834
Conviction for serious violence in the last 5 years (no)	0.953	0.126	7.533	<.001	2.593
Conviction for breach of sentence in the last 5 years (no)	0.852	0.108	7.86	<.001	2.344
Prior incarceration for any offence (no) ^a	0.964	0.177	5.464	<.001	2.623
Prior convictions for any offence					
1-9 priors (0)	0.541	0.312	1.735	.083	1.718
10 or more priors (0)	1.227	0.323	3.805	<.001	3.413
Aged 24 or more at the time of offending (no)	-0.399	0.132	-3.029	.002	0.671

^a There were 16 degrees of freedom for these parameters and 6099 for all other parameters.

Table 8.8: Case level random co-efficient model for the male assaults female in-out decision random effects component (for fixed effects see Table 8.7) ($n = 6,166$)

Variables (reference categories in brackets)	Standard Deviation	Variance Component	<i>d.f.</i>	χ^2	<i>p</i>-value
Intercept	0.462	0.214	16	38.264	.002
Two or more prior convictions for male assaults female (0 or 1)	0.348	0.121	16	28.971	.024
Prior incarceration for any offence (no)	0.444	0.197	16	29.154	.023

The difference in the likelihood of incarceration across the circuits is partially explained by differences in the thresholds for imprisonment depending: on the number of prior convictions; and whether the offender has ever been incarcerated before. None of the seriousness factors modelled had random effects. This indicates that variation between circuits for the decision to incarcerate hinges around different weight being placed on criminal history and the consequences of repeated offending or the risk of re-offending. However, the absence of variables indicating the seriousness of the offence in terms of either the level of violence employed or the impact of the offending on the victim prevents a conclusion being drawn to the effect that offence seriousness factors do not play a part in distinguishing the approaches between the circuits.

Prediction 3 is supported: the interpretation and use of some of the seriousness of offence and criminal history variables differed between circuits for the use of incarceration for male assaults female.

Male Assaults Female Sentence Length Model

No variables were simplified for this model. The finalised model is reported in Tables 8.9 (the fixed effects) and 8.10 (the random effects). The model only identifies the way the charge was laid having a significant random effect between circuits on the determination of sentence length. As only 5.0% (36/722) of the cases include indictably laid male assaults female charges the explanatory value of the effect is limited. The limited result is not surprising given that only two of the six circuits that varied significantly from the reference circuit in the single level models differed to a substantial extent. The effectiveness of the model may be partly due to the smaller number of cases in the sentence length dataset, with 722 cases insufficient to enable differences to be discerned. The result indicates that once a decision has been made to incarcerate there is reasonable consistency between circuits in terms of the influence of criminal history factors on sentence length. If the

inter-circuit inconsistency is explained by differences in the weight afforded to any seriousness factors it is arguably likely to be associated with either or both of the omitted variables representing the level of violence employed or the impact of the offending on the victim.

Table 8.9: Case level random co-efficient model for the male assaults female sentence length decision fixed effects component (for random effects see Table 8.10) ($n = 722$)

Variable	Co-efficient	S.E.	<i>t</i> -ratio	Approx <i>d.f.</i>	<i>p</i> -value
Intercept	6.204	0.385	16.098	16	<.001
Charge laid indictably (no)	3.896	1.335	2.919	16	.01
Multiple male assaults female charges (no)	2.033	0.413	4.917	683	<.001
Breach of bail (no)	-1.106	0.520	-2.129	683	.034
Total number of charges for sentence ^a	0.477	0.067	7.174	683	<.001
Prior convictions for male assaults female ^a	0.266	0.118	2.258	683	.024
Prior incarceration for male assaults female (no)	1.433	0.421	3.407	683	<.001
Recent serious violence conviction (none or > 5 years ago)	1.144	0.341	3.352	683	<.001

^a This variable was grand mean centred.

Table 8.10: Case level random co-efficient model for the male assaults female sentence length decision random effects component (for fixed effects see Table 8.9) ($n = 722$)

Variable	Standard deviation	Variance Component	<i>d.f.</i>	χ^2	<i>p</i> -value
Intercept	1.124	1.264	12	65.705	<.001
Charge laid indictably (no)	3.841	14.757	12	38.227	<.001
level-1, <i>r</i>	4.111	16.898			

Note: The chi-square statistics reported above are based on only 13 of 17 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

There is limited support for Prediction 4: the interpretation and use of some of the seriousness of offence and criminal history variables differed between circuits for the determination of incarceration sentence lengths for male assaults female.

Burglary In-Out Model

Eight variables were simplified for the burglary in-out hierarchical generalised linear model. The changes are listed in Table 8.11.

Table 8.11: Burglary in-out model simplification of case level variables

Variable	Original categories	Simplified categories
Burglary convictions for sentence		
	0 = 1	0 = 1
	1 = 2	1 = 2 or more
	2 = 3 or more	
Approximate value of the lead burglary offence		
	0 = <\$500	0 = <\$500
	1 = \$500 to \$5,000	1 = \$500 or more or unknown
	2 = >\$5,000	
	3 = unknown	
Total charges sentenced		
	0 = 1 charge for sentence	0 = 1 charge for sentence
	1 = 2 charges for sentence	1 = 2, 3 or 4 charges for sentence
	2 = 3 or 4 charges for sentence	2 = 5 or more charges for sentence
	3 = 5, 6 or 7 charges for sentence	
	4 = 8 or more charges for sentence	
Prior convictions for any offence		
	0 = none	0 = none
	1 = 1 to 4	1 = 1 to 4
	2 = 5 to 9	2 = 5 to 19
	3 = 10 to 19	3 = 20 or more
	4 = 20 to 49	
	5 = 50 or more	
Prior convictions for serious dishonesty (including burglary)		
	0 = none	0 = none
	1 = 1 or 2	1 = 1 to 9
	2 = 3 to 9	2 = 10 or more
	3 = 10 or more	
"Longest prior incarceration for burglary" became "Prior incarceration for burglary"		
	0 = never	0 = no
	1 = less than or equal to 12 months	1 = yes
	2 = greater than 12 months and less than or equal to 24 months	
	3 = greater than 24 months	
Elapsed time from charge to charge outcome		
	0 = ≤ 6 weeks	0 = ≤ 6 weeks
	1 = > 6 weeks ≤ 3 months	1 = > 6 weeks
	2 = > 3 months ≤ 6 months	
	3 = > 6 months	

Variable	Original categories	Simplified categories
Ethnicity	0 = European 1 = Māori 2 = Pacific 3 = Other	0 = European or other 1 = Māori or Pacific

The finalised model is reported in Tables 8.12 (the fixed effects) and 8.13 (the random effects).

There were significant random effects for a series of offence seriousness and criminal history variables. The variation in the likelihood of incarceration between circuits is partially explained by differences in the custodial threshold centred around: whether the most recent conviction for burglary was in the last 5 years; whether the offender had been convicted for breach of sentence in the last 5 years; the number of convictions being sentenced; the time taken from charge to charge outcome (sentence); and the number of previous convictions for burglary or other serious dishonesty. With the exception of time from charge to charge outcome these are all self-explanatory offence seriousness or criminal history variables. The time from charge to charge outcome is a mitigating factor influencing the choice of sentence. The shorter the time to the resolution the more likely it is that the offender entered an early guilty plea and the greater the likelihood that the offender is seen to be taking responsibility and is potentially more remorseful and thus the sentence is more lenient. The circuits are distinguished by the extent to which they are more or less willing to sentence leniently in response to these considerations. There are differences in the threshold related to criminal history and the consequences of repeated offending or the risk of re-offending. The distinction between the circuits is only significant for offenders with 1-9 prior convictions with serious dishonesty; there was no random effect for offenders with 10 or more prior convictions. This suggests that circuits differ in how to treat offenders with less extensive histories but not once they cross a threshold. It is also notable that circuits place different emphasis on the offender's recent willingness to comply with sentences. Some circuits are less willing to take the risk of imposing a lesser sentence.

Table 8.12: Case level random co-efficient model for the burglary in-out decision fixed effects component (for random effects see Table 8.13) ($n = 5,740$)

Variables (reference categories in brackets)	Co-efficient	S.E.	<i>t</i>-ratio	<i>p</i>-value	Odds Ratio
Intercept ^a	-4.271	0.227	-18.83	<.001	0.014
Female (male)	-0.569	0.143	-3.98	<.001	0.566
Charge laid indictably (no)	1.707	0.155	11.022	<.001	5.515
Breach of sentence (no)	0.559	0.114	4.911	<.001	1.749
Other serious conviction (no)	0.595	0.098	6.057	<.001	1.812
Reparation ordered (no)	-0.501	0.080	-6.248	<.001	0.606
Conviction for burglary last 5 years (no) ^a	0.833	0.167	5	<.001	2.301
Prior conviction for serious violence (no)	0.641	0.102	6.262	<.001	1.898
Conviction for breach of sentence last 5 years (no) ^a	0.533	0.160	3.332	.004	1.703
Multiple burglary convictions for sentence (no)	0.850	0.098	8.637	<.001	2.339
Total convictions for sentence					
2 to 4 convictions (1) ^a	0.559	0.156	3.587	.002	1.748
5 plus convictions (1)	0.889	0.070	12.724	<.001	2.432
Charge to outcome > 6 weeks (< 6 weeks) ^a	0.179	0.208	0.861	.402	1.197
Prior convictions for burglary and other serious dishonesty					
1 to 9 (0) ^a	0.393	0.159	2.473	.025	1.481
10 or more (0)	0.338	0.089	3.788	<.001	1.402
Prior convictions for any offence					
1 to 4 (0)	0.486	0.166	2.927	.003	1.626
5 to 19 (0)	0.541	0.088	6.131	<.001	1.717
20 or more (0)	0.512	0.071	7.242	<.001	1.669
Previous incarceration for burglary (no)	0.743	0.127	5.87	<.001	2.102
Māori or Pacific Offender (no)	0.239	0.084	2.866	.004	1.270
Approximate value of lead offence \$500 or more or unknown (no)	0.541	0.083	6.525	<.001	1.718

^a There were 16 degrees of freedom for these parameters and 5623 for all other parameters.

Table 8.13: Case level random co-efficient model for the burglary in-out decision random effects component (for fixed effects see Table 8.12) ($n = 5,740$)

Variables (reference categories in brackets)	Standard Deviation	Variance Component	<i>d.f.</i>	χ^2	<i>p</i>-value
Intercept	0.466	0.217	16	27.381	.037
Conviction for burglary last 5 years (no)	0.518	0.268	16	39.237	.001
Conviction for breach of sentence last 5 years (no)	0.522	0.272	16	47.853	<.001
Sentenced for 2 to 4 convictions (1)	0.434	0.188	16	35.262	.004
Charge to outcome > 6 weeks (< 6 weeks)	0.727	0.528	16	48.697	<.001
1 to 9 prior convictions for burglary and other serious dishonesty (0)	0.391	0.153	16	34.136	.005

This was the only multi-level model that tested for random effects in the influence of gender and ethnicity on sentence. These variables were not included in the other models because the variables were not significant in the single level models. Both variables had significant fixed effects as expected, but no significant random effects were demonstrated. The effects of gender and ethnicity on the determination of sentence type were similar between the circuits.

Prediction 3 is supported: the interpretation and use of some of the seriousness of offence and criminal history variables differed between circuits for the use of incarceration for burglary.

Burglary Sentence Length Model

The continuous variables representing prior convictions for burglary and prior convictions for other serious dishonesty were simplified for the burglary sentence length model. The simplification differs from the reduction in the number of categories employed for the in-out modelling. The only change was the creation of a new continuous variable counting the total number of prior convictions for serious dishonesty including burglary. The finalised model is reported in Tables 8.14 (the fixed effects) and 8.15 (the random effects).

Table 8.14: Case level random co-efficient model for the burglary sentence length decision fixed effects component (for random effects see Table 8.15) ($n = 2,338$)

Variable	Co-efficient	S.E.	<i>t</i> -ratio	Approx <i>d.f.</i>	<i>p</i> -value
Intercept	13.148	0.666	19.745	16	<.001
Charge laid indictably (no)	5.795	0.467	12.408	2239	<.001
Days from charge to charge outcome ^a	0.005	0.002	2.831	2239	.005
Judge type (criminal)	-1.295	0.593	-2.182	2239	.029
Plea (guilty plea)	1.799	0.827	2.175	2239	.03
Number of burglary charges ^a	0.762	0.109	6.985	16	<.001
Receiving (no)	2.326	0.492	4.727	2239	<.001
Breach of bail (no)	-1.761	0.451	-3.908	2239	<.001
Breach of sentence (no)	0.822	0.631	1.303	16	.211
Violent offence (no)	1.722	0.416	4.139	2239	<.001
Serious driving offence (no)	1.109	0.465	2.387	2239	.017
Total number of charges for sentence ^a	0.319	0.039	8.218	2239	<.001
Count of offence dates for the burglary ^a	2.200	0.271	8.121	2239	<.001
Reparation order (no)	-0.956	0.369	-2.594	2239	.01
Longest prior term of incarceration for burglary ^a	0.258	0.026	9.766	16	<.001
Prior convictions for serious dishonesty including burglary ^a	0.040	0.014	2.94	16	.01
Prior convictions for other serious offences ^a	0.151	0.031	4.859	2239	<.001
Approximate value of the lead offence					
\$500 to \$5,000 (<\$500 or unknown)	1.115	0.365	3.057	2239	.002
>\$5,000 (<\$500 or unknown)	3.426	0.504	6.801	2239	<.001

^a This variable was grand mean centred.

There were significant random effects for a series of offence seriousness and criminal history variables. The variation in the length of sentence between circuits is partially explained by differences in the custodial threshold centred on additions to sentence length accorded to offenders: being sentenced for multiple burglaries; being sentenced for a breach of sentence; who have served longer prior terms of incarceration for burglary; and who have for more prior convictions for serious dishonesty.

While there is some overlap with the in-out decision there is also some notable differences. The variation between circuits for the determination of sentence length is

more strongly centred on burglary specific factors than the decision on whether to incarcerate. While breach of sentence features in both models it is notable that in the in-out model circuits vary in terms of the influence of a recent prior conviction for breach of sentence but in determining sentence length they vary in terms of the influence of a current conviction.

Table 8.15: Case level random co-efficient model for the burglary sentence length decision random effects component (for fixed effects see Table 8.14) ($n = 2,338$)

Variable	Standard deviation	Variance Component	<i>d.f.</i>	χ^2	<i>p</i> -value
Intercept	2.240	5.016	16	110.962	<.001
Number of burglary charges	0.288	0.083	16	38.494	.002
Breach of sentence	1.958	3.835	16	44.084	<.001
Longest prior term of incarceration for burglary	0.079	0.006	16	40.020	.001
Prior convictions for serious dishonesty including burglary	0.038	0.001	16	45.194	<.001
level-1, <i>r</i>	7.872	61.969			

Prediction 4 is supported: the interpretation and use of some of the seriousness of offence and criminal history variables differed between circuits for the determination of incarceration sentence lengths for aggravated drink driving.

Summary

The models described above support Predictions 3 & 4: the interpretation and use of some seriousness of offence and criminal history variables varied between circuits in terms of their influence on the in-out decision (Prediction 3) and the determination of sentence length (Prediction 4). There was variation in the results between offences. The identification of the factors treated differently was more limited for male assaults female. This potentially reflects the lower level of variation for male assaults female but may also be a consequence of omitted variables, limited numbers for sentence length modelling or the narrowness of the offence within the wider spectrum of physical assaults limiting the observation of sentencing variation.

Chapter 9: The Influence of Circuit Context on Sentencing

The research questions asked whether contextual differences between the circuits were associated with the variation in sentencing between circuits. The question was addressed by expanding the case level random co-efficient models reported in Chapter 8 into circuit level random co-efficient models by including contextual variables applying to the circuits.

The models were severely limited due to the small number of circuits. With only 17 circuits the models can only accommodate one circuit level variable at a time. The influence of the contextual variables had to be examined in a series of bivariate models. Multi-collinearity between the available circuit level variables prevented determination of the influence attributable to any of the contextual variables.

To some extent the limitations were overcome by the inclusion of a variable separating metropolitan and provincial circuits in the range of variables modelled. The two categories of circuit based on the proportion of the circuit population served by city or suburban courts were also distinguished from one another by many of the more substantive contextual variables included in the modelling. Whether a circuit was metropolitan or provincial was significantly associated with variations in sentencing between circuits in three of the six models: the likelihood of incarceration for aggravated drink driving; the likelihood of incarceration for burglary; and the length of incarceration for burglary. The modelling revealed associations between a range of other contextual variables and sentencing variation between circuits which were consistent with the differentiation of metropolitan and provincial circuits for each variable. However, due to the high levels of multi-collinearity between the variables identification of the more substantive contextual variables contributing to the differences between metropolitan and provincial circuits was precluded.

Once again the aggravated drink driving and burglary models were more informative than the male assaults female models. The distinction between the offences was consistent with the results of the earlier models. It is likely to be attributable to the constraints affecting the modelling of male assaults female sentencing: the low incarceration rate; relatively small number of sentence length observations; and the narrowness of the offence. These constraints compound to make the identification of contextual variables associated with sentencing variation difficult.

Development of the Models

The hierarchical models including the circuit level contextual variables were an extension of the case level random co-efficient models. Raudenbush & Bryk (2002) advise against the use of saturation models and deletion of non-significant variables when circuit level variables are added to the finalised case level models. In this research saturation was not possible in any event. With only 17 circuits it was necessary to model each circuit level contextual variable separately by adding it to the finalised case level model.

The selection of the variables was informed by the United States research, with the specific variables limited by the available data. Seventeen contextual variables covering aspects of the demographic and socio-economic context (five variables) and the caseload and caseflow context (11 variables), plus a variable distinguishing metropolitan from provincial circuits, were included in the modelling. A circuit was categorised as metropolitan if 70% or more of the population was served by a court located in a city or suburban area. Except for the categorical variable distinguishing metropolitan and provincial circuits, the variables were continuous and grand mean centred.

Many of the variables were recognised as likely to be inter-related. For instance adult population, new cases per thousand adults, the count of full time equivalent judicial officers and new cases per full time equivalent judicial officer shared some common underlying values. Other variables were likely to be associated due to well-known linkages. For instance, higher incomes are known to be related to education and hence the median personal income variable was likely to be negatively associated with the proportion of adults with no educational qualifications. Before the bivariate models were run correlation matrices were created and scanned to identify the degree to which the variables were correlated. The matrices revealed that there were high levels of collinearity between the variables that would likely have caused problems if multivariate modelling had been possible. Across 252 correlations between the variables, 83 could be categorised as strong (a correlation co-efficient of .5 or higher); 91 as moderate (.3 or higher but less than .5); and 78 as weak or not material (less than .3). Twenty of the 24 correlations between the Metropolitan/Provincial categorical variable and the other variables were strong (12) or moderate (8). The correlation matrices are located in Appendix G. The high levels of collinearity poses a problem in the interpretation of the results from the bivariate models as it is not possible to distinguish which effects are most important.

Descriptive Statistics

Table 9.1 describes the five demographic and socio-economic variables and seven caseload and caseflow variables that apply generically across the three offences. Table 9.2 presents the offence specific values for the remaining four caseload and caseflow variables. The demographic and socio-economic indicators were based on the 2006 census data and the case related data is from 2008 and 2009. The metropolitan and provincial circuits are contrasted in terms of the other contextual variables in Tables 9.3 and 9.4. The minimums and maximums for each variable demonstrate that there is substantial difference between the circuits at the ends of each spectrum. Across the range of variables the spread of the circuits varies with proportionately larger standard deviations for some variables (e.g. adult population) and smaller deviations for others (e.g. median personal income).

Table 9.1: Generic demographic, socio-economic and caseload and caseflow contextual variables

Circuit Contextual Variables	Mean	SD	Minimum	Maximum
<i>Demographic & Socio-economic Variables</i>				
Adult population	176252	110592	40950	425766
Median personal income	24941	2685	21100	30700
Ethnic homogeneity	0.483	0.087	0.371	0.687
Proportion adults with no educational qualifications	0.233	0.052	0.111	0.312
Proportion of adults unemployed	0.030	0.006	0.018	0.040
<i>Caseload & Caseflow Variables</i>				
Number of Judicial Officers (ftes)	10.6	7.8	2.3	29.7
Number of Judges (ftes)	7.1	4.4	1.7	15.6
New criminal cases per 1,000 adults	128.8	31.2	90.6	209.9
New criminal cases per Judicial Officer (ftes)	2038.5	499.4	992.6	2821.6
Average active cases per Judicial Officer (ftes)	175.3	50.8	114.5	327.0
Cases disposed per Judicial Officer (ftes)	2051.7	491.1	1045.1	2796.2
Hearing rate	0.110	0.036	0.057	0.190

Three of the variables are measures of size (adult population, the number of judicial officers and the number of judges). The inclusion of these variables enabled the modelling to test whether, consistent with the United States research; sentencing was more lenient in larger circuits.

The ethnic homogeneity variable enabled the modelling to test the racial threat hypothesis. The hypothesis predicts that larger racial minorities are perceived as a threat to

the established order and are therefore associated with more severe sentencing. In the United States sentencing research there was support for the hypothesis in some research with greater severity affecting all offenders sentenced not just racial minorities. In this research support for the hypothesis, within the limitations of the modelling, would be indicated by more severe sentencing in less ethnically homogenous circuits.

The personal income, educational attainment and unemployment variables enables the modelling to test the economic threat hypothesis. There was mixed support in the United States for the proposition that greater levels of disadvantage are a threat to the established order and are therefore associated with more severe sentencing.

Assuming that there are differences between metropolitan and provincial circuits in terms of broad politics (a combination of political leanings, values and attitudes) a tentative exploration of the effect of local political setting on sentencing can be undertaken. Assuming that provincial circuits can be broadly labelled as more conservative, sentencing may be more severe in provincial circuits.

The effect of caseload pressures on sentencing are explored through the offence specific variables, the measures of total caseload in terms of flow (new cases and disposal of cases) and stock (cases on hand at a point in time) and trial rates. These measures enable exploration of whether there is evidence that higher caseloads per judge are associated with more lenient sentencing, or that higher trial rates are associated with more severe sentencing.

In addition the models presented here include two other variables, the crime rate and the speed of disposal of cases. The crime rate is measured in three ways. The general crime rate is represented by new cases per 1,000 adults, and the offence specific crime rate is represented by the number of cases sentenced per 1,000 adults and the number of cases sentenced per judge. These measures enable the effect of the prevalence of offending on sentencing variation between circuits to be explored. The two perspectives on the offence specific rate provide a contrast between the prevalence of convicted and sentenced offenders relative to the population compared to a judge centric perception based on the number cases seen by a full time judge. The speed of disposal is measured separately for each of the three offences. Two measures are used, the mean time from charge to sentencing and the proportion of cases sentenced within three months of the charge date.

Table 9.2: Offence specific caseload and caseflow contextual variables

Circuit Contextual Variables	Mean	SD	Minimum	Maximum
<i>Aggravated drink driving</i>				
Sentencings per 1,000 adults	4.31	1.81	1.95	8.10
Sentencings per District Court Judge	104.7	40.3	60.3	193.9
Mean days charge to sentencing	104.4	26.1	71.2	155.6
Proportion of cases sentenced within 3 months of charge	0.572	0.133	0.345	0.760
<i>Male assaults female</i>				
Sentencings per 1,000 males	5.58	2.80	1.99	13.77
Sentencings per District Court Judge	64.1	26.6	27.4	124.8
Mean days charge to sentencing	108.4	38.3	54.7	182.6
Proportion of cases sentenced within 3 months of charge	0.542	0.171	0.238	0.840
<i>Burglary</i>				
Sentencings per 1,000 adults	2.65	0.95	1.16	4.15
Sentencings per District Court Judge	65.27	23.51	33.35	101.79
Mean days charge to sentencing	118.8	33.1	68.5	181.5
Proportion of cases sentenced within 3 months of charge	0.504	0.148	0.254	0.754

The metropolitan and provincial circuits were substantially different. The metropolitan circuits were larger with a less racially homogeneous population, the residents were better educated, had higher personal incomes and higher unemployment rates. Consistent with serving a larger population the metropolitan circuits had more judges and more judicial officers (judges plus community magistrates and justices of the peace).

The caseload and caseflow contrast was more complex. The provincial circuits received more new cases per head of population and per judicial officer and disposed of more cases per judicial officer, but despite having lower rates of new cases the metropolitan circuits had a higher level of active cases on hand per judicial officer. This counter-intuitive juxtaposition may be due to the higher hearing rate for the metropolitan circuits. The higher hearing rate may result in cases progressing more slowly and may indicate that cases are more complex in metropolitan circuits.

The provincial circuits sentenced more offenders for each of the three offences per head of population and per judge than the metropolitan circuits. Cases progressed more quickly from charge to sentence in the provincial circuits.

Table 9.3: Means and standard deviations for the metropolitan and provincial groupings for the generic contextual variables

Contextual Variables	Metropolitan Group		Provincial Group	
	Mean	SD	Mean	SD
<i>Demographic & Socio-economic Variables</i>				
Adult population	263771	83724.5	139786	101530.4
Median personal income	28360	1955.2	23517	1253.2
Ethnic homogeneity	0.568	0.090	0.448	0.058
Proportion adults with no educational qualifications	0.173	0.047	0.258	0.028
Proportion of adults unemployed	0.033	0.005	0.029	0.006
<i>Caseload & Caseflow Variables</i>				
Number of Judicial Officers (ftes)	17.62	10.38	7.68	4.10
Number of Judges (ftes)	11.70	4.54	5.17	2.53
New criminal cases per 1,000 adults	108.81	12.12	137.15	33.25
New criminal cases per Judicial Officer (ftes)	1742.99	600.43	2161.68	418.82
Average active cases per Judicial Officer (ftes)	202.67	77.90	163.88	32.63
Cases disposed per Judicial Officer (ftes)	1788.27	589.35	2161.49	424.08
Hearing rate	0.146	0.031	0.094	0.026

Table 9.4: Means and standard deviations for the metropolitan and provincial groupings for the offence specific contextual variables

Contextual Variables	Metropolitan Group		Provincial Group	
	Mean	SD	Mean	SD
<i>Aggravated drink driving</i>				
Sentencings per 1,000 adults	3.164	0.917	4.782	1.908
Sentencings per District Court Judge	72.54	15.97	118.09	40.05
Mean days charge to sentencing	131.9	23.5	93.0	17.5
Proportion of cases sentenced within 3 months of charge	0.448	0.103	0.624	0.109
<i>Male assaults female</i>				
Sentencings per 1,000 males	4.365	1.593	6.084	3.090
Sentencings per District Court Judge	46.97	13.11	71.24	27.83
Mean days charge to sentencing	154.7	25.1	89.1	22.8
Proportion of cases sentenced within 3 months of charge	0.347	0.095	0.623	0.122
<i>Burglary</i>				
Sentencings per 1,000 adults	1.722	0.490	3.038	0.818
Sentencings per District Court Judge	39.15	7.24	76.15	18.59
Mean days charge to sentencing	158.1	24.6	102.4	19.4
Proportion of cases sentenced within 3 months of charge	0.348	0.095	0.569	0.115

The Multi-level Models

For some of the models there were few contextual variables that were significant on a bivariate basis but for others there were many. The challenge in reporting the outcomes of the modelling was to report in an informative and concise manner. The raw results are often unhelpful because they reflect the units of measurement underlying each variable. In models where the adult population was a significant variable the co-efficient is very small. The size of the co-efficient reflects the fact that it represents the shift in the likelihood of incarceration if the population is one person higher than the mean of 176,252. Conversely, some of the variables are based on proportions varying between zero and one, so that a change of one is a nominal concept and results in a very large co-efficient. The parameter estimates for each variable that was significant are reported but the meaning of those estimates required transformation to make the potential influence of each contextual variable clear. The substantive effect of each variable on the likelihood of incarceration was measured by calculating the odds of incarceration one standard deviation either side of the mean and then calculating the odds ratio for the likelihood of incarceration between those two points. The range one standard deviation either side of the mean covers the central two-thirds of the circuits. The odds ratio between the values one standard deviation from the mean is an indicator of the difference between the circuits, and provides a sense of the degree to which a statistically significant effect is substantively significant. In like manner the substantive effect of the variable on sentence length was calculated to determine the change in the length of sentence one standard deviation either side of the mean and the change between those two points. The ranges one standard deviation either side of the mean are an estimate of the difference between the central two-thirds of the circuits. In the case of the in-out models for aggravated drink driving and burglary the results are tabulated but for the other models the statistics for the small number of significant variables are described in the text.

Aggravated drink driving in out model

Ten of the circuit level variables were significant on a bivariate basis when added to the in-out case level random co-efficient model (Chapter 8: Tables 8.2 & 8.3). The variables are listed in Table 9.5.

Table 9.5: Parameter estimates for circuit level variables individually added to the case level model for the in-out decision for aggravated drink drive sentencing

Circuit Contextual Variable	B	S.E.	p - value
Adult population	-0.000	0.000	.036
Median personal income	-0.000	0.000	<.001
Ethnic homogeneity	-6.658	1.778	.002
Proportion adults with no educational qualifications	13.208	3.071	<.001
Proportion of adults unemployed	-70.883	29.348	.029
New criminal cases per 1,000 adults	0.014	0.006	.028
Hearing rate	-18.814	4.161	<.001
Mean days charge to sentencing	-0.030	0.006	<.001
Proportion of cases sentenced within 3 months of charge	6.170	1.190	<.001
Provincial (Metropolitan)	1.809	0.260	<.001

Note: there were 15 degrees of freedom for all these variables.

The results demonstrate that there is an association between some aspects of the circuit context and the use of incarceration for aggravated drink driving. Taken individually each of the variables appears to have a substantial impact on the likelihood of incarceration between circuits. The small number of circuits and high correlations between the variables precluded the formation of a single model. The substantive effect of each variable is shown in Table 9.6. The odds ratios between the mean and one standard deviation either side of the mean and the odds ratio between the standard deviations are reported.

In terms of demographic and socio-economic factors the odds of incarceration were greater in circuits with smaller populations, with lower median personal incomes, that are more racially homogeneous (i.e. with smaller representation of minorities), where a higher proportion of adults had no educational qualifications, or with proportionately fewer people unemployed. In terms of caseload factors the odds of incarceration were greater in circuits with a higher rate of new criminal cases per thousand adults, with a lower rate of cases proceeding to hearing, or with lower mean days from charge to sentence and a higher proportion of cases sentenced within three months of charge. The degree of difference varied across these factors.

The variation between metropolitan and provincial circuits was significant and substantial. The odds of incarceration for offenders sentenced in provincial circuits were 6.101 times greater than offenders sentenced in metropolitan circuits. The results for all the more substantive contextual factors with significant parameter estimates align with the metropolitan and provincial circuits. The greater likelihood of incarceration in circuits with smaller populations, lower median personal incomes, greater ethnic homogeneity, proportionately more adults with no educational qualifications, lower unemployment, proportionately more new cases, lower hearing rates and faster case resolution are all consistent with the application of these measures to provincial and metropolitan circuits. It is apparent that the broad context of provincial versus metropolitan circuits is important but it is not possible to conclude which of the distinguishing factors is most important or how they combine to produce the overall result.

Table 9.6: Odds ratios one standard deviation from the mean for circuit level variables individually added to the case level model for the in-out decision for aggravated drink drive sentencing

Circuit Contextual Variable	Mean	SD	Range of OR +/- 1SD from the mean		OR for circuits +/-1SD from the mean
			Mean-SD	Mean+SD	
Adult population	176252	110592	1.556	0.643	0.413
Median personal income	24941	2685	1.957	0.511	0.261
Ethnic homogeneity	0.483	0.087	1.785	0.560	0.313
Proportion adults with no educational qualifications	0.233	0.052	0.503	1.987	3.950
Proportion of adults unemployed	0.030	0.006	1.530	0.654	0.427
New criminal cases per 1,000 adults	128.8	31.2	0.651	1.536	2.361
Hearing rate	0.110	0.036	1.969	0.508	0.258
Mean days charge to sentencing	104.4	26.1	2.173	0.460	0.212
Proportion of cases sentenced within 3 months of charge	0.572	0.133	0.440	2.272	5.161

Aggravated drink driving sentence length model

Only one circuit level variable was significant on a bivariate basis when added to the case level random co-efficient model for sentence length (Chapter 8: Tables 8.4 & 8.5). The variable was the proportion of offenders sentenced within three months of the date of charge. Higher proportions of cases resolved within three months were associated with shorter sentences ($B = -2.955$, $S.E. = 1.286$, $d.f. = 15$, $p = .036$). A proportion one full percentage point (.01) above or below the mean was associated with an increase or decrease in sentence length of approximately one day. At one standard deviation from the mean (.133) sentence lengths increased or decreased by 12 days. It is likely that the effect of this variable representing the proportion of cases resolved within three months is associated with the larger proportionate discount associated with early guilty pleas.

There was no significant difference between metropolitan and provincial circuits. The substantial difference in the use of incarceration does not carry over to sentence length. One possible explanation is that the relatively low maximum penalty, two years imprisonment, limits the opportunity for variation. This possibility is partially supported by the limited number of circuits (7) that varied significantly from the reference circuit for the length of incarceration compared to the substantially larger number (15) that varied for the use of incarceration.

Male assaults female in-out and sentence length models

Two of the circuit level variables were significant on a bivariate basis when added to the in-out case level random co-efficient model (Chapter 8: Tables 8.7 & 8.8). The variables were the hearing rate for all offences and the number of male assaults female sentencings per thousand adult males. Both variables had a negative effect on the likelihood of incarceration.

A higher hearing rate for all offences was associated with a reduced likelihood of incarceration ($B = -5.213$, $S.E. = 2.211$, $d.f. = 14$, $p = .028$). At one standard deviation (.036) above the mean the odds of incarceration were 0.822 times those at the mean; conversely at one standard deviation below the mean the odds were 1.216 times greater than the mean. The odds of incarceration for an offender sentenced in a circuit that was one standard deviation below the mean were 1.479 times the odds for a circuit that was one standard deviation above the mean.

The number of offenders sentenced for male assaults female per thousand adult males was associated with a reduced likelihood of incarceration ($B = -0.062$, $S.E. = 0.026$,

$d.f. = 14$, $p = .030$). At one standard deviation (2.8 offenders sentenced) above the mean the odds of incarceration were 0.840 times those at the mean; conversely at one standard deviation below the mean the odds were 1.191 times greater than the mean. The odds of incarceration for an offender sentenced in a circuit that was one standard deviation below the mean were 1.418 times the odds for a circuit that was one standard deviation above the mean.

None of the circuit level variables were significant on a bivariate basis when added to the case level random co-efficient model for sentence length (Chapter 8: Tables 8.9 & 8.10).

Burglary in-out model

As shown in Table 9.7 ten of the circuit level variables were significant on a bivariate basis when added to the in-out case level random co-efficient model (Chapter 8: Tables 8.12 & 8.13).

Table 9.7: Parameter estimates for circuit level variables individually added to the case level model for the in-out decision for burglary sentencing

Circuit Contextual Variable	B	S.E.	p - value
Adult population	-0.000	0.000	0.027
Median personal income	-0.000	0.000	0.046
Proportion adults with no educational qualifications	6.734	1.714	0.001
Number of Judicial Officers (ftes)	-0.034	0.011	0.008
Number of Judges (ftes)	-0.086	0.016	<0.001
Hearing rate	-5.787	2.713	0.05
Mean days charge to sentencing	-0.007	0.003	0.05
Sentencings per 1,000 adults	0.342	0.083	<0.001
Sentencings per District Court Judge	0.017	0.004	<0.001
Provincial (Metropolitan)	0.306	0.093	0.005

Note: there were 15 degrees of freedom for all variables.

These results demonstrate that there is an association between some aspects of the circuit context and the use of incarceration. Taken individually each of the variables appears to have a substantial impact on the likelihood of incarceration. The small number of circuits and high correlations between the variables precluded the formation of a single model. The substantive effect of each variable is shown in Table 9.8 reporting the odds ratios between the mean plus and minus one standard deviation and between the two points one standard deviation either side of the mean.

In terms of demographic and socio-economic factors the odds of incarceration was greater in circuits with smaller populations, with lower median personal incomes, where a higher proportion of adults had no educational qualifications, with proportionately fewer people unemployed, or with a lower complement of judicial officers or judges. In terms of caseload and caseflow factors the odds of incarceration was greater in circuits where the rate of cases proceeding to hearing was lower, where cases proceeded to sentencing more quickly (the mean days from charge to sentence were lower), the rate of burglars sentenced per thousand adult population was higher, or the rate of burglars sentenced per judge was higher.

Table 9.8: Odds ratios one standard deviation from the mean for circuit level variables individually added to the case level model for the in-out decision for burglary sentencing

Circuit Contextual Variable	Mean	SD	Range of OR +/- 1SD from the mean		OR for circuits +/-1SD from the mean
			Mean-SD	Mean+SD	
Adult population	176252	110592	1.248	0.802	0.643
Median personal income	24941	2685	1.230	0.813	0.661
Proportion adults with no educational qualifications	0.233	0.052	0.705	1.419	2.014
Number of Judicial Officers (ftes)	10.6	7.8	1.303	0.767	0.589
Number of Judges (ftes)	7.1	4.4	1.459	0.686	0.470
Hearing rate	0.110	0.036	1.232	0.812	0.659
Mean days charge to sentencing	118.8	33.1	1.240	0.806	0.650
Sentencings per 1,000 adults	2.65	0.95	0.723	1.384	1.915
Sentencings per District Court Judge	64.1	23.51	0.667	1.499	2.248

The variation between metropolitan and provincial circuits was significant and substantial. The odds of incarceration for offenders sentenced in provincial circuits were 1.358 times greater than offenders sentenced in metropolitan circuits. The results for all the other contextual factors with significant parameter estimates align with the metropolitan and provincial circuit variable. The greater likelihood of incarceration in circuits with smaller populations, lower median personal incomes, proportionately more adults with no educational qualifications, fewer judicial officers, lower hearing rates, faster case resolution, and more sentencings per adult and per Judge are all consistent with the application of these measures to provincial and metropolitan circuits. It is apparent that the

broad context of provincial versus metropolitan circuits is important but it is not possible to conclude which of the distinguishing factors is most important or how they combine to produce the overall result.

Burglary sentence length model

Three of the circuit level variables were significant on a bivariate basis when added to the case level random co-efficient model for sentence length (Chapter 8: Tables 8.14 & 8.15). The variables were median personal income, the proportion of adults with no educational qualifications and the metropolitan/provincial split.

Higher levels of median personal income were associated with shorter sentences ($B = -0.001$, $S.E. = 0.000$, $d.f. = 15$, $p < .001$). At one standard deviation from the mean (\$2,685) sentence lengths increased or decreased by 51 days. A higher proportion of adults with no educational qualifications was associated with longer sentences ($B = 23.995$, $S.E. = 8.214$, $d.f. = 15$, $p = .011$). At one standard deviation from the mean (.052) sentence lengths increased or decreased by 37 days.

Provincial circuits imposed sentences 3.5 months longer than metropolitan circuits ($B = 3.490$, $S.E. = 0.831$, $d.f. = 15$, $p < .001$). This is a notable finding because offenders are also more likely to be incarcerated in provincial circuits. The significant finding for burglary but not for aggravated drink driving may be attributable to the much higher maximum penalty for burglary (10 years imprisonment) compared to aggravated drink driving (2 years imprisonment).

Discussion

The incorporation of contextual variables into the models adds an extra layer to the prior findings. The significant correlations between contextual variables and both incarceration rates and sentence lengths provide insights into the differences between the circuits that may explain the sentencing variations between circuits observed in the earlier single level and random effects models. However, the results have to be treated with caution because of the need to construct separate models for each variable, and the high levels of multicollinearity between the contextual variables.

The results vary between the models. In two models (aggravated drink driving and burglary in-out models) ten contextual factors are identified that may be linked to variation in sentencing between circuits. At the other end of the scale no variables were identified in the sentence length model for male assaults female. Limited numbers were identified in

the other three models: burglary sentence length (3); male assaults female in-out (2); and aggravated drink driving sentence length (1). The more limited results for the sentence length models is consistent with the United States research which also tended to find more associations between contextual variables and categorical decisions like sentence type or granting of a departure than for decisions represented by a continuous variable like sentence length or the amount of a departure. In addition, the difference between male assaults female in-out model and the in-out models for the other two offences is likely to have been influenced by the constraints on modelling that offence.

Despite these limitations there are a series of observations that can be drawn from the models. Consistent with the United States research the results are mixed and in two instances contradictory. There is evidence of an association between size and sentencing variation between circuits in the in-out models for aggravated drink driving and burglary, but there were no effects in the other models. Larger populations (both models) and larger complements of judicial officers and judges (burglary) were associated with greater leniency. This is consistent with the United States research.

The models provided no support for the proposition that circuits with larger ethnic minorities would be more severe. In the only model that showed a significant effect, the aggravated drink driving in-out model, offenders were sentenced more leniently in circuits that were more ethnically diverse (where ethnic minorities were larger). This is out of step with three of the United States studies but supportive of one (Crow & Gertz, 2008). It may be explained, as Crow & Gertz (2008) suggested, by a threshold at which the effect reverses, when minorities become sufficiently large they are able to exercise political power and effect changes to policy and practices. Alternatively it may reflect a difference in race relations between New Zealand and the United States. In New Zealand greater diversity may promote greater tolerance and that this extends to tolerance for criminal behaviour.

The aggravated drink driving in-out, burglary in-out and burglary sentence length models all provide support for the proposition that greater economic disadvantage will be associated with more severe sentencing. This effect was observed for lower personal incomes and lower educational achievement in all three models. However, the reverse effect occurred for unemployment in the aggravated drink drive in-out model, where higher unemployment was linked to more lenient sentencing between circuits.

None of the three general measures of caseload pressure (new cases per judicial officer, disposals per judicial officer, or active cases per judicial officer) were associated with sentencing variation between circuits. However trial rate was negatively correlated with sentencing variation between circuits in the in-out models for all three offences. This is counter to the expectations tested in the United States where higher trial rates were expected to result in more severe sentences due to the application of a trial penalty. The United States research found no significant association between trial rate and sentences.

The three measures of crime rate all had some effect. Aggravated drink drivers in circuits with a higher general rate of offending (criminal cases per head of population) were more likely to be incarcerated. Burglars in circuits with higher rates of burglary (whether expressed as the rate of burglars sentenced per head of population or the rate sentenced per judge) were also more likely to be incarcerated. However offenders sentenced for male assaults female in circuits with higher rates of male assaults female offenders sentenced per head of population were less likely to be incarcerated.

The speed of disposal also had some effect. Whether expressed in terms of mean days from charge to sentence or the proportion of cases sentenced within three months of charge, speedier case disposal was correlated with a higher likelihood of incarceration for aggravated drink driving and burglary. However, the proportion of cases sentenced within three months of charge resulted in shorter sentences for aggravated drink driving. This contradictory effect is probably due to greater sentencing discounts for guilty pleas the sooner the sentence is imposed after the charge is laid.

The discussion of the metropolitan/provincial variable has been held over to the end because it not only addresses the political context but also brings a perspective to the other significant variables. Sentencing was more severe if an offender was sentenced in a provincial circuit in three models, the in-out models for aggravated drink driving and burglary and the sentence length model for burglary. These findings lend support to the proposition that more conservative political attitudes in provincial areas may influence sentencing approaches resulting in more severe sentencing. Conversely greater tolerance and acceptance of diversity in metropolitan areas may foster more lenient sentencing. The differences between the two categories of circuit were also consistent with the effects for the more detailed contextual variables. The difference may be substantially explained by any one or some combination of those variables.

Summary

There is evidence that the circuit context is associated with variations in sentencing between circuits. However, the evidence of the association with the likelihood of incarceration (Research Question 1) and sentence length (Research Question 2) is mixed, and varied by offence type. The modelling was more revealing in respect of the likelihood of incarceration for burglary and aggravated drink driving. In both instances many of the variables were individually significant but the limited number of circuits precluded the formulation of multivariate models. This would have had limited usefulness in any event because of the high levels of multi-collinearity.

The strongest inference that can be drawn is that a difference or differences in the contexts of metropolitan and provincial circuits condition the formation of local sentencing approaches for aggravated drink driving and burglary. However, a more detailed model in which the demographic, socio-economic and caseload or caseflow factors condition the differences between metropolitan and provincial circuits cannot be advanced.

Chapter 10: Summary of findings, discussion and conclusion

In 2006 the New Zealand Law Commission recommended the establishment of a Sentencing Council to develop sentencing guidelines. The recommendations were controversial. Despite legislation establishing a Sentencing Council passing in 2007 the legislation has not been implemented. The crux of the guidelines debate revolved around the adequacy of the current sentencing regime to deliver consistent sentences and evidence purporting to show that sentencing under the regime is inconsistent.

The Law Commission took the view that inconsistency is virtually inevitable in a sentencing system with very wide sentencing discretion. The argument was backed by an assertion of widespread anecdotal evidence of inconsistency and a comparison of substantial differences in sentencing outcomes for high volume offences between regions of the District Courts. Opponents of guidelines rejected the argument that discretion leads to inconsistency. They argued that discretion was needed to achieve fair treatment and just outcomes across the wide diversity of offending and offenders appearing for sentence. As a corollary they argued that the observed differences from the Law Commission study were attributable to differences in the composition of caseloads between regions.

The debate is largely unresolved but has gone into abeyance following the decision not to implement the legislation. The principal aim of the current research was to examine the competing positions in an attempt to resolve or at least advance the debate. Before embarking on a more rigorous examination of the evidence of inconsistency or the identification of compositional differences between regions, a framework for understanding sentencing disparity was developed. The framework tracked how sentencing discretion could result in inconsistency of sentencing despite the operation of sentencing laws, appellate authorities and the moderating effect of sentence appeals.

The framework has been likened to a chain linking discretion to inconsistency. Sentencing frameworks aim to bring consistency to sentencing outcomes but at the same time afford differing degrees of discretion to judges. Discretion is allowed because of the challenges of sentence decision making to enable sentences to be fitted to circumstances. When decision makers are confronted with the challenges (complexity, incomplete information, uncertainty and time pressure) facing judges they are limited by “bounded rationality” and restricted to “satisficing” or making good enough decisions (Simon, 1976). In response decision makers, in this case judges, develop heuristic strategies or rules of thumb, to simplify and speed up decision making. Sentencing discretion enables concerns

outside formal sentencing frameworks to influence outcomes, sentencing is a balance between formal rationalities based on the law or guidelines and substantive rationalities based on concerns, ideologies, attitudes, biases and individual interests (Savelsberg, 1992). It follows that if sentencing is a balance of rationalities that the content of any sentencing heuristic can be categorised into formal rationalities and substantive rationalities. These concepts explain how sentencing might vary between similar cases and between judges but it does not explain variation between locations. Court community theory (Eisenstein & Jacobs, 1978; Eisenstein et al., 1988) and the commonality and sociality corollaries from personal construct psychology (Ostrom et al., 2004) bridge the gap. Court community theory suggests that sentencing is influenced by the concerns and attitudes of the local courthouse community (the judges, prosecutors and defence counsel who regularly appeared in the same courtrooms) (Eisenstein & Jacob, 1977) and by the concerns and attitudes of the wider community the court served (Eisenstein, Flemming & Nardulli, 1988). The commonality corollary suggests that those who share a common environment are likely to make sense of it in similar ways. Sentencing approaches converge between the judges in a locality because they share a similar set of experiences and observations. Convergence is reinforced by the “sociality corollary” with those who work together on a regular basis sharing their experiences and approaches and anticipating and adopting the behaviour of others.

A theoretical explanation is just that, a theory about how the system might work. It does not of itself establish that the expected outcomes are actually manifested. Based on the theory and the findings of research from the United States four predictions and two research questions were examined to test whether practical sentencing outcomes conformed to the theory. The empirical testing focussed on the circuits of the District Courts and was undertaken in three stages. The first stage addressed the two principal predictions: that after controlling for seriousness of offence, criminal history and offender demographic variables similar offenders will be sentenced differently between circuits. The second stage addressed two subsidiary predictions: that the weights applied to some or all of the seriousness of offence, criminal history or other variables will vary between circuits. The subsidiary predictions were reliant on the support for the principal predictions and an assumption that seriousness of offence and criminal history variables will be the principal influences on the determination of sentence. In the third stage the association between contextual variables at circuit level and sentencing variation between circuits was explored in an attempt to identify contextual variables that may have an

influence on sentencing approaches. Each stage divided into two predictions or research questions because the research separately examined two distinct sentencing decisions: the use of incarceration and the length of incarceration.

The sentencing of aggravated drink driving, males assaults female and burglary offending by resident judges in the 17 circuits of the District Courts was separately examined for each prediction and research question. Differing forms of regression modelling appropriate to the sentencing decision and prediction or research question were employed.

The predictions were tested using a dataset specially assembled for this research. The base data was sourced from the Justice Data Warehouse, a store of raw data on offenders, charges, and charge outcomes maintained by the New Zealand Ministry of Justice. The data in the warehouse was drawn from the Case Management System, an administrative system supporting the courts. The base data was transformed into a set of cases for each offence. Each case included information on the offender, information on the case, the sentence imposed, and the offender's criminal history.

The research questions were examined using the case level data plus a series of circuit level contextual variables based on: socio-demographic data from the 2006 census; reports of caseflow and judicial sitting time; and measures of timeliness of case disposal derived from case dataset.

The present study advances the prior research in several ways. In a New Zealand context it is the first study to examine consistency of sentencing between locations (circuits). In the international context it is believed to be the first substantial study of the consistency of sentencing outside the United States. It is distinct from contemporary United States research because it examines sentencing in a non-guideline system. It differs from most studies because it looks separately at three diverse offences before drawing together the results thus enabling observation of any nuances that may be hidden in more broadly based research.

Like almost any study using statistical models the findings are sensitive to gaps in the models due to an absence of information resulting in variables being omitted from the analysis. All of the models lack information on: positive post-offending behaviour (proactive steps toward reform); negative post-offending behaviour (e.g. attempts to influence witnesses); family responsibilities; employment status; occupation; or whether the offender was remanded in custody or was already a serving prisoner (these circumstances limit the choices available to the judge irrespective of the sentencing merits

of the case). It is unlikely that the inclusion of controls for any of these generic variables factors would make a substantial difference to the differences between circuits. To do so they would have to be disproportionately distributed across the circuits and the distribution would have to have a systematic relationship with incarceration rates or going rates for sentence lengths.

Some offence specific variables were also omitted. This was most significant for male assaults female. The models do not have any variables representing the level of violence or the degree of injury or other harm. The burglary models had no variables representing the degree of planning or premeditation, whether the premises were occupied, or any damage greater than that required to complete the burglary. The aggravated drink driving models only partially accounted for driving behaviour (via the variables on other driving charges) but omitted any recognition of bad driving behaviour presented as an aggravating factor but not the subject of a separate charge.

It was intended that these variables would be included in the sentence length models supported by information extracted from sentence notes. However, as explained (Chapter 5 & Appendix D), problems with the uneven and non-random availability of sentence notes across the circuits precluded their use. The sentence note data was analysed to gain an indication if inclusion of the variables would substantially improve each model and whether it would make a material difference to the estimation of the parameters for the circuit variables. The results are presented in Appendix H. The additional variables contributed to a material improvement to the male assaults female and burglary models but made little difference to the aggravated drink driving model. The findings confirmed that the variables were relevant to the judges. More importantly, however, there was no indication of a material change to the estimation of the parameters for the circuit variables. The lack of effect on the circuit variables supports a tentative conclusion that the omitted variables were relatively evenly distributed across the circuits.

The balance of this chapter addresses the research findings, the policy implications arising from the findings and future research.

Inconsistency of Sentencing between Circuits

The single level models demonstrate that the likelihood of incarceration and length of sentence vary between circuits for all three offences. The findings strongly support Prediction 1 (likelihood of incarceration) and Prediction 2 (length of incarceration): after controlling for seriousness of offence, criminal history and offender demographic variables, similar offenders were sentenced differently between circuits. The differences originally observed by the Law Commission (2006) cannot therefore be dismissed as the result of compositional differences between locations.

The results from the regression models varied between the three offences and between the sentencing decisions; the patterns of inconsistency are not a constant phenomenon irrespective of the offence. More circuits had significantly different incarceration rates than the reference circuit for aggravated drink driving (14) and burglary (13) than male assaults female (7). More circuits had significantly different sentence lengths than the reference circuit for burglary (14) than aggravated drink driving (7) and male assaults female (6). While similar numbers of circuits were significantly different for both decisions for burglary (13 & 14) and male assaults female (7 & 6) there was a substantial difference for aggravated drink driving (14 & 7).

The substance of the differences also varied between the offences and decisions. In almost all instances statistically significant differences between a circuit and the reference circuit were substantial. For aggravated drink driving the rates of incarceration ranged from 60 to 640% higher and one circuit was 40% lower⁸. For burglary the rates of incarceration ranged from 80 to 150% higher. For male assaults female the rates of incarceration ranged from 70 to 130% higher and one circuit was 60% lower. For aggravated drink driving the length of incarceration ranged from 15.4 to 33% longer⁹. For burglary the length of incarceration ranged from 21.2 to 43.2% longer and one circuit was 18.5% lower. For male assaults female the length of incarceration ranged from 1.5 to 52.2% longer. It was only male assaults female sentence length differences that were insubstantial, as four of the six significantly different circuits were less than one month higher than the reference circuit. The differences in the rate of incarceration were proportionately greater than the differences in sentence length.

⁸ The proportionate differences were calculated from the relative risks in Chapter 6.

⁹ The proportionate differences for each offence were calculated based on the unadjusted mean sentence lengths in months: aggravated drink driving (9.1); burglary (14.6); and male assaults female (6.7)

Furthermore the rank order of the circuits in one model was not a good predictor of a rank order in a related model (for the same decision between offences or between different decisions for the same offence). In only two of the nine pairings was the rank order in one model a good predictor of the other. The two were rates of incarceration for aggravated drink driving and burglary, and rate of incarceration and sentence length for burglary.

The summation of the results by offence and decision indicates that the balance of formally rational considerations and substantively rational sentencing concerns varies between offences and decisions. This suggests that judges do not share a common view of the proportionate relationship between offences across the circuits. This is important because it suggests that any steps to increase the levels of consistency need to occur offence by offence rather than by generic change.

As expected the models demonstrated that the circuit (location) was not the major determinant of sentence. In all of the models offence seriousness and criminal history variables were the principal determinants of sentence. The seriousness and criminal history factors dominated the improvement in prediction of incarceration, with 73.6% (aggravated drink driving), 91.8% (burglary) and 95.5% (male assaults female) of the improvement attributable to these factors. Seriousness and criminal history factors accounted for the bulk of the variation in sentence length explained by the linear regression models, with 69.4% (male assaults female), 88.9% (aggravated drink driving) and 92.4% (burglary) of the variation attributable to these factors. The extent to which sentence length was dependent on where an offender was sentenced varied between the offences.

The estimates of the effect of gender and ethnicity on sentencing added to the understanding of sentencing and also varied between the offences. In five of the models neither variable had a significant influence on sentencing. The exception was the decision to incarcerate burglars, with male, Māori and Pacific offenders more likely to be incarcerated than similarly situated female or European offenders. The models with no significant influence for gender and ethnicity variables indicate that the local substantively rational sentencing concerns are expressed via the offence seriousness and criminal history variables. The judges are able to reach sentencing decisions based on legal factors; they do not resort to extra-legal factors because they cannot reach a conclusion. These decisions are not significantly influenced by causal attributions (Albonetti, 1991) or perceptual shorthand (Steffensmeier et al., 1998). While legal factors dominate the determination of sentence for burglary there is evidence that judges have greater difficulty in deciding how

to sentence burglars who are on the cusp of incarceration or who might warrant a more lenient sentence due to the prospect of rehabilitation and reduced likelihood of re-offending.

The support for the predictions is consistent with the United States research, which has almost universally found that net of controls sentencing varies between locations. The predominant role played by offence seriousness and criminal history factors is also consistent with the United States research. The more pervasive role of gender and ethnicity in the determination of sentences in the United States was not matched in this study. More detailed comparisons beyond these high level findings were not possible due to differences in the specifications of the models employed here and the United States research. Because many of the United States studies have covered a range of offence types they may have “averaged” the effect of location on sentencing because the locations are not uniformly more or less severe across all the offence types. In separately analysing individual offences, this study avoids the averaging effects due to inclusion of multiple offence types and may therefore have been able to identify a broader range of differences across locations. The focus on individual offences in this research has allowed the modelling to be based on sentencing factors specifically chosen as relevant to that offence, thus potentially heightening the likelihood of differentiating one case from another. Some United States studies have employed more generic factors that may result in less recognition of differences between cases.

The results conform to expectations based on the framework for understanding sentencing decisions. The effects of seriousness and criminal history variables are a tangible representation of the influence of formally rational considerations derived from the law and appellate authority. The predominant influence of these factors indicates that the formally rational framework established in the law and appellate authorities is the core base for determination of sentences across the circuits. Sentencing can be seen as a balance between formal and substantive rationalities. The differences between circuits suggest that the substantive rationalities are embedded in local conditions.

The significant differences between the circuits are evidence that the formally rational framework is not sufficient to achieve consistency and that sentencing is influenced by substantively rational sentencing concerns whose influence varies between circuits. As predicted by Savelsberg (1992), judicial discretion allows for sentencing variation. At one level this is precisely the purpose of discretion, it is intended to enable sentences to be crafted according to the particular facts of the case and the circumstances

of the offender. Individualised sentencing is highly prized. However, it is not intended that it should result in marked differences in the sentencing of similar cases for offenders in similar circumstances. The finding of marked differences between circuits supports the propositions made by Ostrom et al. (2004) that the sentencing constructs of judges who share a common environment are likely to be similar and that the similarity will be increased to the extent that the constructs are shared. Through these means distinct local sentencing norms and ‘going rates’ are formed.

Variation between Circuits in the Interpretation and Use of Sentencing Factors

Having established the predicted variation in sentencing between circuits and the expected role of offence seriousness and criminal history variables as the principal determinants of sentence the modelling moved to the second stage. The multi-level random co-efficient models demonstrate variations between the circuits in the weight given to a range of variables for the determination of sentences. The findings support Prediction 3 (likelihood of incarceration) and Prediction 4 (length of incarceration): the weight applied to some offence seriousness, criminal history or other variables differed between circuits. The level of support varied substantially between the offences. The aggravated drink driving and burglary models offered strong support but the male assaults female models provided partial support and need to be treated with caution.

When determining whether an aggravated drink driver should be incarcerated the differences in the sentencing approaches between the circuits was located in differences in the weights placed on variables indicating the seriousness of the current drink driving offending or the offender’s history of drink driving. Variables directly addressing drink driving seriousness and criminal history were also significant among the variables whose weight differed between circuits in the determination of sentence length for aggravated drink driving. In addition circuits differed in the weight placed on the overall seriousness of the wider case when determining sentence length. The criminal history variables whose influence varied between circuits differed between the two decisions. The variables of interest for the incarceration decision were the number and timing of prior drink drive convictions whereas treatment of the number of prior convictions and the longest prior term of imprisonment for drink driving varied between circuits in the determination of sentence length.

Circuits differed in the weight placed on the seriousness of the wider case, the speed of resolution, criminal history for burglary and serious dishonesty and recent

convictions for breaches of sentence when determining whether a burglary offender should be incarcerated. The range of variables with different weights for the determination of sentence length differed from the incarceration decision. The only common variable was prior convictions for serious dishonesty. Circuits differed in the weight placed on the number of current burglaries being sentenced and whether the offender was also being sentenced for a breach of sentence. Circuits also differed in the weight placed on the longest prior term of imprisonment for burglary. It is notable that there was no difference between circuits in the weight placed on gender or ethnicity in deciding whether to incarcerate a burglar. This indicates that the effects found in stage one are similar across the circuits with males, Māori and Pacific offenders more likely to be incarcerated than similarly situated females or European offenders irrespective of circuit.

Circuits differed in weights placed in two variables when determining whether an offender should be incarcerated for male assaults female and a single variable in the determination of sentence length. Circuits placed differing weights on prior convictions for male assaults female and the offender's wider criminal history (as evidenced by a prior incarceration for any offence) when deciding whether to incarcerate. Circuits placed differing weight on the way a charge was laid in the determination of sentence length. In general indictable charges are laid for more serious offending; consequently the way a charge is laid can be interpreted as a measure of seriousness.

The findings from the male assaults female models may be due to several limitations. The lower incarceration rate for the offence may restrict the degree of difference when the sentencing decision is cast this way. This effect may have been exacerbated if a substantial proportion of incarcerations were due to restriction of sentencing choices because the offender spent time remanded in custody or was already a serving prisoner. The real incarceration rate representing the exercise of discretion may have been lower. The narrowness of male assaults female as just one slice of the spectrum of physical assaults may have limited the opportunity for ready identification of differences in approach. This effect may have limited the sentence length model. The sentence length model may also have been limited by the smaller number of observations.

The models add depth to the understanding of the influence of formal and substantive rationalities. First, they demonstrate, consistent with theory, that the weight given to formally rational variables can and does vary between locations. The absence of any random effects for either the gender or the ethnicity factors in the in-out model for burglary demonstrates that substantive rationalities can apply generally across locations

and do not always vary between locations. When the influence of substantive rationalities is expressed through formally rational sentencing factors the effect is indirect and largely hidden. Judges appear to be expressing their sentences in similar terms having followed a consistent framework but arrive at a different sentence. A focus on the framework may therefore lead to an erroneous impression that sentencing is consistent. Had the models only shown fixed effects it would have been inferred that the inconsistency was associated with other variables not represented in the models. It is probable nonetheless that a portion of the inconsistency is due to variables that were not modelled.

The models for aggravated drink driving and for burglary each identified four or five factors that were applied differently across the circuits in determining whether to incarcerate and how long a sentence of incarceration should be. For both offences the variables were a subset of those indicated by the models as most influential in determining the sentence. In contrast the male assaults female models were more limited, identifying one variable with random effects for the determination of sentence length and two for the use of incarceration. The variables were not the most influential variables in the models.

Further discussion can only be hypothetical. The differences between circuits in the sentencing constructs for aggravated drink driving and burglary are suggestive of differences in the sentences required to hold an offender to account, or the need to protect the community from the offender, or a combination of both. The differing weights placed on criminal history variables in five of the six models suggests that variation in the custodial threshold between circuits is at least partially based on differences of view about what characterises a recidivist offender, and in the uplift in sentence warranted for past offending. Alternatively, sentencing approaches may differ because views of the seriousness of the offence type differ, with the difference practically expressed in differences in the likelihood of incarceration and lengths of incarceration in similar cases. Sentencing approaches may also vary because of localised differences in views on the best way to deal with some types of offending. The establishment of Family Violence Courts in some locations and not others are a practical example of this effect.

The models indicate that there appears to common agreement on what factors to take into account but there is less agreement on the significance of the factors for the incarceration decision and the determination of the length of sentence.

Associations between Circuit Context and Sentencing Variation

The third stage of this study attempted to identify some of the contextual differences between circuits that may have influenced the differences in approach. The limited number of circuits severely limited the exploratory modelling that could be undertaken, with a separate model having to be estimated for each contextual variable. The high levels of multi-collinearity between the contextual variables limited any conclusions that could be drawn.

Some effect was found for many of the variables that were modelled. Significant effects were found in one or more models for: all five demographic and socio-economic variables; four of the seven generic caseload and caseflow variables; all four offence specific variables; and for the variable distinguishing metropolitan and provincial circuits.

The models demonstrate that whether offenders were sentenced in a metropolitan or provincial context was significant. Aggravated drink drivers and burglars were more likely to be incarcerated and burglars to receive longer sentences in provincial circuits. This was the most significant finding because the metropolitan and provincial circuit categories were distinguished from one another by many of the other variables that had significant bivariate effects. These variables provide some limited insight into the more specific aspects of the two types of circuit associated with sentencing variation between circuits. The weight that can be placed on any particular variable is limited by multi-collinearity between the variables. The differences may be due to a complex combination of the variables or one variable may be dominant.

Two of the considerations raised in the research literature concerning size resonate as partial explanations for more lenient sentencing in metropolitan circuits. More routine sentencing is said to be likely to be less visible in metropolitan (larger) circuits, it is less likely to come under scrutiny and less likely to be subject to criticism for being soft or to calls for judges to be more punitive. Secondly, metropolitan (larger) circuits are likely to attract a more diverse population with attendant wider range of behaviour, including more diverse deviant behaviour, promoting greater tolerance generally and in sentencing in particular. Greater diversity is one of the distinguishing features separating metropolitan and provincial circuits. Under a racial threat hypothesis this would result in more severe sentencing. The modelling results did not support the racial threat hypothesis. The results might indicate however, that the proportionate size of ethnic minorities has passed a threshold as suggested by Crow and Gertz (2008). In combination these factors all point to the possibility that the social politics of residents in metropolitan circuits may be more

liberal and permissive than in provincial circuits which may exhibit more conservative attitudes including a lower tolerance of deviance and being tougher on law and order.

The influence of distinctions in the social politics and socio-economic make-up of metropolitan and provincial circuits may explain the counter-intuitive results found for caseload and caseload factors in the bivariate analyses of effects. For instance, faster resolution of aggravated drink driving and burglary cases was associated with higher use of incarceration. Ordinarily faster resolution would indicate taking of responsibility, remorse and an earlier guilty plea and would be expected to be associated with more lenient sentencing. Conversely, a higher trial rate for all offending in general was associated with lower use of incarceration for all three offences. Ordinarily going to trial suggests the offender has not taken responsibility, is not remorseful, and does not attract any benefit from pleading guilty and would be expected to be associated with more severe sentencing. In this research faster resolution was found in the provincial circuits and higher trial rates were found in the metropolitan circuits. These two findings are linked, more cases going to trial should, all other things equal, result in slower resolution of cases. However, the expected effects on sentence severity were not demonstrated. This strongly suggests that the influence of the views, preferences and attitudes of the local courthouse community and wider local community differs between metropolitan and provincial circuits and outweighs the influence of caseload and caseload factors. This may have flow-on effects with defendants in provincial circuits pleading guilty at higher rates to take advantage of the mitigation effects of an early plea to avoid more severe sentences whereas defendants in metropolitan circuits may be less concerned about sentences and more willing to go to trial.

Sentencing Heuristics

The research enables some limited observations to be made about the operation and content of sentencing heuristics. The first is that separate heuristics apply to the determination of sentence type, in this case whether to incarcerate or not, and sentence length. The separation is evident in the content of the single level models addressing consistency of sentencing; different combinations are used to decide on the two components of the sentence. The distinction between the two decisions is consistent with the operation of a planning heuristic; the determination of sentence is broken into separate decisions.

The second observation is that the content of the heuristics differs between circuits. This is evident in the random effects models that demonstrate differences in the interpretation and use of a range of significant sentencing factors. The models provide an indication of the core variables that lie at the heart of the differences in the heuristics applied in the circuits but do not give any indication of the spread of the differences or the particular nature of the differences. The best means of extending understanding of the sentencing heuristics would be to discuss the research with the judges and to seek from them a description of the influence of the factors on the custodial threshold and the contribution the factors make to the determination of the quantum of the sentence.

The more severe sentencing in provincial circuits compared to metropolitan circuits is a clear indication of differences in heuristics between those two groupings. However, it does not reveal whether the heuristics are similar within each grouping or whether they just result in more similar sentencing outcomes that are arrived at in different ways. Discussion of the results of the research and description of the influence of the various factors on the two sentencing decisions may elucidate greater insight into the degree of similarity between the two groupings. It is possible that the degree of similarity will vary by offence. The wider spread of odds ratios for aggravated drink driving compared to the clustering of odds ratios for burglary suggests this may well be the case.

The Policy Implications

The findings of this research have significant implications for the judiciary and for sentencing policy makers. The current sentencing framework seems to have worked effectively to inform judges about what factors to take into account. This is evidenced by the significant role offence seriousness and criminal history factors play in sentencing. However, the statutory framework and appellate authority have not been sufficient to avoid substantial inconsistency between circuits by bringing about common sentencing approaches. In other words, judges use relevant formally rational sentencing considerations as the basis for sentencing decisions but their use is influenced by substantively rational sentencing concerns based on local conditions. The sentencing discretion seeking to enable individualisation of sentences has had the unintended consequence of allowing systematic differences to arise between circuits. This situation has the potential to bring sentencing and justice into disrepute if it is not corrected. What is needed is a solution that rebalances the influence of formal and substantive rationalities

but does not replace undue flexibility with undue rigidity preventing individualised sentencing in any circumstance.

The discussion of contextual differences between circuits and the distinction between metropolitan and provincial circuits provides an indication of the possible factors underlying the differences in policy between circuits but does not offer direct options for change. None of the contextual factors are amenable to adjustment to achieve greater consistency. Although this may seem self-evident it is important because it points to the need to find a solution by addressing the guidance given to judges when dealing with individual cases.

The observation of statistically and substantively significant variations between circuits in the use and length of incarceration supports the theory that inconsistency can be expected in a sentencing regime that features wide discretion, with guidance primarily focussed on process but not on substantive sentencing. The Sentencing Act 2002 provides the judges with general guidance on the process of sentencing, providing purposes and principles to guide the sentencing process and lists of factors to take into account, but it offers little guidance on how to transform the seriousness of offending, the offender's culpability and the offender's personal characteristics into an actual sentence. In the absence of substantive guidance on the custodial threshold or the weight to be given to sentencing factors in determining the length of sentence, judges form personal sentencing constructs which converge into circuit level constructs.

The offence based models employed in the current research strongly indicate that any additional guidance needs to be offence based. The variation in the relative influence of different categories of sentencing factors suggests that generic guidance is impractical.

In the contemporary context there are two options for the provision for more detailed offence based sentencing guidance. The first option is to extend the coverage of guideline judgments to a wider range of offending moving away from the focus on the most serious offences. The second option is to revisit the Sentencing Council and implement the legislative framework for the Council and the development of sentencing guidelines.

The Sentencing Council option has a number of advantages over an extension of guideline judgments. As Chambers J said, the guideline judgment process does not enable the depth of research or consultation that accompanies other policy making processes (Hall, 2008). If the Sentencing Council guideline development process was implemented as it stands in the legislation the guidelines would be subject to public consultation. It is

likely that a Sentencing Council would seek and receive feedback from the legal profession; the Court of Appeal cannot do this. The Sentencing Council would include representation from the High and District Courts. This is important because it includes sentencing judges who can bring a direct and contemporary experiential voice to the process.

The Court of Appeal has a limited capacity to issue guideline judgments; it must give attention to the individual appeals made to it in a timely fashion. A Sentencing Council is focussed on the issuing of guidance, although part of its focus may be on the determination of guidelines for the release of prisoners on parole.

To be effective the guideline judgment procedure would need major changes and would likely require legislative backing to enable a process that mirrors the advantages of a Sentencing Council. Even with changes to achieve this, to enable the Court to issue guidelines free of any appeal, to consult and consider responses, and to monitor sentencing, it would require an increase in judges and support staff and there would still be a residual difficulty. A continuation and extension of guideline judgments leaves responsibility for the development of most of the content of sentencing policy with the Judiciary. Currently the Judiciary decide where the line separating community sentences from incarceration lies and the appropriate quantum increases in sentence lengths for incarcerated offenders. The Executive (via Parliament) and Parliament can only exercise crude control over policy either by providing general guidance or by imposing more rigid set sentences such as those entailed by the three strikes regime. There is no good reason for judicial control of sentencing policy. The correction of perceived deficiencies in judicial policies by broad brush crude changes is undesirable and likely to result in a new form of inconsistency with one size fits all sentences resulting in vastly different offending and offenders being treated as if they are the same. It would be preferable to have a mechanism that allows the Government of the day and Parliament to exercise more direct control over sentencing policy. A guideline process that requires guidelines to be ratified by Parliament can meet this requirement.

It would be optimistic to expect that all guidelines will achieve consistency or achieve the sentencing outcomes expected by the framers of the guideline. To avoid the pitfalls of one size fits all and the practical difficulty of catering for all cases in the guidelines, judges must have sufficient sentencing discretion to impose sentences outside the guidelines in individual cases but not to formulate distinct sentencing policies that conflict with the guidelines. To cater for more exceptional cases such as those warranting

mercy or a more severe sentence to protect the community a departure regime is required. The American experience is testimony to the reality that a guideline regime with some sentencing discretion allows inconsistency.

There are however strategies that can be followed to minimise any resultant inconsistency. The Sentencing Council must have the capability to monitor sentencing patterns and to review the working of individual guidelines. It must provide feedback on sentencing patterns to judges, prosecutors and defence counsel. Doing so may have the effect of curbing excessive differences. The Council must also be open to and actively seek feedback from the users of guidelines.

Sentencing guidelines are a heuristic device designed to overcome some of the challenges of sentencing. Guidelines provide a framework for organising sentencing decisions in a coherent way across cases, judges and circuits. By focussing on offence and offender based factors intended to be used to determine sentences and giving guidance about the range of sentences applicable for combinations of the central factors for each offence the balance between formal and substantive rationalities can be altered. The influence of formal rationality should increase at the expense of substantive rationalities based on local conditions and social politics. The extent to which justice is done in the sense of treating similar cases in a similar way will increase.

Future Research

The discussion of future research is broken into two sections. First, a discussion of the lessons learnt during the conduct of the research that may benefit others in the design of research is provided. Second, there is a discussion of the future directions and opportunities for sentencing research.

Lessons

Four broad lessons can be taken from this research. The first lesson involves the selection of offences to be addressed in research of this type. Based on the criteria employed male assaults female appeared to be a viable offence. However, insufficient account was given to the importance of the offence standing alone. Male assaults female is just one offence that is part of the wider array of physical assaults. In New Zealand the array of physical assaults is detailed with offences with maximum penalties of 6 months, 1 year, 2 years, 3 years, 5 years, 7 years, 10 years and 14 years. At the margins prosecuting agencies have to

decide which offence to charge, and these decisions may vary systematically from place to place. This variation may account for unexplained variation in sentencing. The obvious solution would be to treat physical assaults as a single offence grouping. This would have the incidental benefit of allowing the inclusion of the offence type as a broad proxy variable for the seriousness of offending.

A second lesson drawn from the modelling of male assaults female is that a 12% incarceration rate may not be sufficient to support effective modelling of the in-out or sentence length decisions. Any investigation of inconsistency may be better directed to a lower threshold, perhaps addressing the split between community sentences and monetary penalties. It is acknowledged that any such research would have to address factors such as the ability to pay which would require access to data that was not available for this research.

The third lesson was the validation of the decision to disaggregate the research to address sentencing according to offence type and decision type. The separation of decision types is common in sentencing research but the consideration of distinct offences is much less common. If the decision types had not been separated it would not have been possible to identify that gender and ethnicity were influential in the use of incarceration for burglary but not the determination of sentence length. The varying results found in this research demonstrate that the factors influencing sentence are not pervasive irrespective of offence. This is exemplified by the finding that gender and ethnicity were only significant for burglary but not aggravated drink driving or male assaults female. Wider groupings of offences should be approached with caution because effects may be averaged out and important distinctions could go unobserved.

The fourth lesson is particular to New Zealand. This research identified the risks of using sentence notes. The uneven availability of sentence notes across the circuits was a practical problem peculiar to this research and could be overcome in future research. However, the variable quality of the content with some sentence notes providing no information linking the offence and the offender with the sentence and others providing a very detailed explanation of the basis for the sentence cannot be easily overcome. In future, researchers may be better off seeking access to the statements of fact, pre-sentence reports and sentencing submissions that inform the determination of the sentence.

Future directions

Because there is no tradition of sentencing research using large scale regression models in New Zealand, there is a multitude of opportunities for future research stemming from this study. The following suggestions are those that are most closely tied to the current research. Researchers could examine whether:

- 1) there is any evidence of inconsistency in sentencing between judges within the same circuit, thus addressing whether personal sentencing constructs vary between judges in the same circuit or whether the convergence at circuit level eliminates any judge level variation;
- 2) there is any evidence of variation in sentencing between courts within the same circuit when offenders are sentenced by resident judges, addressing whether the location of the court in the circuit has any influence on sentencing within the circuit;
- 3) there is any evidence of any change in sentencing approach within a circuit when there are changes of judicial personnel, and in particular whether any changes vary with the size of the circuit;
- 4) guideline judgments influence sentencing consistency between circuits;
- 5) circuit level inconsistency of sentencing extends to the selection of sentences at lower levels in the sentencing hierarchy; or
- 6) circuit level inconsistency extends to the application of specific sentences, in particular the use of supervision sentences, the hours of community work, the size of monetary penalties or the use of conviction and discharge or discharge without conviction.

There are many other possible avenues for research. Research designed to extend understanding of the effects of gender and ethnicity could be undertaken. The research could take in a wider range of offence types and seek to identify whether there are any patterns in the influence of these factors. The research could be extended to take account of interaction effects between demographic variables and between demographic variables and other sentencing variables.

With the agreement of the Judiciary research could be undertaken into the relationship between judges' backgrounds and sentencing outcomes to determine whether any aspect(s) of the judges' backgrounds are influential in the formation of personal sentencing constructs and whether similarities are evident between the judges resident in the same circuit.

There are no real opportunities to overcome the limited ability to investigate contextual differences between circuits. Research using the individual courts as an intermediate level of analysis is not feasible because the judges are not nested within courts, just within circuits. With the agreement of the Judiciary there may be greater opportunity to analyse context at the judge level but this may also be limited by the numbers of offenders sentenced per judge and would entail a need to cover a range of offences in the same model.

The understanding of sentencing disparity could be advanced by undertaking qualitative research to add judicial, prosecutorial or defence counsel perspectives on sentencing and their responses to the findings of this quantitative research. Interviews with these key participants may provide greater understanding of the forces behind differences in sentencing between metropolitan and provincial circuits.

Conclusion

This thesis reports the first large scale systematic study of sentencing consistency in New Zealand. The results are clear. There were statistically significant differences between circuits in the approach to sentencing for all three of the high volume offences sentenced in the District Courts that were the subject of regression models. The sentences imposed on similar offenders who offended in similar situations differed substantially depending on where in New Zealand an offender was sentenced for aggravated drink driving, male assaults female or burglary. The differences were not merely statistically significant, they were also substantively significant. The reliability of this finding is strengthened because of the detailed list of sentencing variables able to be included in the models. Those who rejected the evidence of inconsistency claiming it was the consequence of compositional differences in caseload were wrong.

Consistent with prior research and theory, the decision to incarcerate and the length of incarceration were strongly influenced by legally relevant variables relating to offence seriousness and the criminal history of offenders. Judges do not need guidance on what to take into account, as the models suggest there is already substantial agreement on that. The difficulty lies in the application of the factors, the determination of what level of offending warrants a step up the hierarchy of sentences and the determination of the quantum of sentence.

The mixed results in relation to the impact of gender and ethnicity suggests, consistent with theory, that when judges are uncertain of how to proceed they may be swayed by stereotypical views based on gender and ethnicity.

The findings of this research have significant implications for the judiciary and for sentencing policy makers. The current sentencing framework seems to have worked effectively to inform judges about what factors to take into account. This is evidenced by the significant role offence seriousness and criminal history factors play in sentencing. However, the statutory framework and the results of appeals have not been sufficient to bring about common sentencing approaches for these three offences. This situation has the potential to bring sentencing and justice into disrepute if it is not addressed and corrected.

The distinctive nature of offences, the differing content of the models presented in this thesis, and the history of guideline judgments all point to the need for some form of offence-based guidance. The guidance needs to extend beyond the narrow range of offences currently covered by guideline judgments. The implementation of the Sentencing Council and the development of sentencing guidelines would enable guidelines to be introduced relatively quickly and enable timely periodic rapid revisions.

Without change the current sentencing regime with wide discretion in the pursuit of individualised justice will continue to result in widespread injustice as similar offenders being sentenced for similar offending receive different sentences dependent on where they are sentenced. As the analysis of burglary sentencing shows, the inconsistency can also manifest itself as differing treatment of offenders depending on gender or ethnicity.

None of these findings and recommendations should be taken as a criticism of the judiciary. Judges are not deliberately acting arbitrarily or outside the rules. Inconsistency occurs despite the best efforts of judges to do justice across all the offenders who appear before them. The inconsistency is the outcome of a system that in seeking to ensure individualised justice provides too little guidance. The assumption that wide discretion is needed to deal with the infinite variety of cases is sorely tested. The focus on the notion that every case is unique overlooks the prospect that for many offence types and offenders the similarities outweigh the differences. Failure to recognise this, and the failure to provide guidance, has left judges having to fill in the gap for themselves. The consistency that Hall (1991) described as inherent in the rule of law is not achieved, and the rule of law has broken down, arguably creating the manifest injustice that Ashworth (1998) thought capable of bringing a sentencing system into disrepute. The sentence an offender gets should not be influenced by their gender, ethnicity or where they are sentenced.

APPENDICES

Appendix A

Sentencing Variability Associated with Socio-demographic Factors

Much of the early research work investigating sentencing disparity focussed on demographic factors, and in particular on the race/ethnicity of the offender. A substantial amount has also addressed gender/sex and age has also featured. These variables have appeared have control factors in much of the research on geographic disparity, the importance of these controls is demonstrated by the incidence of statistically significant influence found in the models. There has only been limited research exploring socio-economic status, information on income, employment and educational attainment is not readily available. This review uses the terms employed by the authors when referring to specific studies and otherwise uses the terms ethnicity and gender as these are the terms employed in official conviction and sentencing statistics in New Zealand.

The direct influence of ethnicity on sentencing

Driven, in part, by the overrepresentation of certain ethnic and racial groups in the prison population, many studies have examined whether the ethnicity or race of an offender leads to different sentencing outcomes for similar offending. Johnson (2011) concluded that contemporary research in the United States indicates that black and Hispanic offenders are often disadvantaged at sentencing but the disadvantage is not systemic and seems to be limited to certain contexts. A substantial narrative review (Spohn, 2000) and a meta-analysis (Mitchell & Mackenzie, 2004) illustrate the effects described by Johnson (2011).

Spohn (2000) conducted a narrative review of the effects of race (black vs. white offenders) and ethnicity (Hispanic vs. white offenders) on sentencing following sentencing reforms seeking to establish a racially unbiased system. Spohn (2000) reviewed 40 studies published between 1983 and 2000. Prior reviews had found that race only exerted a modest effect on sentencing once controls for legally relevant factors were employed (Spohn, 2000). However, these reviews tended to be based on studies from the 1970s and earlier and thus reflected a different sentencing system and did not measure contemporary sentencing outcomes reflecting reforms such as the introduction of guidelines, determinate sentencing and mandatory minimums.

Spohn (2000) determined that a group was treated more severely or leniently if the parameter estimate for the group was statistically significant at $p < .05$ level. When this threshold was applied to sentencing in state courts Spohn (2000) found blacks were sentenced more severely than whites 41 times, more leniently 4 times and there was no

significant difference in 50 of 95 contrasts. Hispanic offenders were sentenced more severely than whites 8 times, more leniently 2 times and there was no significant difference in 19 of 29 contrasts. At Federal level blacks were sentenced more severely in 15 of 22 contrasts and Hispanics more severely in 10 of 21 contrasts. Neither blacks nor Hispanics were sentenced more leniently in any of the Federal level studies.

The findings varied between contrasts of the in-out decision, sentence length and those based on other measures of severity. At state level 30 of the 57 contrasts of the in-out decision resulted in blacks or Hispanics being sentenced more severely, whereas only 10 of 53 contrasts resulted in longer sentences. At Federal level the results were more balanced with 6 of 14 in-out decision contrasts and 8 of 17 sentence length contrasts resulting in more severe sentences for blacks or Hispanics. In 10 of 14 contrasts at state level and 11 of 12 contrasts at Federal level blacks or Hispanics were less likely to receive downward departures, more likely to be treated as habitual offenders, less likely to avoid a mandatory minimum sentence, and less likely to receive a reduction for acceptance of responsibility.

Based on the totality of these results Spohn (2000) concluded that black and Hispanic offenders sentenced at state level and in the Federal system face higher odds of incarceration and in some jurisdictions longer sentences and lower chances of downward departure.

Mitchell and Mackenzie (2004) took the assessment of the wider body of research a step further, undertaking a meta-analysis of sentencing studies at state and Federal level. The analysis covered 122 distinct sentencing contexts drawn from 71 studies, yielding 430 effect sizes or contrasts that could be separately coded and standardised to enable results for dichotomous dependent variables (the in-out decision or departure decisions) and interval level or continuous dependent variables (sentence length, the amount of a reduction for departure) to be combined. Two-thirds of the contrasts compared African Americans and whites and 25% compared Latinos and whites.

The results across the studies were mixed, with a high proportion finding that minorities were sentenced more severely. African Americans were sentenced more severely than whites in 83% of the contrasts across the Federal system and 77% of contrasts across state systems and Latinos were sentenced more severely than whites in 72% of the contrasts across both Federal and state systems. Using multivariate analysis Mitchell and Mackenzie (2004) found that the random effects mean odds ratio for African-Americans was 1.28 ($p < .01$) across the studies of state systems. Across the Federal

system studies the odds ratio was 1.15 but it was not statistically significant at conventional levels ($p = .093$). For Latinos it was 1.18 ($p < .01$) across Federal and state systems combined. Whites were the reference category for these contrasts.

Mitchell and Mackenzie (2004) described the results as demonstrating that minorities were sentenced more severely but also described the effect as small and highly variable. To provide an indication of the relative size of the effect they said that if the incarceration rate for white offenders was 50% the rate for African-Americans would be 56% in the state courts and 53% in the Federal courts. The rate across both systems combined would be 54% for Latinos compared to the hypothetical 50% for whites.

The importance of context is demonstrated by results disaggregated by offence and time period by Mitchell and Mackenzie (2004) and Mitchell (2005). Mitchell and Mackenzie (2004) found that when the odds ratios for Federal studies were calculated by offence type the odds ratio for property offences was 1.32 ($p > .01$) but for drugs it was only 1.08 and not significant. When Mitchell (2005) re-analysed the data taking account of the introduction of the Federal guidelines, limiting the analysis to studies based on data series with a mid-point in or after 1987 (the year the guidelines were introduced) the odds ratio for drugs increased to 1.44 and was higher than for property offences 1.35.

The reviews of the literature generally conclude that racial disparities in sentencing exist (e.g., Mitchell & Mackenzie, 2004; Mitchell, 2005; Spohn, 2000), but that the overall magnitude of racial effects tends to be relatively small. The variation in studies reported by both Spohn (2000) and Mitchell and Mackenzie (2004) confirms Johnson's (2011) general conclusion that the sentencing disadvantage is not systemic but seems to be limited to certain contexts.

Outside of the United States there has only been a limited amount of research. Research on ethnic differences in sentencing outcomes has generated mixed results in Australia. Two studies of the in-out decision in Western Australia (Bond & Jeffries, 2011), and New South Wales (Snowball & Weatherburn, 2006) found no statistically significant differences after controlling for legal factors. In South Australia Jeffries and Bond (2009) found that indigenous offenders were less likely to be incarcerated, reporting an odds ratio of 0.49 ($p < .05$), but were likely to receive sentences 1.21 times longer ($p < .05$) than non-indigenous offenders after controlling for legal factors. The small size and weak controls for offence seriousness suggests these findings should be treated with caution. A second New South Wales study (Snowball & Weatherburn, 2007) reported that a model controlling for indigenous status and sex showed that the odds of incarceration for

indigenous offenders were 2.76 greater than non-indigenous but when controls for offence seriousness and criminal history were added the odds were 1.17 ($p < .05$) times greater. A study of sentence length on New South Wales offenders sentenced to imprisonment in 2003 and 2004 found that indigenous offenders received slightly shorter sentences in the lower courts (incident risk ratio = 0.963, $p < .05$) but no significant difference in the higher courts (Bond et al., 2011).

Two Canadian studies on the influence of indigeneity on sentencing across a variety of contexts found little evidence of a direct effect (Weinrath, 2007; Welsh & Ogloff, 2008). In a study of the sentencing of young offenders Latimer and Foss (2005) found no effect in the likelihood of incarceration but did find that indigenous young offenders were likely to receive sentences 1.5 times longer than non-indigenous ($p < .01$). However, this result was challenged by Doob and Sprott (2007) who found that the model was mis-specified because it failed to take account of regional differences, sentence lengths were similar in different places but the proportion of indigenous offenders varied from place to place.

Only one New Zealand study (Triggs, 1999) has addressed the direct influence of ethnicity on sentencing. Ethnicity was one of 14 factors in a logistic regression model addressing the probability of imprisonment for offenders sentenced in 1995. Neither Māori nor Pacific offenders were more likely to be imprisoned than Pakeha (European), though both Māori and Pacific offenders were significantly ($p < .001$) less likely to be fined and more likely to receive other non-custodial sentencing options like periodic detention and community service (Triggs, 1999). Triggs (1999) recognised that these results may be due to differences in the ability to pay fines, which was a factor that could not be controlled for in her study.

The Canadian, Australian and New Zealand research on indigenous offenders differs from the United States research on race. In a review of research on the impact of indigenous status on sentencing outcomes in the United States, Canada, and Australia, Jeffries and Bond (2012) found that the empirical evidence was mixed with some studies finding that indigenous offenders were treated more leniently and others that they were treated more severely and suggested that the differences may be reflective of differences in the political and social contexts for indigenous people in different jurisdictions. Jeffries and Bond (2012) also cautioned that the findings in some studies might be modified if there was better measurement of key variables describing criminal history, aggravating and mitigating factors, and the offenders' social background.

Sentencing variability associated with gender or sex

Research examining or controlling for the influence of gender or sex on sentencing outcomes has consistently found that women tend to receive less severe sanctions than men for like offences (e.g., Blackwell, Holleran & Finn, 2008; Jeffries et al., 2003; Steffensmeier & Demuth, 2001), leading Farrell, Ward and Rousseau (2010, pp. 85-86) to suggest that “leniency toward women has become an almost accepted phenomenon among scholars studying criminal case processing.” Zatz (2000) in a review of the literature found that the difference between the sexes was greatest for the in-out decision with lesser effects on sentence length.

Twenty-four of the studies from the United States included in this review controlled for gender (Albonetti, 1997; Albonetti & Baller, 2010; Britt, 2000; Crow & Gertz, 2008; Everett & Wojtkiewicz, 2002; Farrell et al, 2010; Fearn, 2005; Johnson, 2005 & 2006; Kautt, 2002; Kautt & Delone, 2006; Ostrom et al., 2004; Pasko, 2002; Sharp, 2007; Spohn, 2000; Steffensmeier et al., 1998; Ulmer, 2005; Ulmer & Johnson, 2004; Ulmer & Kramer, 1996; Ulmer & Light, 2010; Weidner & Frase, 2003; Weidner, Frase & Schultz, 2005; Wheeler, Weisburd & Bode 1982).

Females were significantly less likely to be incarcerated in 29 out of 33 contrasts in the 20 studies that addressed the in-out decision. None of the studies found females were sentenced more severely. Females were sentenced to significantly shorter terms in 22 out of 26 contrasts in the 21 studies that addressed sentence length. None of the studies found females were sentenced more severely in terms of the likelihood of incarceration or sentence length. Females were found to be more likely to receive downward departures in all three of the studies that addressed departures from sentencing guidelines (Johnson, 2005; Johnson, Ulmer & Kramer, 2008; and Sharp, 2007). Johnson (2005) also found females were less likely to receive upward departures. Johnson et al. (2008) found that the magnitude of downward departures was larger for females.

The Australian studies that focussed on the impact of indigeneity on sentencing also controlled for gender. Three of the four studies that included the in-out decision found that men were significantly more likely to be imprisoned in Western Australia (Bond & Jeffries, 2011) and in New South Wales (Snowball & Weatherburn, 2006, 2007). Jeffries and Bond (2009) found no significant effect in South Australia. Both studies that addressed sentence length found that women received shorter sentences after controlling for seriousness and criminal history. In the South Australian study sentences for women were 65% of the length of male sentences (IRR = 0.647, $p < .01$) (Jeffries & Bond, 2009).

In New South Wales, Bond et al. (2011) found that sentences for women were 81% of sentences for men in the lower courts and 78% in the higher courts (IRR = .806, $p < .001$ [lower courts] and IRR = .781, $p < .001$ [higher courts]). A Canadian study (Latimer and Foss, 2005) of the sentencing of young offenders found males were more likely to be incarcerated than females, and if incarcerated, to be incarcerated for twice as long ($p < .01$).

Two New Zealand studies have addressed the influence of sex on sentencing. Triggs (1999) found that females were much less likely to be imprisoned, with an odds ratio of 0.518 ($p < .001$). Jeffries et al. (2003) reported the results of a study of sentencing for a matched sample of 388 male and female offenders for more serious offences in Christchurch. The variables were entered in a series of blocks and the results for the partial models were reported. The full model did not achieve statistical significance in relation to the in-out decision. However, the sixth and last block of variables in the model included the pre-sentence report recommendation and that variable dominated all others in the model. Prior to inclusion of the final block of variables and the recommendation, the model indicated that men were significantly and substantially more likely to be imprisoned. When the model included demographics, criminal history, offence variables, court processing and remand status but excluded the recommended sentence, the odds ratio was 1.94 ($p < .05$). The full model did show that women received statistically significantly shorter terms of imprisonment (21 months compared to 28 months for males, $p < .05$).

Unlike race, the findings about the influence of the offender's sex on sentencing are common between the United States research and research elsewhere. With limited exceptions the research finds that, after controlling for legal factors, women are less likely to be incarcerated and when incarcerated their sentences are shorter.

Sentencing variability associated with age

Researchers often control for age. However, they have not been particularly interested in the direct effects of age and have instead taken greater interest in how it conditions other social variables, especially race (Ulmer, 2011). The direct effects of age were reported in 20 of the studies that focussed on or controlled for location. (Britt, 2000; Everett and Wojtkiewicz, 2002; Farrell et al., 2010; Fearn, 2005; Iles, Bumphus & Zehel, 2011; Johnson, 2006; Johnson et al., 2008; Kautt & Delone, 2006; Ostrom et al., 2004; Pasko, 2002; Steffensmeier et al., 1998; Ulmer, 2005; Ulmer & Johnson, 2004; Ulmer & Kramer, 1996; Ulmer & Light, 2010; Weidner et al., 2005; Wheeler et al., 1982). Most of the

studies treated age as a continuous variable and found that the likelihood of incarceration reduced with age. One study (Wheeler et al., 1982) departed from this pattern, finding that the likelihood of imprisonment increased with age. This may be a consequence of that study focussing on white collar crime. Three studies (Iles, Bumphus & Zehel, 2011; Pasko, 2002; Weidner et al., 2005) found no significant effects and 17 studies found at least some effects for one or more decisions (in-out, sentence length or departure).

However, these results must be treated with some caution due to the findings from four studies that treated age as a categorical variable. Everett and Wojtkiewicz (2002), Fearn (2005), Ostrom et al. (2004) and Steffensmeier et al. (1998) found that middle-aged offenders were more likely to be incarcerated than younger and older offenders. This suggests that age breaches the assumption of linearity and should be treated as a categorical variable or otherwise adjusted (e.g. Iles et al. (2011) included the square of the offender's age as an extra variable).

Triggs (1999) included age in logistic regression modelling of the probability of imprisonment for offenders sentenced in New Zealand in 1995. Compared to 20 - 29 year olds, very young offenders (< 17) were very unlikely to be imprisoned (OR 0.158, $p < .001$), whereas offenders aged 17 – 19 were not significantly different and offenders aged 30 or more (OR 1.090, $p < .05$) were a little more likely to be imprisoned. The finding that very young offenders were unlikely to be imprisoned was consistent with sentencing legislation, the Criminal Justice Act 1985 and the Children, Young Persons, and Their Families Act 1989 favoured the use of non-custodial sentences for young offenders.

The direct effects of age are difficult to interpret, and are complicated if it is correct that those studies treating age as a continuous variable may be mis-specified because the assumption of linearity is not met.

The interaction effects of gender, ethnicity and age

More recently researchers have focused on the influence of interactions between gender, ethnicity and age on disparities in sentencing by including interaction variables in their models (e.g., Doerner & Demuth, 2010; Steffensmeier & Demuth, 2006; Steffensmeier et al., 1998). Others have addressed interactions less directly by specifying separate models and comparing and contrasting the results (e.g. Albonetti, 1997; Kautt & Spohn, 2007). Most of the research focuses on how ethnicity interacts with gender and age to produce disparities in sentencing, with a number of studies finding that young black and Hispanic males tend to be subject to more severe sentences (e.g., Doerner & Demuth, 2010;

Steffensmeier & Demuth, 2006). However, main effects for ethnicity, gender and age are also usually reported (e.g., Steffensmeier et al., 1998), with younger offenders more likely to be incarcerated (Ulmer & Bradley, 2006).

Steffensmeier et al. (1998) demonstrated that the consideration of the interaction effects provides a richer and more subtle depiction of disparities, finding that young black men received the most severe sentences in terms of both type and length. The sentences they received were more severe than for younger white males. Younger males of both races were sentenced more severely than older males and there was no significant difference between older black and white males. Conversely there were only marginal effects for age for females but at all ages black females were sentenced more severely than white females.

The findings of interaction effects for offender race, sex and age are consistent with the direct effects for each of these factors and demonstrate a compounding effect so that young black males are treated more severely than blacks in general, males in general or offenders of any age.

Sentencing variability associated with socio-economic status

Research on potentially important social factors such as unemployment and social class remains limited and few definitive conclusions can be drawn at this stage (Johnson, 2011). Comparatively few researchers have attempted to address any difference between socio-economic groupings and, when they have, it has most often been operationalised using educational achievement or employment as a proxy measure (Zatz, 2000). The major challenge for researchers is the lack of information that is routinely captured. Information on socio-economic status in any form is generally contained in documentary sources like pre-sentence reports and sentencing submissions and requires substantial effort to be assembled. This contrasts with information on race, sex and age which are regularly captured in, or readily calculable from, administrative databases.

The direct effects of educational attainment were reported in nine of the studies that focussed on or controlled for location (Albonetti, 1997; Albonetti & Baller, 2010; Everett & Wotjkieicz, 2002; Iles et al., 2011; Johnson et al., 2008; Kautt & Delone, 2006; Pasko, 2002; Sharp, 2007; Ulmer & Light, 2010). Employment was addressed in a single study (Spohn, 2000) and occupational status and social contribution were both addressed by Wheeler et al. (1982). Higher educational attainment was associated with lower sentence severity in 20 out of 22 contrasts including 4/4 in-out, 8/10 sentence length (with two

contrasts not significant) and 7/7 departure decisions. The remaining contrast combined sentence type and length into a single sentence seriousness scale.

Spohn (2000) found no significant effects for employment status on sentence length or on the in-out decision. Wheeler et al. (1982) found significant effects for both occupational status and social standing on the sentencing of white collar offences. The higher an offender's occupational status the more likely he or she was to be incarcerated and the sentence was likely to be marginally longer. Conversely, the more significant the offender's social contribution (voluntary works, past behaviour) the less likely he or she was to be incarcerated and the shorter the sentence.

Two of the Australian studies (Bond & Jeffries, 2011; Jeffries & Bond, 2009) included employment as a control. Neither study found that employment status had a significant influence on sentence. In New Zealand, Jeffries et al. (2003) included employment as a variable in the examination of sentencing by sex in Christchurch. Offenders not in paid employment were significantly more likely to be imprisoned but there was no effect on sentence length.

Appendix B

Comparison of Parameter Estimates for Sentencing Variables in Single Level Regressions Models and Multi-level Random Intercept Models

The tables in this appendix compare the parameter estimates for single level logistic and linear regression models for each of the six offence and decision scenarios modelled in this research with multi-level hierarchical generalised linear and hierarchical linear random intercept models to determine if there is any evidence that the parameter estimates in the single level models are under-estimated.

Examination of the tables indicates no evidence of systematic under-estimation of standard errors or over-estimation of statistical significance due to nesting of the data. The differences between the two types of model are generally small and the relative magnitude of the standard errors is sometimes lower in the hierarchical models when the standard error is divided by the associated co-efficient. The only large proportionate differences occurred in relation to variables that were not significant in either model.

There was one instance of a variable that was statistically significant in a single level model that became non-significant in the hierarchical model. The change occurred in relation to the likelihood of a Pacific offender being incarcerated for burglary relative to a European offender with the *p*-value increasing .049 to .070.

Table B1: Comparison of parameter estimates for logistic regression and hierarchical generalised linear models for aggravated drink driving ($n = 9,282$)

Variable (reference categories in brackets)	Single level model			Multi-level model		
	B	S.E.	Sig.	B	S.E.	Sig.
Constant	-8.666	.285	.000	-5.762	.382	<.001
Alcohol level	.529	.064	.000	.526	.063	<.001
Multiple drink drive (no)	1.932	.152	.000	1.916	.152	<.001
Disqualified driving (no)	.347	.108	.001	.344	.108	.001
Serious driving charge (no)	.380	.143	.008	.378	.143	.008
Breach of sentence (no)	.746	.155	.000	.744	.154	<.001
Total convictions for sentence						
Two convictions (1)	.239	.097	.014	.239	.096	.013
Three or more convictions (1)	.742	.122	.000	.739	.122	<.001
Elapsed time - charge to charge outcome						
> 6 weeks ≤ 3 months (≤ 6 weeks)	1.026	.119	.000	1.022	.119	<.001
> 3 months ≤ 6 months (≤ 6 weeks)	1.168	.120	.000	1.161	.120	<.001
> 6 months (≤ 6 weeks)	1.012	.127	.000	.998	.127	<.001
Plea (guilty)	-.512	.230	.026	-.510	.229	.026
Prior convictions for drink drive						
3 priors (2)	1.118	.106	.000	1.114	.106	<.001
4 priors (2)	1.873	.125	.000	1.864	.125	<.001
5 or 6 priors (2)	2.413	.138	.000	2.400	.137	<.001
7 or more priors (2)	3.136	.170	.000	3.114	.169	<.001
Drink drive convictions in the past 10 years						
1 conviction (0)	1.213	.153	.000	1.211	.153	<.001
2 convictions (0)	1.757	.159	.000	1.753	.159	<.001
3 or more convictions (0)	1.967	.174	.000	1.958	.174	<.001
Elapsed time since the last drink drive conviction						
> 18 months ≤ 5 years (over 5 years)	.714	.086	.000	.710	.086	<.001
≤ 18 months (over 5 years)	1.052	.120	.000	1.045	.119	<.001
Prior incarceration			.000			
Yes, not for drink drive (no)	.326	.097	.001	.327	.097	<.001
Yes, including for drink drive (no)	.577	.092	.000	.582	.091	<.001
Conviction for breach of sentence in last 5 years (no)	.503	.088	.000	.500	.087	<.001
Female (male)	-.091	.100	.362	-.090	.100	.368
Race			.			
Māori (European)	.067	.076	.375	.066	.075	.378
Pacific (European)	.078	.137	.571	.062	.137	.65
Other or unknown (European)	-.013	.172	.940	-.021	.172	.901
Age at time of offending						
Aged 32 - 39 (17 - 31)	-.225	.103	.028	-.225	.102	.028
Aged 40 - 47 (17 - 31)	-.342	.114	.003	-.341	.114	.003
Aged 48 or older (17 - 31)	-.493	.124	.000	-.491	.124	<.001

Table B2: Comparison of parameter estimates for linear regression and hierarchical linear models for aggravated drink driving ($n = 2,312$)

Variable (reference categories in brackets)	Single level model			Multi-level model		
	B	S.E.	Sig.	B	S.E.	Sig.
(Constant)	1.769	.648	.006	9.061	.295	<.001
Prior convictions for drink driving	.473	.047	.000	.466	.047	<.001
Years since the last drink drive conviction	-.131	.032	.000	-.129	.032	<.001
Longest term of imprisonment for drink driving	.284	.024	.000	.287	.024	<.001
Alcohol levels standardised	1.135	.159	.000	1.130	.159	<.001
Drink drive charges in the case	2.704	.292	.000	2.706	.291	<.001
Case includes disqualified driving charge(s)	.503	.207	.015	.511	.207	.014
Case includes breach(es) of bail (no)	1.498	.356	.000	-1.494	.355	<.001
Case includes breach(es) of sentence (no)	-.692	.303	.023	-.686	.303	.024
Total number of charges for sentence	.501	.057	.000	.494	.057	<.001
The number of days from charge to charge outcome	.003	.001	.013	.003	.001	.012
Sex (male)	-.176	.262	.501	-.184	.262	.481
Age when offended	-.010	.010	.323	-.009	.010	.351
Māori (cf European)	-.358	.191	.061	-.352	.190	.065
Pacific (cf European)	.022	.381	.953	-.025	.378	.948
Other (cf European)	-.468	.488	.338	-.486	.488	.319

Table B3: Comparison of parameter estimates for logistic regression and hierarchical generalised linear models for male assaults female ($n = 6,166$)

Variable (reference categories in brackets)	Single level model			Multi-level model		
	B	S.E.	Sig.	B	S.E.	Sig.
Constant	-5.679	.371	.000	-5.265	.349	<.001
Multiple male assaults female (no)	.530	.152	.000	.527	.151	<.001
Threat to kill or do GBH (no)	1.882	.618	.002	1.958	.615	.001
Breach of sentence (no)	.579	.174	.001	.571	.172	<.001
Total convictions for sentence						
Two convictions (1)	.783	.139	.000	.782	.138	<.001
Three convictions (1)	1.031	.172	.000	1.020	.171	<.001
Four or more convictions (1)	2.064	.163	.000	2.044	.162	<.001
Plea (guilty)	.758	.184	.000	.740	.183	<.001
Elapsed time - charge to charge outcome						
> 6 weeks $\leq$ 3 months ($\leq$ 6 weeks)	.495	.151	.001	.515	.149	<.001
> 3 months $\leq$ 6 months ($\leq$ 6 weeks)	.346	.153	.024	.340	.151	.024
> 6 months ($\leq$ 6 weeks)	-.004	.177	.981	-.003	.174	.988
Charge laid indictably (no)	1.870	.306	.000	1.872	.303	<.001
Prior convictions for male assaults female						
1 prior (0)	.118	.175	.501	.105	.174	.545
2 priors (0)	.642	.211	.002	.641	.210	.002
3 or more priors (0)	1.221	.217	.000	1.196	.216	<.001
Conviction for male assaults female in last 5 years (no)	.609	.165	.000	.603	.164	<.001
Conviction for serious violence in last 5 years (no)	.908	.133	.000	.897	.132	<.001
Conviction for breach of protection order in last 5 years (no)	.515	.151	.001	.505	.150	<.001
Conviction for breach of sentence in last 5 years (no)	.723	.117	.000	.705	.117	<.001
Prior convictions for any offence						
1-9 priors (0)	.640	.315	.042	.656	.314	.037
10-19 priors (0)	1.308	.331	.000	1.315	.331	<.001
20 or more priors (0)	1.229	.355	.001	1.233	.354	<.001
Prior incarceration for any offence						
Once (0)	.716	.166	.000	.729	.165	<.001
2-4 times (0)	.914	.185	.000	.933	.184	<.001
5 or more times (0)	1.694	.211	.000	1.707	.210	<.001
Māori (European)	.142	.132	.283	.107	.129	.407
Pacific (European)	.350	.198	.077	.272	.193	.158
Other or unknown (European)	.277	.304	.362	.246	.301	.414
Age at time of offending						
Aged 24-30 (17-23)	-.422	.153	.006	-.434	.152	.004
Aged 31-38 (17-23)	-.371	.163	.023	-.389	.162	.017
Aged 39 or older (17-23)	-.892	.185	.000	-.904	.184	<.001

Table B4: Comparison of parameter estimates for linear regression and hierarchical linear models for male assaults female ($n = 722$)

Variable (reference categories in brackets)	Single level model			Multi-level model		
	B	S.E.	Sig.	B	S.E.	Sig.
(Constant)	3.162	.894	.000	6.156	.466	<.001
Multiple male assaults female charges	1.989	.412	.000	2.137	.418	<.001
Case includes breach(es) of bail	-1.172	.526	.026	-1.075	.529	.043
Total number of charges for sentence	.478	.067	.000	.478	.067	<.001
Charge laid indictably (cf summarily)	2.387	.733	.001	2.994	.734	<.001
Prior convictions for male assaults female	.269	.120	.025	.278	.121	.021
Prior incarceration for male assaults female	1.387	.424	.001	1.477	.430	<.001
Recent serious violence conviction (cf none or > 5 years ago)	1.122	.344	.001	1.176	.349	<.001
Age when offended	-.016	.020	.408	-.022	.020	.262
Māori (cf European)	.399	.413	.334	.060	.408	.883
Pacific (cf European)	.109	.308	.724	-.019	.306	.951
Other (cf European)	-.311	.357	.384	-.363	.361	.316

Table B5: Comparison of parameter estimates for logistic regression and hierarchical generalised linear models for burglary ($n = 5,740$)

Variable (reference categories in brackets)	Single level model			Multi-level model		
	B	S.E.	Sig.	B	S.E.	Sig.
Constant	-5.567	0.253	.000	-4.425	.261	<0.001
2 burglary convictions for sentence (one)	.461	.117	.000	.457	.116	<0.001
3 burglary convictions for sentence (one)	1.087	.140	.000	1.081	.139	<0.001
Other serious conviction (no)	.463	.102	.000	.460	.101	<0.001
Breach of sentence conviction (no)	.528	.115	.000	.524	.115	<0.001
Value of lead burglary \$500 to \$5,000 (< \$500)	.449	.089	.000	.439	.088	<0.001
Value of lead burglary \$5,000 + (< \$500)	1.022	.137	.000	1.006	.136	<0.001
Value of lead burglary unknown (< \$500)	.367	.155	.018	.364	.154	0.018
Two convictions for sentence (one)	.422	.127	.001	.416	.127	0.001
Three or four convictions for sentence (one)	.900	.126	.000	.893	.126	<0.001
Five to seven convictions for sentence (one)	1.409	.151	.000	1.397	.150	<0.001
Eight or more convictions for sentence (one)	2.413	.171	.000	2.378	.169	<0.001
Reparation ordered (no)	-.513	.081	.000	-.502	.081	<0.001
Charge laid indictably (no)	1.638	.160	.000	1.633	.159	<0.001
Elapsed time - charge to charge outcome						
> 6 weeks ≤ 3 months (≤ 6 weeks)	.351	.116	.003	.346	.116	0.003
> 3 months ≤ 6 months (≤ 6 weeks)	.212	.117	.071	.202	.117	0.083
> 6 months (≤ 6 weeks)	.161	.131	.217	.138	.130	0.288
Prior convictions for serious dishonesty including burglary						
One or two (none)	.296	.133	.026	.294	.132	0.026
Three to nine (none)	.447	.156	.004	.446	.155	0.004
Ten or more (none)	.613	.206	.003	.606	.205	0.003
Conviction for burglary last 5 years (no)	.838	.107	.000	.826	.106	<0.001
Prior conviction for serious violence (no)	.595	.103	.000	.590	.103	<0.001
Conviction for breach sentence last 5 years (no)	.434	.094	.000	.434	.094	<0.001
1 – 4 prior convictions for any offence (none)	.595	.172	.001	.594	.171	<0.001
5 - 9 prior convictions for any offence (none)	1.169	.191	.000	1.164	.190	<0.001
10 - 19 prior convictions for any offence (none)	1.354	.212	.000	1.348	.211	<0.001
20 - 49 prior convictions for any offence (none)	1.671	.241	.000	1.654	.240	<0.001
50 + prior convictions for any offence (none)	1.835	.291	.000	1.809	.290	<0.001
Longest prior incarceration for burglary						
≤12 months (none)	.529	.138	.000	.538	.137	<0.001
>12 months ≤24 months (none)	.732	.177	.000	.733	.176	<0.001
> 24 months (none)	1.509	.251	.000	1.512	.250	<0.001
Female (male)	-.529	.144	.000	-.526	.144	<0.001
Māori (European)	.274	.089	.002	.275	.089	0.002
Pacific (European)	.323	.165	.049	.297	.163	0.07
Other or unknown (European)	.075	.234	.749	.062	.233	0.789
Aged 20-29 (13-19)	.033	.102	.744	.030	.102	0.769
Aged 30-39 (13-19)	.030	.147	.839	.012	.073	0.871
Aged 40 or older (13-19)	-.282	.188	.135	-.093	.062	0.135

Table B6: Comparison of parameter estimates for linear regression and hierarchical linear models for burglary ($n = 2,338$)

Variable (reference categories in brackets)	Single level model			Multi-level model		
	B	S.E.	Sig.	B	S.E.	Sig.
(Constant)	-1.687	1.087	.121	8.881	.895	<.001
Number of burglary charges for sentence	.713	.074	.000	.719	.074	<.001
\$500 to \$5,000 cf Less than \$500 or unknown	1.236	.366	.001	1.195	.365	.001
More than \$5,000 cf Less than \$500 or unknown	3.541	.502	.000	3.501	.502	<.001
Count of offence dates for the burglaries	2.314	.259	.000	2.291	.259	<.001
Case includes receiving	2.440	.490	.000	2.426	.490	<.001
Case includes breach of bail	-1.633	.449	.000	-1.623	.449	<.001
Case includes breach of sentence	.888	.388	.022	.892	.388	.022
Case includes violent offence	1.887	.416	.000	1.892	.416	<.001
Case includes serious driving offence	1.136	.463	.014	1.123	.463	.015
Total number of charges for sentence	.328	.039	.000	.321	.039	<.001
Charge laid indictably or summarily	5.738	.463	.000	5.773	.463	<.001
Plea	1.951	.822	.018	1.989	.822	.016
Days from charge to charge outcome	.005	.002	.005	.005	.002	.009
Whether a reparation order was made	-.831	.370	.025	-.781	.369	.035
Categorisation of judges by warrant	-1.391	.594	.019	-1.323	.593	.026
One cf no prior convictions for burglary	.497	.546	.364	.515	.546	.346
Two cf no prior convictions for burglary	.837	.640	.191	.870	.640	.175
3, 4 or 5 cf no prior convictions for burglary	2.569	.576	.000	2.571	.576	<.001
6 to 9 cf no prior convictions for burglary	3.571	.728	.000	3.569	.728	<.001
10 to 19 cf no prior convictions for burglary	5.355	.849	.000	5.327	.848	<.001
20 or more cf no prior convictions for burglary	6.374	1.062	.000	6.337	1.062	<.001
Longest prior term of incarceration for burglary	.128	.022	.000	.129	.022	<.001
One cf no prior convictions for serious dishonesty excl burglary	.399	.601	.507	.346	.600	.565
Two cf no prior convictions for serious dishonesty excl burglary	1.036	.646	.109	1.026	.646	.112
3, 4 or 5 cf no prior convictions for serious dishonesty excl burglary	2.693	.578	.000	2.656	.578	<.001
6 to 9 cf no prior convictions for serious dishonesty excl burglary	2.672	.676	.000	2.641	.676	<.001
10 to 19 cf no prior convictions for serious dishonesty excl burglary	2.240	.665	.001	2.192	.665	<.001
20 or more cf no prior convictions for serious dishonesty excl burglary	2.957	.750	.000	2.933	.749	<.001
Criminal history for serious violent, drugs or driving offences	.100	.034	.003	.100	.034	.003
Sex of the offender	-1.137	.723	.116	-1.150	.723	.112
Age at the time of offence	-.014	.027	.604	-.015	.026	.569
Māori cf European	.359	.390	.358	.355	.388	.361
Pacific cf European	.504	.743	.498	.401	.741	.589
Other or unknown cf European	-.106	1.064	.921	-.118	1.064	.912

Appendix C

Sentence notes as a data source

It was originally planned that judges' sentence notes would provide a second source of data to inform the modelling of sentence length. A sentence note is a record of the judge's remarks on sentencing. In the District Courts these are not created for every case. An audio recording is always taken of the judge's remarks but only transcribed in certain circumstances. The remarks are routinely transcribed if an offender is imprisoned or sentenced to home detention. For other sentence types the remarks are only transcribed if the sentence is appealed or the judge otherwise orders a transcript. As a consequence transcripts only exist for a minority of cases but always for incarceration. Ordinarily the transcripts are not easily accessible because they are stored in the paper file for the case at the court. However, for a limited period of time, including 2008-2009, a large holding of sentence notes was assembled at the Law Commission to inform the drafting of sentencing guidelines. The Chief District Court Judge gave permission for transcripts held at the Law Commission to be accessed and for data on the sentencing factors to be coded.

The sentencing notes summarise the judge's view of the case, and may contain information on the offending, any aggravating or mitigating factors the judge saw fit to mention, the purposes of sentencing and the sentencing mathematics. The sentencing mathematics may include the judge's view of the starting point given the seriousness of the offending, plus any addition for aggravating factors (individually or collectively), and then a reduction for mitigating factors (individually or collectively) and a proportionate discount for a guilty plea. Not all sentence notes have all these features. Some judges provide considerable detail and some give very little. The briefest do no more than state the sentence.

A set of guidance material and coding sheets were developed for the extraction of data from the sentence notes. The extraction covered information on the offences being sentenced, sentencing factors, the offender, and the determination of the sentence. The selection of the sentencing factors was based on a review of reference texts on sentencing, appellate decisions, the draft sentencing guidelines developed by the Sentencing Establishment Unit and review of a sample of sentence notes for each offence. The selection was reviewed in light of experience of use after initial experience of coding. The guidance material, copies of the coding sheets and the results of reliability testing are available from the author.

There was always a risk that the sentence notes would not prove to be a viable source of data because the collection was incomplete and would not constitute a random sample drawn across the circuits thus biasing any models based on the data. The proportions of sentence notes available by circuit and the differences between the mean sentence length for the comprehensive administrative dataset and the limited sentence note dataset provided a ready indication of the distortion. There was wide variation in the proportion of sentence notes available per circuit for aggravated drink driving (see Table C1).

Table C1: The number of cases and mean sentence lengths by circuit for the administrative and sentence note datasets for aggravated drink driving

Circuit	Administrative data		Sentence note data		Proportion with sentence notes	Proportionate change in mean sentence
	Cases	Mean sentence (months)	Cases	Mean sentence (months)		
0	95	9.09	69	9.07	73%	-0.2%
1	311	9.71	138	10.21	44%	5.1%
2	80	9.69	52	10.13	65%	4.5%
3	134	10.22	87	9.51	65%	-7.0%
4	46	9.02	38	9.16	83%	1.5%
5	186	10.35	69	8.48	37%	-18.1%
6	249	6.95	164	6.54	66%	-5.8%
7	79	10.34	41	9.39	52%	-9.1%
8	138	7.30	75	7.21	54%	-1.2%
9	46	9.91	8	11.25	17%	13.5%
10	138	8.07	75	7.54	54%	-6.6%
11	136	8.98	103	9.30	76%	3.6%
12	357	9.44	269	9.48	75%	0.5%
13	27	9.48	14	9.00	52%	-5.1%
14	60	9.88	35	9.90	58%	0.3%
15	43	9.36	20	8.88	47%	-5.2%
16	187	11.32	51	11.19	27%	-1.1%
Total	2312	9.28	1308	8.95	57%	-3.5%

The proportion ranged from a low of 17% (Circuit 9) to a high of 83% (Circuit 4) with the other 15 circuits ranged between these extremes. The effect of differing levels of availability may have been limited if the sentencing outcomes for the available sample was similar to the population of incarcerated offenders. The comparison of mean sentence lengths for the sample and the population showed that the distribution of sentence lengths in the samples varied substantially between circuits. At one extreme, Circuit 5, the mean sentence for the sample was almost 2 months (18.1%) below the population of incarcerated

offenders for the circuit whereas in Circuit 9 the sample was about 6 weeks higher (13.5%) than the population for the circuit.

There was a similarly wide variation in the proportion of sentence notes available per circuit for male assaults female (see Table C2). The proportion ranged from a low of 24% (Circuit 16) to a high of 81% (Circuit 11). Excluding the change in Circuit 9 (a difference of 5.11 months (83.3%) based on low sentence volumes), the differences in mean sentence lengths between the sample and the population by circuit ranged from a decrease of about 6 weeks (-20.7%) in Circuit 8 to an increase of 7 weeks (21.4%) in Circuit 5.

Table C2: The number of cases and mean sentence lengths by circuit for the administrative and sentence note datasets for male assaults female

Circuit	Administrative dataset		Sentence note dataset		Proportion with sentence notes	Proportionate change in mean sentence
	Cases	Mean sentence (months)	Cases	Mean sentence (months)		
0	82	6.68	52	7.39	63%	10.7%
1	71	9.44	38	10.12	54%	7.1%
2	34	8.04	19	9.41	56%	17.1%
3	63	8.85	40	9.31	63%	5.2%
4	18	8.56	14	8.71	78%	1.9%
5	76	7.68	31	9.32	41%	21.4%
6	46	6.56	31	5.91	67%	-9.9%
7	20	9.53	14	9.54	70%	0.1%
8	21	6.60	11	5.23	52%	-20.7%
9	9	6.14	4	11.25	44%	83.3%
10	44	7.22	28	6.89	64%	-4.6%
11	53	6.92	43	7.44	81%	7.5%
12	82	5.82	52	6.00	63%	3.1%
13	14	7.21	10	8.20	71%	13.7%
14	16	5.55	7	5.57	44%	0.4%
15	40	6.98	13	6.77	33%	-3.0%
16	34	8.15	8	9.13	24%	12.0%
Total	723	7.44	415	7.85	57%	5.5%

There was wide variation in the proportion of sentence notes available per circuit for burglary (see Table C3). The proportion ranged from a low of 26% (Circuit 9) to a high of 81% (Circuit 6). The differences in mean sentence lengths between the sample and the population by circuit ranged from a decrease of a little less than 3 weeks (-3.6%) in Circuit 5 to an increase of more than 4 months (25.3%) in Circuit 16.

Table C3: The number of cases and mean sentence lengths by circuit for the administrative and sentence note datasets for burglary

Circuit	Administrative dataset		Sentence note dataset		Proportion with sentence notes	Proportionate change in mean sentence
	Cases	Mean sentence (months)	Cases	Mean sentence (months)		
0	150	14.62	118	15.87	79%	8.5%
1	285	18.08	168	19.85	59%	9.8%
2	79	15.98	49	17.68	62%	10.6%
3	243	14.82	136	15.97	56%	7.7%
4	45	15.51	35	15.91	78%	2.6%
5	279	16.74	109	16.14	39%	-3.6%
6	149	17.01	118	18.68	79%	9.9%
7	57	14.89	39	15.76	68%	5.8%
8	80	12.03	52	13.92	65%	15.7%
9	46	13.57	12	17.25	26%	27.2%
10	146	15.23	95	15.93	65%	4.6%
11	133	15.56	95	16.85	71%	8.3%
12	203	14.95	151	14.58	74%	-2.5%
13	67	12.54	32	13.44	48%	7.2%
14	57	16.82	35	18.83	61%	11.9%
15	197	17.76	103	19.60	52%	10.4%
16	122	16.39	39	20.53	32%	25.3%
Total	2338	15.90	1386	16.98	59%	6.8%

These differences indicated that the sentence note data was unlikely to be a viable source of sentencing factor data to inform the sentence length models.

Appendix D

Summary of the tests employed to determine primary offence status

A hierarchy of tests was devised to accept or reject observations. It began with a simple count of the number of charges the offender was sentenced on. If there was only one charge it followed that it must be the most serious offence and the observation was retained.

If there were two or more charges it was necessary to determine which was most serious. Three objective tests based on the charge outcome, the sentence type and the sentence amount were used to attempt to resolve the status of cases with multiple convictions. If only one charge resulted in a conviction and sentence and the others were discharged, the convicted and sentenced charge was the most serious and was retained if it was a subject offence. Otherwise it was rejected.

If the charge outcome did not identify one charge as the most serious, the sentence types were examined to identify if one charge received a more serious sentence in terms of the hierarchy of types set out in the Sentencing Act 2002. If one charge attracted a more serious sentence type than any of the others that charge was recognised as the most serious and was retained if it was a subject offence. Otherwise it was rejected.

If the sentence type did not identify one charge as the most serious sentence, the magnitude of the sentences of that type were examined to determine if one charge received a larger or longer sentence than the other(s). If one charge did attract a larger or longer sentence that charge was the most serious and was retained if it was a subject offence. Otherwise it was rejected.

If none of these tests resolved which charge was most serious the selection was determined by the generic seriousness of the tied offence types. The generic seriousness was based on a seriousness scale maintained by the Ministry of Justice that ranked offences according to the average days of imprisonment, with non-custodial sentences scoring zero. If a single charge was identified as the most serious the case was accepted if the charge was for burglary, male assaults female or drink driving and rejected if it was any other offence. If multiple charges remained tied the case was rejected if none of the tied offences were burglary, male assaults female or drink driving offences.

Some cases survived this sequence of tests because two or more burglaries or two or more drink driving offences or two or more male assaults female charges had attracted the same charge outcome and/or sentence and were of equal seriousness. Although such

cases were retained it was necessary to identify one of the charges as the lead offence to build the case around. The first test favoured any charge that was laid indictably on the assumption that indictably laid charges were more likely to be more serious than summarily laid charges. If some charges were still tied the offence committed most recently was chosen on the basis that the offender had gone on to re-offend having already offended. If there was still a tie a random choice was made between the remaining charges.

Appendix E

The predictor variables included in the modelling

Table E1: Demographic variables and the circuit variable

Variable	Coding	Expected Effect
Gender	Male (0) Female (1)	Uncertain
Ethnicity	European (0) Māori (1) Pacific People (2) Other or unknown (3)	Uncertain
Age at Offence¹	Continuous variable or a categorical variable separately specified for each offence type. Drink driving: Aged 17-31 (0) Aged 32-39 (1) Aged 40-47 (2) Aged 48+ (3) Male assaults female: Aged 17-23 (0) Aged 24-30 (1) Aged 31-38 (2) Aged 39+ (3) Burglary: Aged 17-19 (0) Aged 20-29 (1) Aged 30-39 (2) Aged 40+ (3)	Uncertain
Circuit	Dummy variables coded 0-16 covering the 17 circuits	Uncertain effect but expected to vary by circuit

¹ Age was specified as a continuous variable for the sentence length model but was categorised for the in-out models because it violated the assumption of linearity. For male assaults female and drink driving the age of offenders was relatively evenly spread and the variable was formed by splitting the offenders into four “quarters” as close to equal as possible while placing all offenders of the same year of age in the same quarter. This split into quarters did not make sense for burglary as a high proportion of offenders were very young (39% 19 or younger). Teenagers formed the first category, followed by two ten year bands and a catch-all category for older offenders

Table E2: Drink driving offence seriousness, case processing and criminal history variables

Variable	Coding	Expected Effect
<i>Offence seriousness and case processing variables</i>		
Alcohol level	A continuous variable with alcohol levels standardised based on the proportionate amount above the legal limit. Refusal of a test was assigned a nominal level 2.25 times above the limit. Offenders with no reading included with at a nominal reading 1.9 times the limit.	Aggravating
Multiple drink drive offences	Categorical variable: 0 (no) 1 (yes)	Aggravating
Concurrent driving while disqualified offence	Categorical variable: 0 (no) 1 (yes)	Aggravating
Concurrent serious driving offence	Categorical variable: 0 (no) 1 (yes)	Aggravating
Concurrent minor driving offence	Categorical variable: 0 (no) 1 (yes)	Aggravating
Concurrent breach of sentence offence	Categorical variable: 0 (no) 1 (yes)	Aggravating
Concurrent breach of bail offence	Categorical variable: 0 (no) 1 (yes)	Aggravating
Total convictions for sentence¹	Continuous variable or a categorical variable: 0 (one charge) 1 (two charges) 2 (three or more charges)	Aggravating
Time from charge to sentence¹	Continuous variable in days or a categorical variable coded: 0 (≤ 6 weeks) less) 1 (> 6 weeks ≤ 3 months) 2 (> 3 months ≤ 6 months) 3 (> 6 months)	Aggravating
Plea	Categorical variable coded: 0 guilty plea 1 not guilty plea	Aggravating ²
Judge type	Categorical variable coded: 0 Criminal 1 Family	Uncertain

Variable	Coding	Expected Effect
<i>Criminal history variables</i>		
Prior convictions for drink driving¹	Continuous variable or a categorical variable coded: 0 (2 prior convictions) 1 (3 priors) 2 (4 priors) 3 (5 or 6 priors) 4 (7 or more priors)	Aggravating
Convictions for drink driving in the last ten years	Categorical variable coded: 0 (none) 1 (1) 2 (2) 3 (3 or more)	Aggravating
Elapsed time since the last conviction¹	Continuous variable in years, or a categorical variable reverse coded: 0 (over five years ago) 1 (over 18 months ago and five or fewer years ago) 2 (18 months or more recent)	Mitigating Aggravating
Prior incarceration	Categorical variable coded: 0 (no prior incarceration) 1 (prior incarceration but not for drinking and driving) 2 (prior incarceration, including for drinking and driving)	Aggravating
Prior incarceration for drink driving	Categorical variable: 0 (no) 1 (yes)	Aggravating
Longest prior incarceration for drink driving	Continuous variable in months	Aggravating
Prior convictions for disqualified driving¹	Continuous variable or a categorical variable coded: 0 (no prior convictions) 1 (1) 2 (2) 3 (3 or 4) 4 (5 or more)	Aggravating
Prior convictions for any offence^{1,3}	Categorical variable coded: 0 (2 - 5) 1 (6 - 11) 2 (12 -24) 3 (25)	Aggravating
Recent conviction for breach of sentence	Categorical variable coded: 0 (no) 1 (yes, convicted in the previous five years)	Aggravating

¹ The variable was converted from a continuous to a categorical variable because it breached the assumption of linearity.

² The explanation differs from the other factors because a not guilty plea is not an aggravating factor in sentencing terms, the effect of a not guilty plea is to lose the benefit of mitigation afforded to a guilty plea.

³ An offender's total convictions for any offence were capable of being expressed in many different ways. An offender's overall criminal history expressed as the number of convictions could be divided in many categories with choices as to the number of categories and the boundaries between them. The formulation presented in the table splits the data into four quarters.

Table E3: Male assaults female offence seriousness, case processing and criminal history variables

Variable	Coding	Expected Effect
<i>Offence seriousness and case processing variables</i>		
Multiple male assaults female offences	Categorical variable: 0 (no) 1 (yes)	Aggravating
Concurrent threat to kill or do GBH offence	Categorical variable: 0 (no) 1 (yes)	Aggravating
Concurrent breach of protection order offence	Categorical variable: 0 (no) 1 (yes)	Aggravating
Concurrent other violent offence	Categorical variable: 0 (no) 1 (yes)	Aggravating
Concurrent wilful or intentional damage offence	Categorical variable: 0 (no) 1 (yes)	Aggravating
Charge type indicates weapon involved	Categorical variable: 0 (no) 1 (yes)	Aggravating
Concurrent breach of bail offence	Categorical variable: 0 (no) 1 (yes)	Aggravating
Concurrent breach of sentence offence	Categorical variable: 0 (no) 1 (yes)	Aggravating
Time from charge to sentence¹	Continuous variable in days or a categorical variable coded: 0 (≤ 6 weeks) less) 1 (> 6 weeks ≤ 3 months) 2 (> 3 months ≤ 6 months) 3 (> 6 months)	Aggravating
Plea	Categorical variable coded: 0 guilty plea 1 not guilty plea	Aggravating ²

Variable	Coding	Expected Effect
Charge laid indictably	Categorical variable coded: 0 charge laid summarily 1 charge laid indictably	Aggravating
Sentenced in an active family violence court	Categorical variable coded: 0 no 1 yes	Uncertain
Judge type	Categorical variable coded: 0 Criminal 1 Family	Uncertain
<i>Criminal history variables</i>		
Prior convictions for male assaults female¹	Continuous variable or a categorical variable coded: 0 (none) 1 (1) 2 (2) 3 (3 or more)	Aggravating
Conviction for male assaults female in the last five years	Categorical variable coded: 0 (no) 1 (yes)	Aggravating
Prior incarceration for male assaults female	Categorical variable coded: 0 (no) 1 (yes)	Aggravating
Longest prior incarceration for male assaults female	Continuous variable in months	Aggravating
Prior conviction for serious violence	Categorical variable coded: 0 (no) 1 (yes)	Aggravating
Conviction for serious violence in the last five years	Categorical variable coded: 0 (no) 1 (yes)	Aggravating
Prior conviction for breach of sentence in the last five years	Categorical variable coded: 0 (no) 1 (yes)	Aggravating
Prior convictions for any offence^{1,3}	Continuous variable or a categorical variable coded: 0 (none) 1 (1 - 9) 2 (10 -19) 3 (20 or more)	Aggravating
Prior incarceration for any offence	Categorical variable coded: 0 (no) 1 (1) 2 (2 – 4) 3 (5 or more)	Aggravating

Variable	Coding	Expected Effect
Longest prior incarceration for any offence	Continuous variable in months	Aggravating

¹ The variable was converted from a continuous to a categorical variable because it breached the assumption of linearity.

² The explanation differs from the other factors because a not guilty plea is not an aggravating factor in sentencing terms, the effect of a not guilty plea is to lose the benefit of mitigation afforded to a guilty plea.

³ An offender's total convictions for any offence were capable of being expressed in many different ways. An offender's overall criminal history expressed as the number of convictions could be divided in many categories with choices as to the number of categories and the boundaries between them. The formulation presented in the table splits the data into four quarters.

Table E4: Burglary offence seriousness, case processing and criminal history variables

Variable	Coding	Expected Effect
<i>Offence seriousness and case processing variables</i>		
Number of burglary charges for sentence	Continuous variable or categorical variable coded: 0 (1) 1 (2) 2 (3 or more)	Aggravating
Approximate value of lead burglary offence	Categorical variable coded: 0 (less than \$500 or unstated) 1 (\$500 to \$5,000) 2 (more than \$5,000)	Aggravating
Burglary by day or night	Categorical variable coded: 0 (day) 1 (night) 2 (unknown)	Aggravating (if at night)
Number of burglary offence dates	Categorical variable coded: 0 (one) 1 (two) 2 (three or more)	Aggravating
Concurrent receiving charge	Categorical variable coded: 0 (no) 1 (yes)	Aggravating
Concurrent unlawful taking charge	Categorical variable coded: 0 (no) 1 (yes)	Aggravating
Concurrent other dishonesty charge	Categorical variable coded: 0 (none) 1 (1) 2 (2 or more)	Aggravating

Variable	Coding	Expected Effect
Concurrent violent offence charge	Categorical variable coded: 0 (no) 1 (yes)	Aggravating
Concurrent drug charge	Categorical variable coded: 0 (no) 1 (yes)	Aggravating
Concurrent serious driving charge	Categorical variable coded: 0 (no) 1 (yes)	Aggravating
Concurrent breach of bail offence	Categorical variable coded: 0 (no) 1 (yes)	Aggravating
Concurrent breach of sentence offence	Categorical variable coded: 0 (no) 1 (yes)	Aggravating
Concurrent other serious (receiving, unlawful taking, serious violence or serious driving) charges	Categorical variable coded: 0 (no) 1 (yes)	Aggravating
Total convictions for sentence	Continuous variable or categorical variable coded: 0 (1) 1 (2) 2 (3 or 4) 3 (5, 6 or 7) 4 (8 or more)	Aggravating
Time from charge to sentence	Continuous variable or categorical variable coded: 0 (≤ 6 weeks) less) 1 (> 6 weeks ≤ 3 months) 2 (> 3 months ≤ 6 months) 3 (> 6 months)	Aggravating
Plea	Categorical variable coded: 0 guilty plea 1 not guilty plea	Aggravating ²
Charge laid indictably	Categorical variable coded: 0 charge laid summarily 1 charge laid indictably	Aggravating
Reparation order made	Categorical variable coded: 0 no 1 yes	Mitigating

Variable	Coding	Expected Effect
Reparation amount	Categorical variable coded: 0 (none) 1 (<\$500) 2 (≥\$500 <\$1,000) 3 (≥\$1,000 <\$2,500) 4 (≥\$2,500)	Mitigating
Judge type	Categorical variable coded: 0 Criminal 1 Family	Uncertain
<i>Criminal history variables</i>		
Prior convictions for burglary	Continuous variable or categorical variable coded: 0 (no prior convictions) 1 (1 prior) 2 (2 – 4 priors) 3 (5 or more)	Aggravating
Conviction for burglary in the previous five years	Categorical variable coded: 0 no 1 yes	Aggravating
Prior convictions for serious dishonesty excluding burglary	Continuous variable or categorical variable coded: 0 (no prior convictions) 1 (1 prior) 2 (2 or 3 priors) 3 (4 – 9 priors) 4 (10 or more)	Aggravating
Prior convictions for serious dishonesty including burglary	Continuous variable or categorical variable coded: 0 (no prior convictions) 1 (1 or 2 priors) 2 (3 – 9 priors) 3 (10 or more)	Aggravating
Prior convictions for serious violence	Categorical variable coded: 0 no 1 yes	Aggravating
Prior convictions for serious violence, drug dealing or serious driving	Continuous variable	Aggravating
Conviction for breach of sentence in the past five years	Categorical variable coded: 0 no 1 yes	Aggravating
Prior convictions for any offence^{1, 3}	Continuous variable or categorical variable coded: 0 (no prior convictions) 1 (1 - 9 priors) 2 (10 -19 priors) 3 (20 or more priors)	Aggravating

Variable	Coding	Expected Effect
Prior incarceration for any offence	Categorical variable coded: 0 (no prior incarceration) 1 (one prior incarceration) 2 (2 – 4 prior incarcerations) 3 (5 or more prior incarcerations)	Aggravating
Longest prior incarceration for burglary	Categorical variable coded: 0 (no prior incarceration) 1 (≤ 12 months) 2 (>12 months ≤ 24 months) 3 (>24 months)	Aggravating
Longest prior incarceration for any offence	Continuous variable	Aggravating

¹ The variable was converted from a continuous to a categorical variable because it breached the assumption of linearity.

² The explanation differs from the other factors because a not guilty plea is not an aggravating factor in sentencing terms, the effect of a not guilty plea is to lose the benefit of mitigation afforded to a guilty plea.

³ An offender's total convictions for any offence were capable of being expressed in many different ways. An offender's overall criminal history expressed as the number of convictions could be divided in many categories with choices as to the number of categories and the boundaries between them. The formulation presented in the table splits the data into four quarters.

Appendix F

The Impact of Cumulative Sentencing and Home Detention on the Aggravated Drink Driving and Male Assaults Female Sentence Length Models

When the high values (>3) for the standardized residual for both aggravated drink driving and male assaults female were scrutinised a pattern was readily apparent. The cases were associated with two sentencing outcomes; the offender either received a cumulative sentence, or was sentenced to home detention (rather than imprisonment).

For aggravated drink driving the standardised residuals for 24 cases exceeded an absolute value of 3. All bar one of these was due to actual sentences substantially higher than the predicted sentence. The cases were spread across the circuits with no apparent concentrations and they were spread across the demographic variables. Scrutiny of the criminal history seriousness variables confirmed that these were cases likely to result in high actual sentences. Ten of the 24 cases had the sentence type of home detention. Of the remaining 14 cases, 13 had cumulative terms of imprisonment.

For male assaults female the standardised residuals for 6 cases exceeded an absolute value of 3. All involved actual sentences substantially higher than the predicted sentence. The cases were all from different circuits, they were spread across the demographic variables, and scrutiny of the seriousness and criminal history variables confirmed that these were cases likely to result in high actual sentences. Two of the six cases had the sentence type of home detention and three of the four remaining cases had cumulative terms of imprisonment.

The imposition of cumulative sentences and the choice of home detention over imprisonment are significant characteristics; they are part of the sentencing decision and are not what would normally be thought of as sentencing factors. Cumulative versus concurrent sentencing is a choice of how to express the sentence when there are multiple charges. Faced with multiple charges the judge has to decide whether to impose separate sentences that cumulate to form the total sentence or whether to impose a longer sentence on the lead charge to reflect the totality of the offending and concurrent regular sentences on the other charge(s). The judge must also choose the sentence type within the incarceration grouping. If a judge has decided that a short term of imprisonment (two years or less) is appropriate the judge then determines whether the sentence should be imprisonment or home detention taking account of the Sentencing Act requirement to impose the least restrictive sentence that is appropriate. For both aggravated drink driving

and male assaults female this decision is regularly required because when a judge decides to imprison, home detention is always an option. The only exception would be when the judge has determined that the totality of offending requires a cumulative sentence in excess of two years.

In retrospect it is not surprising that some cumulative sentences would be associated with high residuals. Cumulative sentences are sometimes imposed when the maximum sentence available for the lead offence is not sufficient to reflect the totality of the offending in the case.

The association of high residuals with cases where home detention was imposed was unexpected. Retrospective analysis revealed that the mean home detention sentence was 1.5 months longer than the mean imprisonment sentence (10.2 c.f. 8.7 months) for aggravated drink driving and 2.1 months longer (9.3 c.f. 7.2 months) for male assaults female.

This was surprising for two reasons. First, it was assumed that home detention was more likely to be “the least restrictive sentence appropriate” for a higher proportion of lower seriousness cases and that it would be judged inappropriate for a higher proportion of more serious cases. Secondly, there may be concerns about the capacity of offenders to comply with longer terms of home detention. The self-discipline required and the pressures of being restricted to the home may be too great if the term is too long.

Investigation of the data indicated that the initial assumption that home detention would form a higher proportion of incarceration sentences for less serious offending was wrong. Review of the sentences revealed that the difference in average sentences was due to the substantial proportion of very short imprisonment sentences (less than 6 months) and the small number of home detention sentences in that space. The relative proportions are listed in Tables F1 (aggravated drink driving) and F2 (male assaults female).

Table F1: Relative proportions of imprisonment and home detention sentences for aggravated drink driving by categories of sentence length

Sentence length	Imprisonment		Home detention	
	#	%	#	%
3 months or less	164	11.8%	5	0.5%
> 3 months < 6 months	220	15.9%	33	3.6%
6 months	248	17.9%	146	15.8%
> 6 months ≤ 9 months	320	23.1%	273	29.5%
> 9 months < 12 months	88	6.3%	96	10.4%
12 months	129	9.3%	223	24.1%
> 12 months	219	15.8%	148	16.0%
TOTAL	1388	100%	924	100%

For aggravated drink driving 27.7% of imprisonment terms were less than 6 months compared to 4.1% of home detention sentences. For male assaults female 38.7% of imprisonment terms were less than 6 months compared to 9.8% of home detention sentences. A proportion of the very short imprisonment sentences will have been in circumstances where the offender was remanded in custody and the sentence reflects the time in custody. Otherwise it would seem that judges consider that a prison sentence of 3 months (actual time served 6 weeks) can serve a valid purpose but they do not attach the same value to serving 6 weeks on home detention. However, the distribution also supports the notion that long home detention sentences may be too difficult. After adjustment to match the term of home detention to imprisonment 79.8% of home detention sentences for aggravated drink driving and 78.0% for male assaults female fell within the range 6 to 12 months (meaning that the offenders serve 3 to 6 months).

Table F2: Relative proportions of imprisonment and home detention sentences for male assaults female by categories of sentence length

Sentence length	Imprisonment		Home detention	
	#	%	#	%
3 months or less	138	21.5%	3	3.7%
> 3 months < 6 months	110	17.2%	5	6.1%
6 months	115	17.9%	15	18.3%
> 6 months ≤ 9 months	114	17.8%	26	31.7%
> 9 months < 12 months	42	6.6%	7	8.5%
12 months	57	8.9%	16	19.5%
> 12 months	65	10.1%	10	12.2%
TOTAL	641	100%	82	100%

Although none of the observations were a cause for concern on their individual characteristics the recognition that they were either cumulative sentences or home detention was a major concern because the rates of both are uneven across the circuits.

The proportions of sentences imposed as home detention for aggravated drink driving (Table F3) varied widely across the circuits ranging from 15% (Circuit 8) to 67% (Circuit 9) with no substantial clustering. The proportion of sentences imposed as part of a cumulative package of sentences also varied, from 0% (Circuit 13) to 23% (Circuit 10) but there was a cluster of 10 circuits in the 11-15% range near the centre of the spectrum.

Table F3: The proportion of incarcerated aggravated drink drivers sentenced to home detention or who received a cumulative sentence of incarceration

	Incarcerated	Home detention		Cumulative sentence	
		#	%	#	%
Circuit 0	95	31	33%	12	13%
Circuit 1	311	131	42%	47	15%
Circuit 2	80	20	25%	10	13%
Circuit 3	134	43	32%	16	12%
Circuit 4	46	11	24%	7	15%
Circuit 5	186	71	38%	13	7%
Circuit 6	249	53	21%	10	4%
Circuit 7	79	35	44%	10	13%
Circuit 8	138	21	15%	18	13%
Circuit 9	46	31	67%	5	11%
Circuit 10	138	46	33%	32	23%
Circuit 11	136	51	38%	19	14%
Circuit 12	357	228	64%	31	9%
Circuit 13	27	12	44%	0	0%
Circuit 14	60	14	23%	13	22%
Circuit 15	43	27	63%	3	7%
Circuit 16	187	99	53%	24	13%
TOTAL	2312	924	40%	270	12%

The proportions of sentences imposed as home detention for male assaults female (Table F4) also varied widely across the circuits. Leaving aside Circuit 9 (no use of home detention), the proportions ranged from 2.2% (Circuits 6) to 33.3% (Circuit 4) with no substantial clustering. The proportion of sentences imposed as part of a cumulative package of sentences also varied. Again leaving aside Circuit 9 (no cumulative sentences), the proportions ranged from 7.1% (Circuit 13) to 47.7% (Circuit 10) with a limited cluster of six circuits in the 22-28% range near the centre of the spectrum.

In order to understand the implications for the model and its interpretation, in particular the parameter estimates for the circuit variables, the models for both offences were re-run twice. The first re-run included a cumulative sentence variable (0 = no cumulative sentence, 1 = cumulative sentence) and the second re-run omitted the cumulative sentence variable and included a home detention variable (0 = imprisonment, 1 = home detention). In theory neither of the variables should affect the models and neither should be statistically significant. In theory the choice of cumulative versus concurrent sentencing should be neutral; the choices are just alternate ways of packaging what should be the same sentence reflecting the totality of the offending and the circumstances of the

offender. In theory the choice of home detention versus imprisonment should be neutral in terms of sentence length; they are just alternative ways of serving the same term of incarceration. However, the widely varied rates of use of home detention and cumulative sentences across the circuits suggested that they may not be neutral because different approaches were being followed in different places.

Table F4: The proportion of offenders incarcerated for male assaults female sentenced to home detention or who received a cumulative sentence of incarceration

	Incarcerated	Home detention		Cumulative sentence	
		#	%	#	%
Circuit 0	82	4	5%	13	16%
Circuit 1	71	11	15%	20	28%
Circuit 2	34	4	12%	13	38%
Circuit 3	63	4	6%	22	35%
Circuit 4	18	6	33%	4	22%
Circuit 5	76	3	4%	7	9%
Circuit 6	46	1	2%	7	15%
Circuit 7	20	3	15%	7	35%
Circuit 8	21	3	14%	3	14%
Circuit 9	9	0	0%	0	0%
Circuit 10	44	3	7%	21	48%
Circuit 11	53	4	8%	7	13%
Circuit 12	82	17	21%	19	23%
Circuit 13	14	3	21%	1	7%
Circuit 14	16	1	6%	4	25%
Circuit 15	40	4	10%	10	25%
Circuit 16	34	11	32%	8	24%
TOTAL	723	82	11%	166	23%

None of the adjusted models supported the proposition that the effects would be neutral. When the cumulative sentence variable was added to the aggravated drink driving model as an additional fifth block it made a significant difference to the model. The model R^2 increased from .385 to .414, and the R^2 change (.029) was significant ($p < .001$). The cumulative sentence variable was statistically significant ($B = 3.120$, $S.E. = 0.295$, $p < .001$).

When the home detention variable was added to the aggravated drink driving model it also made a significant and more substantial difference to the model. The model R^2 increased from .385 to .474, and the R^2 change (.089) was significant ($p < .001$). The home detention variable was statistically significant ($B = 3.529$, $S.E. = 0.180$, $p < .001$).

When the cumulative sentence variable was added to the male assaults female model as a fifth block it made a significant difference to the model. The model R^2 increased from .263 to .296, the R^2 change (.032) was significant ($p < .001$). The cumulative sentence variable was statistically significant ($B = 2.321$, $S.E. = 0.410$, $p < .001$).

When the home detention variable was added to the male assaults female model it also made a significant and more substantial difference to the model. The model R^2 increased from .263 to .305, the R^2 change (.042) was significant ($p < .001$). The home detention variable was statistically significant ($B = 3.240$, $S.E. = 0.502$, $p < .001$).

These findings do not conclusively invalidate the proposition that the effects of the choices of concurrent or cumulative sentences and home detention or imprisonment will be neutral. It is possible the results reflect the limitations of the models and the cumulative sentence and home detention variables may be proxies for other variables unable to be included in the models.

The critical question was whether the cumulative sentence and/or the home detention variable should be included in the final model. The argument against doing so is that both variables are elements of the sentencing outcome and that including them means that an element of the outcome is being used to predict the outcome. The decision to sentence cumulatively may be a consequence of the judge wishing to impose a more severe sentence than concurrent sentencing would allow. Because the rate of use of cumulative sentences varies between circuits, including the cumulative sentence variable may result in mis-estimation of the parameters for the circuit variables and suppress the degree of variability between circuits. A comparison of the parameter estimates for the circuit variables between the original model and the adjusted model including the cumulative sentence variable confirmed that the degree of variability reduced. In general if the circuit used home detention or cumulative sentences at a higher rate than the reference circuit the parameter estimate for the circuit reduced and if they were used at a lower rate the parameter estimates increased.

Although the inclusion of the cumulative sentence variable improved the model it did so at the cost of obscuring or suppressing variability between the circuits that is at least partially due to differences in sentencing approach. Therefore no adjustment to the model was justified.

Three reasons why use of home detention might vary between circuits and affect the parameter estimates were identified and assessed to determine if they justified

adjustment of the model. The availability of home detention is likely to vary between circuits for technical reasons. Some circuits may have a higher proportion of residences that cannot support the sentence resulting in a higher rate of use of imprisonment irrespective of the preference of the judge in any individual case. Assuming that incarceration sentence type has no or little influence on sentence amount (an assumption supported by the data collected from sentence notes and reported in Chapter 6 showing that the ratio of sentences expressed as prison terms before selecting home detention to the actual home detention sentence was 1:1.01) the inclusion of the home detention variable would result in mis-estimation of parameter estimates because an effect actually due to the availability of the sentence would be treated as if it were a sentencing factor used to determine sentence length.

The second possibility was that the variation reflects differences between circuits about the appropriate use of home detention in general or for this offence in particular. If different circuits were more or less liberal in the use of home detention, more or less inclined to consider it the least restrictive sentence appropriate and the differences were driven by different thresholds of seriousness and/or criminal history factors, the choice of sentence type could appear to contribute to variation in sentencing. However, similar to the cumulative sentencing choice the choice of sentence type is an intrinsic part of the sentencing outcome and including it would result in the outcome being used to predict the outcome and result in mis-estimation of the parameter estimates for other variables.

The third possibility considered was that home detention was a proxy for other sentencing variables that were not included. If that were the case inclusion of the home detention variable would improve the model and would be justified on the basis that it stood in the place of some unspecified valid predictor variable(s). Doing so would be problematic for two reasons. Describing what the variable is a proxy for can only be done in vague general terms. Secondly, to have a significant influence on the parameter estimates for the circuit variables the unknown variables would have to be unevenly distributed across the circuits. There is no reasonable basis for assuming that to be the case. In the absence of any alternative indications it is more reasonable to assume that such factors are evenly distributed.

Although the inclusion of the home detention variable improved the model none of the possible explanations for the effect justify its inclusion. There is a risk that the resultant parameter estimates for the circuit variables would be mis-estimated.

Appendix G

Correlation Matrices for Circuit Contextual Variables for Multi-level Modelling

The correlations between the various circuit variables are set out in Tables G1 & G2 below. Table G1 covers the correlations between each of the generic variables that apply across the models irrespective of offence. Table G2 covers the correlations between the generic variables and the four offence specific variables plus the correlations between the respective offence specific variables for each offence.

Table G1: Correlation matrix for the generic contextual variables

Circuit Contextual Variables	1	2	3	4	5	6	7	8	9	10	11	12	13
1 Metropolitan or provincial court	1	-.527*	-.847**	-.648**	.770**	-.295	-.602*	-.301	.427	.394	-.359	.357	-.668**
2 Population aged 17+		1	.583*	.325	-.618**	.108	.853**	.531*	-.482	-.417	.137	-.402	.561*
3 Median personal income			1	.433	-.827**	.083	.658**	.134	-.438	-.445	.236	-.425	.683**
4 Racial homogeneity				1	-.321	.802**	.557*	.269	.197	-.296	.059	-.260	.588*
5 Proportion adults with no educational qualifications					1	-.022	-.636**	-.151	.535*	.432	-.456	.419	-.698**
6 Proportion of adults unemployed						1	.361	.426	.467	-.195	-.074	-.168	.309
7 Number of Judicial Officers (FTEs)							1	.535*	-.289	-.711**	-.126	-.693**	.700**
8 Number of Judges (FTEs)								1	-.153	-.378	-.088	-.350	.308
9 New criminal cases for 000 adults									1	.370	-.203	.353	-.251
10 New criminal cases per Judicial Officer FTE										1	.325	.997**	-.509*
11 Average active cases per Judicial Officer FTE											1	.319	.344
12 Cases disposed per Judicial Officer FTE												1	-.498*
13 Hearing rate													1

*. Correlation is significant at the 0.05 level (2-tailed).

**.. Correlation is significant at the 0.01 level (2-tailed).

Table G2: Correlation matrix for the offence specific contextual variables and including the correlations between the generic variables and each offence specific variable

	Aggravated drink driving				Male assaults female				Burglary			
	14	15	16	17	14	15	16	17	14	15	16	17
1 Metropolitan or provincial court	-.700**	.623**	.157	.219	-.806**	.758**	.265	.376	-.790**	.698**	.308	.374
2 Population aged 17+	.370	-.377	-.307	-.158	.483*	-.491*	-.346	-.285	.427	-.386	-.410	-.211
3 Median personal income	.690**	-.611**	-.259	-.241	.816**	-.801**	-.221	-.266	.700**	-.584*	-.307	-.296
4 Racial homogeneity	.526*	-.537*	.469	.246	.486*	-.498*	.395	.217	.536*	-.484*	.279	.011
5 Proportion adults with no educational qualifications	-.738**	.673**	.342	.296	-.821**	.754**	.388	.406	-.785**	.723**	.539*	.461
6 Proportion of adults unemployed	.264	-.324	.470	.124	.132	-.180	.602*	.322	.240	-.258	.316	-.044
7 Number of Judicial Officers (FTEs)	.490*	-.491*	-.171	-.232	.569*	-.581*	-.120	-.194	.508*	-.463	-.324	-.349
8 Number of Judges (FTEs)	-.007	-.003	-.334	-.426	.056	-.031	-.188	-.302	.105	-.081	-.522*	-.561*
9 New criminal cases for 000 adults	-.263	.202	.777**	.500*	-.473	.443	.926**	.763**	-.379	.371	.655**	.323
10 New criminal cases per Judicial Officer FTE	-.471	.482	.313	.585*	-.554*	.590*	.250	.487*	-.470	.485*	.382	.605*
11 Average active cases per Judicial Officer FTE	.527*	-.490*	.000	.091	.407	-.338	-.253	-.245	.585*	-.550*	-.274	-.158
12 Cases disposed per Judicial Officer FTE	-.471	.488*	.298	.573*	-.537*	.575*	.242	.479	-.461	.482*	.377	.601*
13 Hearing rate	.725**	-.685**	.007	-.027	.678**	-.660**	-.069	-.105	.798**	-.734**	-.367	-.391
14 Mean days charge to sentencing	1	-.982**	.061	-.131	1	-.982**	-.293	-.484*	1	-.967**	-.336	-.451
15 Proportion of cases sentenced within 3 months of charge		1	-.134	.081	-.982**	1	.237	.449	-.967**	1	.326	.447
16 Sentencings per 000 adults			1	.807**	-.293	.237	1	.855**	-.336	.326	1	.809**
17 Sentencings per Judge FTE				1	-.484*	.449	.855**	1	-.451	.447	.809**	1

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix H

Using the sentence notes to identify omitted variables

Although the data drawn from the sentence notes could not be used to improve the linear regression models for sentence length it could be used to gain insight into potential omitted variables. This was done by forming a new dataset that combined the variables from the administrative dataset with the additional variables available from the sentencing notes and modelling sentence length to identify any sentence note variables that were significant. This process enables a tentative recognition of any variables that were important to the judges and were not included in the models. The findings are summarised below.

Aggravated drink driving

Sentence note variables were available for 1,339 of the 2,312 cases included in the administrative dataset. However, 31 were excluded because the sentence note revealed the offender was a sentenced prisoner at the time of sentencing (16 cases); the sentencing described a re-sentencing situation (13 cases) and did not cover the original sentence considered to be warranted for the offending; or the sentence note was incomplete (2 cases) - leaving a final combined dataset of 1,308 cases. An additional 16 variables (see Table H1) were available.

Table H1: Descriptive statistics for the sentence length decision sentencing variables sourced from sentence notes for aggravated drink driving ($n = 1,308$)

Categorical Variables	Sentence notes	
	N	%
Risk to others	277	21.2%
Damage caused	92	7.0%
Attempt to avoid detection	127	9.7%
Offended while on bail	108	8.3%
Offended while subject to a sentence	89	6.8%
Other aggravation	50	3.8%
Positive steps to reform	155	11.9%
Remorse	68	5.2%
Other mitigation	89	6.8%
Poor response to a prior sentence	224	17.1%
Home detention inappropriate	182	13.9%
Leave to apply for a substituted sentence	63	4.8%
Offender on a final warning	123	9.4%
Offender personal mitigating circumstances	203	15.5%

To determine if any of these variables were significant, an augmented model including an additional block of sentence note seriousness variables was run. Three of the variables were found to be significant: risk to other road users (a tangible recognition by the judge that the offender was a risk); offending while on bail (that all or some of the offending being sentenced occurred while the offender was on bail); and the granting of leave for the offender to seek to substitute home detention for imprisonment. The latter decision typically occurs when the offender could not offer an acceptable residence but the judge agrees that home detention is the least restrictive sentence appropriate. It may be a proxy for unobserved mitigating circumstances; the negative co-efficient supports this proposition. However, there is also a risk that including it at least partially uses an element of the sentence outcome to predict sentence length.

The additional variables only made a modest contribution to the overall model ($\Delta R^2 = .011$, $p < .001$). Parameter estimates for the three variables are set out in Table H2.

Table H2: Parameter estimates for the statistically significant sentence note variables added to the sentence length model for aggravated drink driving ($n = 1,308$)

Variables	B	S.E.	t	Sig.
Posed risk to other road users	.703	.275	2.557	.011
Offended on bail	1.968	.485	4.058	.000
Leave granted to apply to substitute HD for imprisonment	-1.138	.510	-2.230	.026

Inclusion of the variables would improve the main model but on these results the improvement would be marginal. As it is unlikely that the additional variables would be unevenly distributed across the circuits it is not plausible that addition of the variables would have a substantial impact on the estimates of variation between the circuits.

Male assaults female

Sentence note variables were available for 415 of the 723 cases included in the administrative dataset. An additional 17 variables (see Table H3) were available.

Table H3: Descriptive statistics for the sentence length decision sentencing variables sourced from sentence notes for male assaults female (*n* = 415)

Categorical Variables	Sentence notes	
	N	%
Degree of Physical Injury		
None, none mentioned	199	48.0%
Minor injury	167	40.2%
Moderate or serious injury	49	11.8%
Level of violence		
Minor or unknown level of violence	156	37.6%
Moderate violence	129	31.1%
Serious violence	130	31.3%
Victim suffered psychological or emotional impact		
No, not mentioned	358	86.3%
Yes	57	13.7%
Prior assault on the same victim by this offender		
None, not mentioned	318	76.6%
Yes	97	23.4%
Weapon used or threatened to be used		
No, not mentioned	393	94.7%
Yes	22	5.3%
Offence involved particular cruelty		
No, not mentioned	393	94.7%
Yes	22	5.3%
Child present during the assault		
No, not mentioned	353	85.1%
Yes	62	14.9%
Offending occurred while the offender was on bail		
No, not mentioned	367	88.4%
Yes	48	11.6%
Offending occurred while the offender was subject to a sentence		
No, not mentioned	333	80.2%
Yes	82	19.8%
Other aggravation (includes		
None, none mentioned	387	93.3%
Yes	28	6.7%

Categorical Variables	Sentence notes	
	N	%
Other mitigation (includes		
None, none mentioned	330	79.5%
Yes	85	20.5%
Offender described as responding poorly to previous sentence(s)		
No, not mentioned	329	79.3%
Yes	86	20.7%
Leave to apply to substitute home detention for imprisonment granted		
No	402	96.9%
Yes	13	3.1%
Risk of re-offending		
Low, not mentioned	332	80.0%
Moderate or high	83	20.0%
Personal circumstances		
None, none mentioned	363	87.5%
Yes	52	12.5%
Offending occurred under influence of alcohol or drugs		
No, not mentioned	314	75.7%
Yes	101	24.3%
Offender had alcohol or drug problem		
No, not mentioned	312	75.2%
Yes	103	24.8%

To determine if any of these variables were significant an augmented model including an additional block of sentence note seriousness variables was run. Three of the variables were found to be significant: the psychological or emotional impact on the victim; the degree of injury (expressed as two dummy variables: minor vs. no or unknown injury; and moderate or serious vs. no or unknown injury); and the degree of violence (expressed as two dummy variables: moderate vs. minor or unknown violence; and serious vs. minor or unknown violence).

The additional variables made a substantial contribution to the overall model ($\Delta R^2 = .098$, $p < .001$). Parameter estimates for the variables are set out in Table H4.

Table H4: Parameter estimates for the statistically significant sentence note variables added to the sentence length model for male assaults female ($n = 415$)

	B	S.E.	t	Sig.
Victim suffered psychological or emotional impact	1.817	.592	3.067	.002
Minor injury cf No or unknown injury	.307	.460	.669	.504
Moderate or serious injury cf No or unknown injury	1.697	.693	2.449	.015
Moderate cf Minor or unknown violence	1.577	.519	3.038	.003
Serious cf Minor or unknown violence	2.581	.529	4.882	.000

Inclusion of the variables would substantially improve the main model. The three variables are all intuitively significant, the level of violence and the physical or psychological harm resulting from it would likely be identified by most people as important in determining sentence length. The significance of the omission of these factors for this research is dependent on the degree to which these factors are unevenly distributed across the circuits.

An indication of the prospect that the inclusion of these variables would affect the assessment of variation between circuits can be gained by comparing the co-efficients for the circuit variables across three models. The three models are: Model 1 (the main model reported in Chapter 8); Model 2 (the same model structure but restricted to the cases for which sentence note variables were available; and, Model 3 (Model 2 augmented by the significant sentence note variables). A comparison across the three models is necessary because there are two changes at work. The first is the change in the mix of cases due to the unavailability of the sentence notes (the difference between Models 1 & 2 is an indication of the effect of this change) and the second is the addition of the sentence note variables (the difference between Models 2 & 3 is an indication of the effect of this change). The co-efficients are reported in Table H5.

Table H5: Comparison of the regression co-efficients for the circuit variables for models including and excluding sentence note variables for male assaults female

	Model 1		Model 2		Model 3	
	(n = 722)		(n = 415)		(n = 415)	
Variable	B	Sig.	B	Sig.	B	Sig.
Circuit 1 v Circuit 0	3.514	.000	2.923	.002	2.408	.007
Circuit 2 v Circuit 0	1.119	.009	1.337	.023	1.085	.049
Circuit 3 v Circuit 0	.834	.000	.776	.010	.765	.007
Circuit 4 v Circuit 0	.308	.262	.079	.813	.029	.926
Circuit 5 v Circuit 0	.346	.011	.610	.002	.542	.004
Circuit 6 v Circuit 0	-.044	.732	-.059	.717	-.104	.502
Circuit 7 v Circuit 0	.580	.000	.593	.002	.472	.008
Circuit 8 v Circuit 0	.120	.350	-.123	.497	-.135	.428
Circuit 9 v Circuit 0	.027	.867	.370	.139	.183	.443
Circuit 10 v Circuit 0	.118	.129	.069	.495	-.012	.898
Circuit 11 v Circuit 0	.048	.471	.063	.445	.050	.514
Circuit 12 v Circuit 0	-.034	.534	-.046	.514	-.055	.409
Circuit 13 v Circuit 0	.091	.325	.111	.327	.088	.410
Circuit 14 v Circuit 0	-.014	.859	-.091	.468	-.052	.657
Circuit 15 v Circuit 0	.061	.259	.086	.344	.040	.644
Circuit 16 v Circuit 0	.115	.030	.139	.175	.136	.167

Comparison of the co-efficients indicates that the additional variables may reduce the variation between circuits but not substantially. Five of the six circuits (1, 2, 3, 5, & 7) that were significantly different remain significantly different but with marginally reduced parameter estimates between Models 2 & 3. The other circuit (16) is not significantly different but this can be attributed to the change in mix and reduction in the number of observations. Overall the results suggest it is unlikely that the additional variables are significantly unevenly distributed across the circuits and therefore it is not likely that addition of the variables would substantially impact on the estimates of variation between the circuits.

Burglary

Sentence note variables were available for 1,471 of the 2,338 cases included in the administrative dataset. However, 85 were excluded because the sentence note revealed the offender was a sentenced prisoner at the time of sentencing (65 cases); or the sentencing described a re-sentencing situation (15 cases) and did not cover the original sentence considered to be warranted for the offending; the sentence note was incomplete (4 cases);

or, the offender was found to be mentally impaired (1 case) leaving a final combined dataset of 1,386 cases. An additional 21 variables (see Table H6) were available.

Table H6: Descriptive statistics for the sentence length decision sentencing variables sourced from sentence notes for burglary (*n* = 1,386)

Categorical Variables	N	%
Value of property taken or damage caused		
Nothing or not mentioned	434	31.3%
\$1,000 or less	234	16.9%
More than \$1,000 but less than \$10,001	463	33.4%
More than \$10,000	255	18.4%
Damage done		
None or none mentioned	801	57.8%
Minimum necessary to effect the offence	495	35.7%
Property trashed	90	6.5%
Level of planning		
Opportunistic or not mentioned	412	29.7%
Some degree of planning	974	70.3%
Residential premise		
Non-residential or not mentioned	503	36.3%
Residential	883	63.7%
Premises occupied, occupants confronted		
No or not mentioned	1146	82.7%
Occupants present	189	13.6%
Confrontation occurred	51	3.7%
Offending occurred while on bail		
No or not mentioned	1222	88.2%
Yes	164	11.8%
Offending occurred while subject to an earlier sentence		
No or not mentioned	1080	77.9%
Yes	306	22.1%
No or not mentioned	1182	85.3%
Yes	204	14.7%
Breach of trust		
No or not mentioned	1280	92.4%
Yes	106	7.6%
Other aggravating factors present		
No or not mentioned	1275	92.0%
Yes	111	8.0%
Offender co-operated with the authorities		
No or not mentioned	1258	90.8%
Yes	128	9.2%

Categorical Variables	N	%
Assist in recovering property or make amends		
No or not mentioned	1133	81.7%
Yes	253	18.3%
Positive actions post offending		
No or not mentioned	1240	89.5%
Yes	146	10.5%
Remorse		
No or not mentioned	1264	91.2%
Yes	122	8.8%
Other mitigation		
No or not mentioned	1303	94.0%
Yes	83	6.0%
Poor response to previous sentences		
No or not mentioned	1133	81.7%
Yes	253	18.3%
Granted leave to apply to substitute home detention		
No or not mentioned	1336	96.4%
Yes	50	3.6%
Risk of re-offending		
Low or not mentioned	1078	77.8%
Moderate or high	308	22.2%
Reference to personal circumstances		
No or not mentioned	1169	84.3%
Yes	217	15.7%
Offended under the influence of alcohol or drugs		
No or not mentioned	1224	88.3%
Yes	162	11.7%
Offender had alcohol or drug problem		
No or not mentioned	871	62.8%
Yes	515	37.2%

To determine if any of these variables were significant an augmented model including an additional block of sentence note seriousness variables was run. Two of the variables were found to be significant: an estimate of the value of the property taken and damage caused and the degree of planning associated with the burglary offending. The insignificant aggravating and mitigating factors were transformed into two new variables by summing the number of aggravating and mitigating factors present respectively. The count of aggravating factors was found to be significant.

The additional variables made a moderate contribution to the overall model ($\Delta R^2 = .047$, $p < .001$). Parameter estimates for the variables are set out in Table H7.

Table H7: Parameter estimates for the statistically significant sentence note variables added to the sentence length model for burglary (*n* = 1,386)

	B	SE	t	Sig
Estimated value of property taken/damage caused	.007	.001	5.066	.000
Degree of planning	2.182	.492	4.433	.000
Number of aggravating factors excluding value and planning	.922	.140	6.576	.000
Number of mitigating factors	-.249	.233	-1.069	.285

Inclusion of the variables would substantially improve the main model. The two offence specific variables are intuitively significant. The estimated value of property taken or damage caused and the degree of planning would likely be identified by most people as important in determining sentence length. The composite generic variables for the number of aggravating and mitigating factors would assist to differentiate the cases in terms of relative seriousness reflecting variations in the offending and offenders. The significance of the omission of these factors for this research is dependent on the degree to which these factors are unevenly distributed across the circuits.

An indication of the prospect that the inclusion of these variables would affect the assessment of variation between circuits can be gained by repeating the approach used for male assaults female and comparing the co-efficients for the circuit variables across three models. The three models are: Model 1 (the main model reported in Chapter 7); Model 2 (the same model structure but restricted to the cases for which sentence note variables were available; and, Model 3 (Model 2 augmented by the significant sentence note variables). The co-efficients are reported in Table H8.

Three of the circuits (4, 7 & 13) that were previously significantly different are no longer significantly different. Comparison of Models 2 & 3 for these circuits indicates that the additional variables could have had some influence but the reduction in case numbers and change in the mix of cases were also influential. It is likely that the additional variables explain a portion of the difference. The remaining eleven circuits that were significantly different remain significantly different and the changes in co-efficient between Models 2 & 3 are marginal with no pattern in terms of increase/decrease.

Table H8: Comparison of the regression co-efficients for the circuit variables for models including and excluding sentence note variables for burglary

Variable	Model 1		Model 2		Model 3	
	(<i>n</i> = 2,338)		(<i>n</i> = 1,386)		(<i>n</i> = 1,386)	
	B	Sig.	B	Sig.	B	Sig.
Circuit 1 v Circuit 0	6.315	.000	6.314	.000	6.172	.000
Circuit 2 v Circuit 0	5.038	.000	5.293	.000	4.989	.000
Circuit 3 v Circuit 0	4.197	.000	4.432	.000	4.439	.000
Circuit 4 v Circuit 0	5.132	.000	4.480	.004	2.851	.055
Circuit 5 v Circuit 0	4.768	.000	3.189	.003	4.212	.000
Circuit 6 v Circuit 0	5.146	.000	5.153	.000	5.162	.000
Circuit 7 v Circuit 0	3.594	.004	2.968	.044	2.757	.052
Circuit 8 v Circuit 0	3.118	.005	3.377	.011	3.654	.004
Circuit 9 v Circuit 0	-1.803	.183	-1.360	.570	-.930	.688
Circuit 10 v Circuit 0	4.691	.000	3.981	.000	3.308	.002
Circuit 11 v Circuit 0	5.788	.000	5.467	.000	5.524	.000
Circuit 12 v Circuit 0	5.535	.000	4.149	.000	4.392	.000
Circuit 13 v Circuit 0	-2.663	.024	-4.152	.009	-2.345	.130
Circuit 14 v Circuit 0	5.841	.000	6.226	.000	5.711	.000
Circuit 15 v Circuit 0	.056	.949	-.467	.669	-.404	.702
Circuit 16 v Circuit 0	4.553	.000	5.099	.000	4.629	.001

On balance while the additional variables would likely improve the performance of the model they would not significantly change the impact of circuit on sentencing.

References

- Albonetti, C.A. (1991). An integration of theories to explain judicial discretion. *Social Problems*, 38, 247-266.
- 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 & Society Review*, 31, 789-822.
- Albonetti, C.A. & Baller, R.D. (2010). Sentencing in federal drug trafficking/manufacturing cases: A multi-level analysis of extra-legal defendant characteristics, guidelines departures, and continuity of culture. *Journal of Gender, Race and Justice*, 14, 41-72.
- Ashworth, A. (1998). Four techniques for reducing sentence disparity. In A. von Hirsch & A. Ashworth (Eds). *Principled sentencing: Readings on theory and policy*. Oxford: Hart Publishing.
- Austin, T.L. (1981a). The influence of court location on type of criminal sentence: The rural-urban factor. *Journal of Criminal Justice*, 9, 305-316.
- Austin, T.L. (1981b). Criminal sentencing in urban, suburban, and rural counties. *Criminal Justice Review*, 6, 31-37.
- Berk, R.A. (1983). An introduction to sample selection bias in sociological data. *American Sociological Review*, 48, 386-398.
- Blackwell, B.S., Holleran, D. & Finn M.A. (2008). The impact of the Pennsylvania sentencing guidelines on sex differences in sentencing. *Journal of Contemporary Criminal Justice*, 24, 399-418.
- Bond, C. & Jeffries, S. (2011). Indigeneity and the judicial decision to imprison. *British Journal of Criminology*, 51, 256-277.
- Bond C., Jeffries, S. & Weatherburn, D. (2011). How much time? Indigenous status and the sentenced imprisonment term decision in New South Wales. *Australian and New Zealand Journal of Criminology*, 44, 272-290.
- Britt, C.L. (2000). Social context and racial disparities in sentencing decisions. *Justice Quarterly*, 17, 707-732.
- Brookbanks, W. (2007). The criminal justice process. In J. Tolmie & W. Brookbanks (Eds), *Criminal Justice in New Zealand*. Wellington, New Zealand: LexisNexis.
- Bushway, S.D., Johnson, B.D. & Slocum, L.A. (2007). Is the magic still there? The use of the Heckman two-step correction for selection bias in criminology. *Journal of Quantitative Criminology*, 23, 151-178.
- Bushway, S.D., & Piehl, A.M. (2001). Judging judicial discretion: Legal factors and racial discrimination in sentencing. *Law & Society Review*, 35, 733-764.

- Chatterjee, S. & Hadi, A. S. (2006). *Regression analysis by example*. Hoboken, New Jersey: John Wiley & Sons.
- Chhana, R., Spier, P., Roberts, S., and Hurd, C. (2004). *The Sentencing Act 2002: Monitoring the first year*. Wellington, New Zealand: Ministry of Justice.
- Criminal Procedure Act 2011
(<http://legislation.govt.nz/act/public/2011/0081/latest/DLM3359962.html>)
- Crow, M.S., & Gertz, M. (2008). Sentencing policy and disparity: Guidelines and the influence of legal and democratic subcultures. *Journal of Criminal Justice*, 36, 362-371.
- Dixon, J. (1995). The organizational context of criminal sentencing. *American Journal of Sociology*, 100, 1157-1198.
- Doerner, J.K. & Demuth, S. (2010). The independent and joint effects of race/ethnicity, gender, and age on sentencing outcomes in U.S. federal courts. *Justice Quarterly*, 27, 1-27.
- Doob, A.N. & Sprott, J.B. (2007). The sentencing of aboriginal and non-aboriginal youth: Understanding local variation. *Canadian Journal of Criminology and Criminal Justice*, 45, 109-123.
- Eisenstein, J., Flemming, R.B., & Nardulli, P.F. (1988). *The contours of justice*. Boston: Little, Brown & Co.
- Eisenstein, J., & Jacob, H. (1977). *Felony justice: An organisational analysis of criminal courts*. Boston: Little, Brown & Co.
- Englich, B., & Mussweiler, T. (2001). Sentencing under uncertainty: Anchoring effects in the courtroom. *Journal of Applied Social Psychology*, 31, 1535-1551.
- Englich, B., Mussweiler, T., & Strack, F. (2005). The last word in court - A hidden disadvantage for the defense. *Law and Human Behavior*, 29, 705-722.
- Englich, B., Mussweiler, T., & Strack, F. (2006). Playing dice with criminal sentences: The influence of irrelevant anchors on experts' judicial decision making. *Personality and Social Psychology Bulletin*, 32, 188-200.
- Everett, R.S., & Wotjkieicz, R. (2002). Difference, disparity, and race/ethnic bias in federal sentencing. *Journal of Quantitative Criminology*, 18, 189-211.
- Farrell, A., Ward, G. & Rousseau, D. (2009). Does racial balance in workforce representation yield criminal justice? Race relations of sentencing in federal court organisations. *Law and Society Review*, 43, 757-805.
- Fearn, N.E. (2003). Community context and sentencing decisions: A multi-level analysis. Unpublished doctoral dissertation, University of Missouri – St Louis, US.

- Fearn, N.E. (2005). A multi-level analysis of community effects on criminal sentencing. *Justice Quarterly*, 22, 452-487.
- Field, A. (2009). *Discovering Statistics Using SPSS*. London, Sage Publications.
- Gigerenzer, G. (2006). Heuristics. In Gigerenzer, G. & Engel, C., *Heuristics and the law*. Cambridge, Massachusetts: The MIT Press.
- Gigerenzer, G. & Goldstein, D.G. (1999). Betting on one good reason: The take the best heuristic. In Gigerenzer, G. Todd, P.M and the ABC Research Group, *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.
- Gigerenzer, G. & Todd, P.M. (1999). Fast and Frugal Heuristics: The Adaptive Toolbox. In Gigerenzer, G. Todd, P.M and the ABC Research Group, *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.
- Goodall, A.W. & Durrant, R. Regional variation in sentencing: The incarceration of aggravated drink drivers in the New Zealand District Courts. *Australian and New Zealand Journal of Criminology*, 46(3), 422-447.
- Hall, G. (1991). Reducing disparity by judicial self-regulation: Sentencing factors and guideline judgements. *New Zealand Universities Law Review*, 14, 208-228.
- Hall, G. (2007). Sentencing. In J. Tolmie & W. Brookbanks (Eds), *Criminal justice in New Zealand*. Wellington, New Zealand: LexisNexis.
- Hall, G.G. (2008). *Hall's Sentencing*. Wellington, New Zealand: LexisNexis.
- Hessell v R [2010] NZSC 135; [2011] 1 NZLR 607.
- Hofer, P.J., Blackwell, K.R., & Ruback, R.B. (1999). The effect of the federal sentencing guidelines on inter-judge sentencing disparity. *Criminology*, 90, 239-306.
- Hupfeld-Heinemann, J. & von Helversen, B. (2009). Models of decision making on guilt and sanctions. In M. E. Oswald, S. Bieneck & J. Hupfeld-Heinemann, *Social psychology of punishment of crime*. New York: J Wiley & Sons
- Iles, G., Bumphus, V.W. & Zehel, L. (2011). Assessing the relevance of legal and extralegal factors in felony drug sentencing: The case of Michigan. *Criminal Justice Studies*, 24, 291-306.
- Imprisonment rates by court region for 2004-2006. (2007). Unpublished paper. Law Commission, Wellington, New Zealand.
- Jeffries, S. & Bond, C. (2009). Does indigeneity matter? Sentencing indigenous offenders in South Australia's higher courts. *Australian and New Zealand Journal of Criminology*, 42, 47-71.
- Jeffries, S. & Bond, C. (2012). The impact of indigenous status on adult sentencing: A review of the statistical research literature from the United States, Canada, and Australia. *Journal of Ethnicity and Criminal Justice*, 10, 223-243.

- Jeffries, S., Fletcher, G.J.O. & Newbold, G. (2003). Pathways to sex-based differentiation in criminal court sentencing. *Criminology*, 41, 329-354.
- Johnson, B.D. (2005). Contextual disparities in guidelines departures: courtroom social contexts, guidelines compliance, and extralegal disparities in criminal sentencing. *Criminology*, 43, 761-797.
- Johnson, B.D. (2006). The multi-level context of criminal sentencing: Integrating judge and county-level influences. *Criminology*, 44, 259-298.
- Johnson, B.D. (2011). Sentencing. In M. Tonry (Ed), *The Oxford handbook of crime and criminal justice*. Oxford: Oxford University Press.
- Johnson, B.D., Ulmer, J.T., & Kramer, J.H. (2008). The social context of guidelines circumvention: The case of federal district courts. *American Society of Criminology*, 46, 737-783.
- Joseph, P.A. (2007). *Constitutional and administrative law in New Zealand*. Wellington, New Zealand: Brookers.
- Kautt, P.M. (2002). Location, location, location: Interdistrict and intercourt variation in sentencing outcomes for federal drug-trafficking offenses. *Justice Quarterly*, 19, 633-671.
- Kautt, P.M. & Delone, M.A. (2006). Sentencing outcomes under competing but coexisting sentencing interventions: Untying the Gordian knot. *Criminal Justice Review*, 31, 105-131.
- Kautt, P.M. & Mueller-Johnson, K. (2009). Cross-jurisdictional disposition variability under sentencing guidelines: The example of equivalent sex offenses. *Criminal Justice Review*, 34, 350-381.
- Kautt, P.M., & Spohn, C. C. (2007). Assessing blameworthiness and assigning punishment: Theoretical perspectives on judicial decision making. In D. E. Duffee & E. R. Maguire (Eds), *Criminal justice theory: Explaining the nature and behaviour of criminal justice*. New York: Routledge.
- Kramer, J.H., & Ulmer, J.T. (2002). Sentencing disparity and departures from guidelines. *Justice Quarterly*, 13, 401-425.
- Latimer, J. & Foss, L.C. (2005). The sentencing of aboriginal and non-aboriginal youth under the Young Offenders Act: A multivariate analysis. *Canadian Journal of Criminology and Criminal Justice*, 45, 481-499.
- Law Commission (2002). *Seeking solutions: Options for change to the New Zealand court system*. Wellington, New Zealand: Law Commission.
- Law Commission (2006). *Sentencing guidelines and parole reform*. Wellington, New Zealand: Law Commission.

- Localio, A.R., Margolis, D.J., & Berlin, J.A. (2007). Relative risks and confidence intervals were easily computed indirectly from multivariable logistic regression. *Journal of Clinical Epidemiology*, 60, 874-882.
- Mason, T., de Silva, N., Sharma, N., Brown, D., & Harper, G. (2007). *Local variation in sentencing in England and Wales*. Ministry of Justice: London.
- Mitchell, O. & Mackenzie D.L. (2004). *Relationship between race, ethnicity, and sentencing outcomes: A meta-analysis of sentencing research*. Unpublished Report prepared for the U.S. Department of Justice.
- Mitchell, O. (2005). A meta-analysis of race and sentencing research: Explaining the inconsistencies. *Journal of Quantitative Criminology*, 21, 439-466.
- Ministry of Justice (2010). *The District Courts*. Retrieved October 26, 2011, from http://www.courtsofnz.govt.nz/district/district/1558_DC_Courts_of_NZ_FINALpdf.pdf/
- Myers, M.A. & Talarico, S.M. (1987). *The social contexts of criminal sentencing*. New York: Springer-Verlag.
- New Zealand Parliament (2006). *New Zealand Parliamentary Debates Volume 636*. Wellington, New Zealand: Government Printer.
- New Zealand Parliament (2007). *New Zealand Parliamentary Debates Volume 640*. Wellington, New Zealand: Government Printer.
- Ostrom, C. W., Ostrom, B. J., & Kleiman, M. (2004). *Judges and Discrimination: Assessing the Theory and Practice of Criminal Sentencing*. Unpublished Report prepared for the U.S. Department of Justice.
- Ostrom, C.W., Ostrom, B.J., Hanson, R.A., & Kleiman, M. (2008). *Assessing Consistency and Fairness in Sentencing: A Comparative Study in Three States*. Unpublished report prepared for the U.S. Department of Justice.
- Pasko, L. (2002). Villain or victim: Regional variation and ethnic disparity in federal drug offense sentencing. *Criminal Justice Policy Review*, 13, 307-328.
- Priestley, J.M. HCJ (2009). Chipping away at the judicial arm? Harkness Henry Lecture. Available at: <http://www.courtsofnz.govt.nz/from/speeches-and-papers>
- R v Barrett [1999] 1 NZLR 146.
- R v Hessel [2009] NZCA 450; [2010] 2 NZLR 298, (2009) 24 CRNZ 612.
- R v Elliot [1981] 1 NZLR 295.
- R v Taueki [2005] 3 NZLR 372.
- R v Ward [1976] 1 NZLR 588.

- Raudenbush, S. W. & Bryk, A.S. (2002). *Hierarchical linear models: applications and data analysis methods (2nd ed)*. Thousand Oaks, California: Sage Publications.
- Robertson, J.B. (Ed). (2007). *Adams on criminal law – sentencing*. Wellington, New Zealand: Brookers.
- S v Police (1989) 5 CRNZ 8.
- Sankoff, P. (2007). Constituents in the trial process. In J. Tolmie & W. Brookbanks (Eds), *Criminal justice in New Zealand*. Wellington, New Zealand: LexisNexis.
- Savelsberg, J.J. (1992). Law that does not fit society: Sentencing guidelines as a neoclassical reaction to the dilemmas of substantivized law. *American Journal of Sociology*, 97, 1346-1381.
- Sentencing Act 2002
(<http://www.legislation/act/public/2002/0009/latest/DLM135342.html>)
- Sentencing Council Act 2007
(<http://legislation.govt.nz/act/public/2007/0025/latest/DLM410029.html>)
- Sentencing Establishment Unit (2008). Summarised numbers and percentages on the type of sentence by guideline. Unpublished paper. (Law Commission), Wellington, New Zealand.
- Sharp, BA. (2007). *A longitudinal study of the US sentencing guidelines: a decade of balancing judicial discretion and unwarranted disparity (1993-2003)*. Unpublished doctoral dissertation, University of Maryland, US.
- Simon, H. A. (1976). *Administrative Behaviour* (3rd Ed). New York: The Free Press
- Simon, H.A. (1992). *Bounded Rationality and the Cognitive Revolution*. Aldershot, United Kingdom: Edward Elgar.
- Snowball, L. & Weatherburn, D. (2006). Indigenous over-representation in prison: The role of offender characteristics. *Contemporary Issues in Crime and Justice*, 99. Sydney: New South Wales Bureau of Crime Statistics and Research.
- Snowball, L. & Weatherburn, D. (2007). Does racial bias in sentencing contribute to indigenous over representation in prison? *Australian and New Zealand Journal of Criminology*, 40, 272-290.
- Soboleva, N., Kazakova, N., & Chong, J. (2006). *Conviction and sentencing of offenders in New Zealand: 1996 to 2005*. Wellington, New Zealand: Ministry of Justice.
- Spier, P. (1989). *An examination of regional differences in the use of custodial sentences in the District Court*. Unpublished Report prepared for the Department of Justice, Wellington.
- Spohn, C. (2000). *Offender race and case outcomes: Do crime seriousness and strength of evidence matter?* Unpublished Report prepared for U.S. Department of Justice.

- Steffensmeier, D., & Demuth, S. (2001). Ethnicity and judges' sentencing decisions: Hispanic-Black-White comparisons. *Criminology*, 39, 145-179.
- Steffensmeier, D., & Demuth, S. (2006). Does gender modify the effects of race-ethnicity on criminal sanctions? Sentences for male, female, white, black, and hispanic defendants. *Journal of Quantitative Criminology*, 22, 241-261.
- Steffensmeier, D., Ulmer, J., & Kramer, J. (1998). The interaction of race, gender and age in criminal sentencing: The punishment cost of being young, black and male. *Criminology*, 36, 763-798.
- Stolzenberg, R.M., & Relles, A.R. (1997). Tools for intuition about sample selection bias and its correction. *American Sociological Review*, 62, 494-507.
- Supreme Court Act 2003
(<http://legislation.govt.nz/act/public/2003/0053/latest/DLM214028.html>)
- Tarling, R. (1979). *Sentencing Practice in Magistrates' Courts*. London: Home Office Research Study No. 56, Her Majesty's Stationery Office.
- Tarling, R. (1979). Sentencing Practice in Magistrates' Courts Revisited. *Howard Journal*, 45, 29-41.
- Triggs, S. (1999). *Sentencing in New Zealand: a statistical analysis*. Wellington: Ministry of Justice.
- Ulmer, J. T. (2005). The localized uses of federal sentencing guidelines in four U.S. District Courts: Evidence of processual order. *Symbolic Interaction*, 28, 255-279.
- Ulmer, J.T. (2012). Recent developments and new directions in sentencing research. *Justice Quarterly*, 29, 1-40.
- Ulmer, J. T., & Bradley, M.S. (2006). Variation in trial penalties among serious violent offenses. *Criminology*, 44, 631-670.
- Ulmer, J.T., Eisenstein, J., & Johnson, B.D. (2009). Trial penalties in federal sentencing: Extra-guidelines factors and district variation. *Justice Quarterly*, 20, 1-33.
- Ulmer, J. T., & Johnson, B. (2004). Sentencing in context: A multi-level analysis. *Criminology*, 42, 137-177.
- Ulmer, J. T. & Kramer, J. H. (1996). Court communities under sentencing guidelines: Dilemmas of formal rationality and sentencing disparity. *Criminology*, 34, 383-407.
- Ulmer, J. T., & Light, M. T. (2010). The stability of case processing and sentencing post-Booker. *The Journal of Gender, Race & Justice*, 14, 143-177.
- von Helversen, B., & Reiskamp, J. (2009). Predicting sentencing for low-level crimes: Comparing models of human judgment. *Journal of Experimental Psychology: Applied*, 15, 375-395

- Wasik, M. & von Hirsch, A. (1990). Statutory sentencing principles: The 1990 White Paper. *Modern Law Review*, 53, (508-517)
- Weidner, R. R., & Frase, R. S. (2003). Legal and extralegal determinants of intercounty differences in prison use. *Criminal Justice Policy Review*, 14, 377-400.
- Weidner, R. R., Frase, R., & Schultz, J. S. (2005). The impact of contextual factors on the decision to imprison in large urban jurisdictions: A multi-level analysis. *Crime & Delinquency*, 51, 400-424.
- Weinrath, M. (2007). Sentencing disparity: Aboriginal Canadians, drunk driving and age. *Western Criminology Review*, 8, 16-28.
- Welsh, A. & Ogloff, J.R.P. (2008). Progressive reforms or maintaining the status quo? An empirical evaluation of the judicial consideration of aboriginal status in sentencing decisions. *Canadian Journal of Criminology and Criminal Justice*, 48: 491-517.
- Wheeler, S., Weisburd, D., & Bode, N. (1982). Sentencing the white-collar offender: Rhetoric and reality. *American Sociological Review*, 47, 641-659.
- Wiseman, A., & Connelly, M. (2008). *Judicial Discretion and Sentencing Outcomes: Incorporating Data from the Courtroom*. Unpublished report prepared for the U.S. Department of Justice.
- Zatz, M.S. (1987). The changing forms of racial ethnic biases in sentencing. *Journal of research in crime and delinquency*, 24, 69-92.
- Zatz, M.S. (2000). The Convergence of race, ethnicity, gender and class on court decisionmaking: Looking toward the 21st century. In *Policies, processes and decisions of the criminal justice system. Volume 3, Criminal Justice 2000*. Washington, DC: United States Department of Justice.