*Asia Pacific Law and Economics Review 2011 Vol 1, pp 5-34*

# An Empirical Analysis of Health and Safety in Employment Sentencing in New Zealand

**by**

**Andrea Menclova and Alan Woodfield**

**Abstract:** Apparent inconsistency in criminal sentencing at District Court level in New Zealand (NZ) might also be expected for health and safety in employment (HSE) offences. We review relevant legislation and the guidelines established in the *de Spa* appeal case, and estimate a model of HSE sentencing variability distinguishing the *de Spa* criteria (and a subset similar to those used in the formal U.S. criminal sentencing guidelines) from a more comprehensive list of sentencing factors routinely used. When the *de Spa* case-mix variables are controlled for, a weak increase in inter-district sentencing variability is observed but with a reduction in intra-district variability, while both inter and intra-judge variability is mitigated. We show that a number of the *de Spa* (and other) criteria are significant determinants of sentencing variation, although some results (e.g., for the presence of remorse) are puzzling. The results seem quite robust to the choice between a dataset including the common s 6 offences only and a dataset of cases as a whole as well as to several other sensitivity checks. We also show that the model retrospectively predicts the sentence in the *de Spa* appeal case well, and suggest how the model might be used as a basis for more consistent future sentencing decisions.

**Keywords:** Health & Safety Offences, Judicial Guidelines, Sentencing Determinants.

**JEL Classification:** K32

**Acknowledgements:** We pay tribute to our research assistant, Grace Lee, for her efforts (well above and beyond the call of duty) in constructing the major components of the database used in this study. We thank the Department of Labour for providing primary data on HSE prosecutions and access to many judicial decisions, sentencing notes, and returns on prosecutions. Particular thanks to Tim Beverstock, Mike Hargreaves, Ross Hill, George Mason, and Rex Moir for their generous ongoing assistance in clarifying interpretations of the data and locating some elusive elements of it. We thank participants in the Law and Economics Association of New Zealand sessions of the 2009 New Zealand Association of Economists conference for their helpful comments on an earlier version of this paper, and the College of Business and Economics at the University of Canterbury for financial support from their Contestable Research Fund. Any errors, naturally, are ours.

**Authors' Affiliation:** Department of Economics and Finance, College of Business and Economics, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand.

**Corresponding Author:** Alan Woodfield, email: **alan.woodfield@canterbury.ac.nz,** phone +64 3 351 5893, fax +64 3 364 2635.

## 1 Introduction

There is considerable ongoing interest and controversy surrounding the consistency of judicial sentencing in New Zealand (NZ). In February 2006, the (then) NZ Labour Government invited the NZ Law Commission to investigate reforms to the sentencing and parole systems for criminal offences, leading to the publication of its report NZLC R94 in August 2006.[1] Evidence suggestive of sentencing inconsistency among Court Districts was influential in the Commission's argument that a core problem was the highly discretionary nature of current sentencing and parole arrangements. Parliament subsequently introduced the Sentencing Council Act 2007 with a view (*inter alia*) of promoting consistency in sentencing practice among different courts and judges. By July 2008, the Sentencing Establishment Unit (SEU) within the Law Commission had drafted seventeen generic guidelines and forty-two offence-based guidelines for criminal sentencing policy (although they have no official status). Following the change in government in November 2008, however, the Sentencing Council has not been established and formal criminal sentencing guidelines have not been adopted.

Criminal sentencing guidelines have previously been adopted in a number of other countries, examples of which include the United States (U.S.), England, Canada, and some states of Australia. For example, the empirically-based U.S. Federal Sentencing Guidelines introduced in 1987 imposed much tighter limits on judicial discretion. Previously, sentencing had been indeterminate with broad punishment ranges specified. The U.S. Guidelines instead spelled out tight sentencing ranges for offences, largely depending on seriousness and a defendant's criminal history.

Apparent inconsistency in criminal sentencing in NZ is likely to remain important in public debate. Further, the issue of sentencing consistency seems to be as pertinent to prosecutions for offences under health and safety legislation as for criminal offences. New Zealand's health and safety 'guidelines' contain two distinct components. The first, introduced by the Health and Safety in Employment Act 1992 (HSE Act) implemented on 1 April, 1993, is a specified range (with a lower bound of zero) of financial penalties for employers convicted of offences under the Act. Notably, the upper bound was increased significantly in the HSE Amendment Act 2002, and account of new sentencing principles specified in the Sentencing Act 2002 was also required. The second component was provided in the guideline judgment in the High Court (HC) case *Department of Labour v de Spa and Co Ltd*. [1994] 1 ERNZ 339 involving a successful appeal against the level of sentence imposed in the Christchurch District Court.[2] Nine relevant sentencing factors were specified. Notably, the *de Spa* Guidelines were adopted in the Sentencing Act with only minor changes.

In this paper, we examine empirically the NZ Courts' sentencing criteria and the associated financial liability for employers convicted of offences under the HSE Act. We focus on Section 6 offences that are by far the most common. We also examine the aggregation of sentences to the case level in order to be able to

[1] Cf. New Zealand Law Commission (2006).
[2] Cf. *Department of Labour v de Spa and Co Ltd*. (DC Christchurch, CRN 30090213/93, 8 October 1993, (Holderness, J).

investigate all multiple-charge and/or multiple-victim cases. To our knowledge, this is the first empirical study of the determinants of health and safety sentencing.

The remainder of the paper is organized as follows: In section 2, we outline the development of sentencing guidelines for offences under New Zealand's health and safety legislation while in section 3, we discuss these developments in the context of the nascent economic theory of sentencing and of empirical assessments of the impact of the stringent U.S. Guidelines for criminal offences. Further, we examine some empirical claims of sentencing inconsistency for criminal offences in NZ. In section 4, we describe the construction of the database used in this study. Section 5 discusses the analytical methods used, while section 6 presents our empirical results. The paper reaches conclusions and implications for sentencing policy in section 7.

## 2 The Development of New Zealand's Health and Safety Guidelines

In the first decade of the application of the HSE Act 1992, the maximum fine that could be imposed was $100,000 if the offender knew the relevant act or omission relevant to the offence was reasonably likely to cause serious harm (s 49) or $50,000 for other offences involving non-compliance with various provisions of the Act (s 50).[3] Subsequently, a strong signal to the Courts to significantly increase penalties was provided by a fivefold increase in maximum penalties in the amendments of 2002 which were implemented from 5 May 2003. Section 51(A) of the amended Act, however, also required the Courts to pay particular regard to sections 7–10 of the Sentencing Act 2002 (implemented from 30 June 2002) dealing with the purposes and principles of sentencing, and to the requirements of s 35 and s 40.[4]

The Sentencing Act repealed s 28 of the Criminal Justice Act 1985 which had permitted the payment of all or part of any fine to a victim of an offence. A new sentence of reparation that could be awarded when a victim suffered loss or damage to property, emotional harm, or loss or damage consequential on any emotional or physical harm, or loss or damage to property, was introduced.[5] Courts were generally required to impose a sentence of reparation where applicable, and reparation could also be imposed in conjunction with other sentences. Unlike sentences for HSE Act offences, reparation awards are not capped.[14] If an offender can meet reparation but not fine payments, reparation takes precedence, while, under s 35, the Courts can discount losses or provide time to pay where the offender is impecunious. Further,

[3] All financial amounts are in NZ dollars (denoted by '$'). On average, one NZ dollar purchased a little less than 60 U.S. cents during the period studied (years ending June 1994-2007) and purchased 76 U.S. cents at the end of this period. For comparison, the maximum s 50 penalty of $50,000 was somewhat higher than the median annual household income of $41,500 in year 2000.

[4] Hall (2009 at para I.1.1) argued that maximum penalties are reserved for the most serious offences of their type and offer little guidance to sentencing judges, and also noted (at para I.1.4) that the statutory language of the Sentencing Act "is frequently expressed in such general terms that it does not place any substantial fetter upon the sentencing discretion." Hughes (2002, p. 125) noted that the new clause 51(A) in the amended HSE Act largely repeated certain sentencing criteria in the Sentencing Act (including references to the ability to pay reparation) and otherwise recited the established *de Spa* Guidelines.

[5] Awards of reparation were previously possible under s 22 of the Criminal Justice Act, but were rare in comparison to awards of fines to victims facing more general losses than those for which reparation could be awarded under the Sentencing Act.

under s 40, if reparation is ordered, the Court must take its magnitude into account in assessing the level of any fine imposed.[6]

In respect of *de Spa*, the Department of Labour successfully appealed the level of sentence. The defendant company was charged (*inter alia*) with failing to comply with s 6 of the HSE Act in that it failed to take all practicable steps to ensure the safety of its employees. The charge arose when an 18 year old employee was fatally crushed by a moving bar while working with a wool bale elevator machine. A fine of $6,500 was imposed. On appeal, this was considered to be manifestly inadequate and was raised to $15,000, representing 30 percent of the maximum fine which in turn was arguably a very small proportion of the loss.[7]

The HC judges then turned their attention to criteria deemed to be relevant in sentencing. They were stated not to be intended to be exhaustive. The *de Spa* Guidelines are summarised in the holding and are listed below:

1. The degree of culpability;
2. The degree of harm resulting;
3. The financial circumstances of the offender;
4. The attitude of the offender, including remorse, co-operation, and taking remedial action;
5. Any guilty plea;
6. The need for deterrence;
7. Compensation to the victim under s 28 Criminal Justice Act 1985;
8. The employer's safety record;
9. The facts of the particular case.

Hall (2009 at para I.2.2(c)) argued that little attempt is made in NZ guideline judgments to analyse, categorise and weigh the factors influencing sentence, with what is seen to be relevant sentencing factors merely being identified. In this respect, *de Spa* is typical. No indication of whether these factors are listed in any particular order of importance is given, nor whether different weights should be applied to the different criteria. Not even the signs of the effects of various criteria on penalties are provided, although these may be implicit in general NZ sentencing principles.[8]

The HC judges applied their sentencing criteria to the facts in *de Spa*. They considered the employer to have exhibited a medium level of culpability in that the hazard was readily foreseeable. The ultimate (i.e., fatal) level of harm resulted. The employer was not impecunious, and their attitude was favourable, with prompt

[6] Gordon and Woodfield (2006) argued that the HSE Act amendments appeared to be tempered by the application of the Sentencing Act. The Courts appeared initially reluctant to increase total financial liabilities much at all, and fines were substantially substituted by reparation awards (see also Mason (2008)). Although more recent evidence suggested that liability levels were increasing markedly, in most cases they remained well below the levels signaled by legislation.

[7] Access Economics (2006) reported a 'conservative' estimate for the value of a statistical life of $3.9 million in NZ for 2003 (the mean estimate being $6.9 m.), representing an estimated value per life year of $184,216 using a discount rate of 3.8 percent and a 40 year (remaining) life expectancy.

[8] According to Hall (2009 at para I.4.3), judges need to preserve "a reasonable degree of uniformity of penalty between offenders convicted of the same offence at or about the same level of culpability." Sentences are meant to reflect the gravity of an offence, in part signalled by the size of the maximum penalty but also reflecting acts of commission and omission relevant to the facts of the particular case.

remedial action taken. No guilty plea was entered, so no consideration was necessary on this account or on the issue of victim compensation. The defence of a complete absence of fault had not been established at DC level, and the defendant did not appeal the decision. Although particular deterrence was not seen to be required, the sentence needed to provide general deterrence. A good safety record without incident or complaint was evident over a 17 year period. On balance, the HC judges considered a sentence representing only 13 percent of the statutory maximum to be manifestly inadequate given that the maximum penalty was to cover the worst possible case, and they could not imagine such a case to be eight times worse than *de Spa*. The fine was therefore increased by 130 percent to $15,000. It was also noted that an initial fine of as much as $20,000 would have been unlikely to be overturned on appeal.

The *de Spa* Guidelines and the accompanying reasoning provided in the judgment became a major reference point for subsequent sentencing decisions, and were later incorporated in the Sentencing Act without significant change.

## 3 New Zealand's Health and Safety Guidelines in Context

In this section, we first discuss some recent developments in the economic theory of sentencing and then briefly review empirical evaluations of aspects of the U.S. Guidelines for criminal offences. We consider the properties of NZ's health and safety 'guidelines' in the context of these discussions. Finally, we briefly examine some empirical claims of sentencing inconsistency for criminal offences in NZ.

### 3.1 Economic Theory of Sentencing

Shavell (2007) and Miceli (2008) examined whether and when it is in society's interest to permit judicial discretion. They allow sentencing rules or guidelines to depend only on included (publicly observable) variables, so that granting discretion allows judges to make decisions reflecting the remaining unincluded variables that are unobservable prior to trial. Discretion is desirable if decisions should depend on unobservable relevant information that is only available *ex post* the rule setting. Discretion, however, can be harmful if judges use their acquired information to set sentences that reflect their own objectives rather than those of society.[9]

Judicial guidelines directly seek to restrict the scope of discretion. Considerable room for discretionary deviation may remain, and this may be beneficial since both Shavell and Miceli (among others) show that a positive scope for discretion is desirable under quite general conditions. Overall, however, scope restriction is considered a rough control device by Shavell since it has no effect on deviation within the allowed scope but precludes further discretion when further discretion is warranted.

Evidently, New Zealand exhibits considerable scope for judicial discretion. The lower bound of the permissible fine is zero which allows for convictions without

[9] Reinganum (1988, 2000) and Bar-Gill and Gazal Ayal (2006) also establish conditions under which guidelines can constrain prosecutorial discretion in plea-bargained cases.

any financial liability in particular circumstances. These, however, are rare in our database; for example, they amount to only 1.4 percent of cases that include a s 6 offence through mid-2007. Absent reparation, the zero lower bound may send a signal to judges that a minor penalty will suffice. This was especially likely where the maximum fine was initially set at the relatively low level of $50,000 for a single-charge breach of s 50 of the HSE Act. The 2002 amendments to the Act maintained the lower bound at zero and raised the maximum fine to $250,000.[10] Notably, this maximum has never been reached, and neither was its predecessor, in each case leaving a substantial margin for the 'worst possible case'. And while total financial liabilities have gradually increased, in some cases substantially so, there may be circumstances where a fine well in excess of zero is warranted (although a small penalty is imposed) and (different) circumstances where a fine well in excess of $250,000 would be warranted (although a penalty well less than $250,000 is imposed to leave room for the elusive 'worst possible case').

In Miceli's (2008) model, if the primary goal of punishment is deterrence, the familiar optimal risk-neutral sanction equals the harm imposed times the inverse of the probability of apprehension. This 'deterrence ideal' is independent of the characteristics of the offender, implying that a sentencing range is otiose. In the *de Spa* Guidelines, both the degree of harm and the degree of employer culpability are identified characteristics of offences to be considered at sentencing but there is no clear link to be followed between the characteristics and the sentence. In any case, it is difficult to credit that the current maximum penalty in NZ even begins to approach the harm arising from fatal or other very serious accidents in typical cases. Miceli (1991) and Polinsky and Shavell (2000), however, recognize that other values such as fairness may have a role in penalty-setting, requiring the characteristics of the offender as well as the crime to be taken into account in sentencing. Judicial decisions, however, are identified solely with fairness in Miceli's 2008 model. Given sentencing guidelines, judicial discretion is constrained by lower and upper bounds on the sanction which are required given that a legislature would find it too costly to anticipate and provide for all possible realizations of offender characteristics and could not condition the sanction on characteristics that are unobservable to them in any case. In the *de Spa* Guidelines, most of the characteristics relevant to sentencing convicted offenders only become known at the time of trial. The *de Spa* Guidelines, however, made no attempt to specify penalties that were functionally related to measures of these characteristics, allowing judges considerable room to deviate from what society would consider appropriate penalties, and also considerable room for inter-judicial variation in sentencing.

For Miceli, optimal sentencing guidelines afford judges some discretion in order to promote fairness, but the role of deterrence limits their discretion to a range that includes the deterrence ideal.[11] It is hard to credit that the NZ legislative

[10] Notably, the real value of the maximum penalty would have been maintained by an increase of $10,000 rather than the actual increase of $200,000.

[11] There is an alternative perspective that argues in favour of discretion to promote deterrence. One reason is that if offenders are risk-averse, greater sentencing variation with a fixed expected punishment will enhance deterrence. Miceli (2008) has reservations about this argument since risk-attitudes of offenders are not directly observable. Attitudes to risk, however, might be inferred from offenders' purchases of insurance in unrelated markets. Another argument is that discretion permits a form of price discrimination, allowing judges to impose harsher sentences on those more difficult to deter. The expected punishment could have included a weight reflecting the marginal cost of the

guidelines include this ideal except in cases of less than serious injuries involving relatively brief rehabilitation periods. Further, as the weight given to deterrence increases, the sentencing constraints tighten on both sides of the ideal. Yet when NZ legislation increased the maximum penalty five-fold, the cap was still far below the deterrence ideal for serious harms while the minimum penalty remained at zero. Soft sentencing could continue while harsh sentencing effectively remained difficult if not impossible for serious harm injuries. Also, offences that are more harmful socially increase the deterrence ideal so both bounds on sanctions must move to accommodate this. In NZ, for serious harms, the amended maximum penalty is now closer to this ideal, but still remains well below it. The lower bound was not adjusted.

If sentencing guidelines are rejected, an alternative approach is to permit the judge a decision-based payoff whereby judges setting penalties outside a specified interval would be penalized sufficiently to deter them from so doing. In NZ, however, District Court judges are paid uniform salaries so that their compensation is not decision-based. Another approach considered by Shavell is an appeals process whereby judges face a sanction for reversed decisions that would deter them from deviating from the socially optimal sentence. In NZ, HSE offence appeals are relatively rare. In the database used in the present study, fewer than 1 percent of DC cases involving an injury were subsequently appealed. This is suggestive of the presence of meaningful constraints on judicial sentencing variability. In addition, although flexibility in sentencing is a precursor to sentencing variability, it is not a sufficient condition. Although the *de Spa* Guidelines are thin on sentencing detail, the reasoning adopted in *de Spa* is explicitly followed by the judiciary in the vast majority of cases. Further, it is common for sentences to be benchmarked to those imposed in other, similar cases and thereby encouraging reduced sentencing variability. And even though little sentencing guidance is available within prescribed ranges, the HSE Act does provide precise ranges for sentencing.

### 3.2 The United States Criminal Sentencing Guidelines

Following efforts by various U.S. states in earlier years, the Comprehensive Crime Control Act 1984 created the U.S. Sentencing Commission, which introduced empirically-based Federal Sentencing Guidelines that imposed tight limits on judicial discretion. Determinate (mandatory) sentencing was introduced, many offender characteristics were given little if any weight, and judges faced a fairly strict menu of punishment levels for offences involving conviction. Judges were typically bound to set sentences within an appropriate range and ranges were set sufficiently tightly so that judicial discretion should have been significantly attenuated. The objectives of the U.S. Guidelines were to promote deterrence, incapacitate dangerous offenders, impose just punishments, and rehabilitate convicted offenders.

It is evident that the *de Spa* Guidelines stand in stark contrast to the above. The U.S. Guidelines provide judges with limited discretion in ranges of sentences which are computed in a mechanical fashion and which give most weight to a single characteristic for both offences (seriousness) and offenders (criminal history). The *de Spa* Guidelines do make reference to the extent of harm suffered by victims of

---

sanction to an offender, and low marginal cost offenders would require a larger dose of sanction to compensate. Yet in practice, only criminal history seems to intrude given that repeat offenders appear harder to deter.

workplace accidents and to the employer's safety record but no schedule of penalty ranges was specified for these or the remaining criteria. While the *de Spa* Guidelines appear more comprehensive in terms of offence and offender characteristics, judicial discretion seems much more proscribed in the narrowly focused, more stringent U.S. Guidelines.[12]

### 3.3 Criminal Sentencing Variability in New Zealand

Empirical analyses of criminal sentencing in NZ have focused on the variation in average sentences imposed by District Court judges across the different court districts. For example, the appendix to NZLC R94 is a synopsis of a commissioned report by Taylor Duignan Barry Ltd. (2006) which investigated the degree of variation in prison sentencing across seventeen NZ District Court regions for a variety of offences. They analysed 65 different offence types and 12 different offence sub-categories for 2004-05 in terms of the percentage of offenders sent to prison for a given offence type, and, conditional on being imprisoned, the nominal term of imprisonment. While considerable variation in sentencing seems apparent in their data, little statistical analysis of such variation is reported. Further, there are no controls for offence or offender characteristics, and sentencing variation across individual judges is not addressed. NZLC R94, however, accepted these results as demonstrating "substantial variations in practice" which "are unlikely to be explicable on the basis of differences in offence or offender variables. Instead, they clearly indicate that some courts are systematically more severe than others, at least in relation to the percentage of convicted offenders who are imprisoned" (p. 20).

The NZ Law Commission's Sentencing Establishment Unit (SEU) (2008) further addressed the issue of judicial preference at DC level. They referred to two papers (SEU (2007a), SEU (2007b)) dealing with DC variation in the use of imprisonment which concluded that substantial variation existed over the period 2004-06. SEU (2007a) did not control for variation in the nature of offending or offenders and found that the highest imprisonment rate was over 50 percent greater than that for all DCs and more than double the lowest rate. SEU (2007b) controlled for the variation in the composition of offending by creating 157 groups of offenders sentenced with some resulting reduction in variation across regions but the highest imprisonment rate remained over twice that of the lowest rate. SEU (2008) instead

[12] Empirical analyses of the U.S. Guidelines focused on whether or not they achieved a goal of reducing undesirable sentencing disparity among individual judges. The results are mixed, in part because of important conceptual differences among the studies. Sentencing proportionality and disproportionality are frequently distinguished. Proportionality involves sentencing variation that is justified by relevant differences among offenders and their offences. Disproportionality reflects sentencing variation that is not attributable to relevant sentencing factors, and is usually seen to constitute 'unwarranted' disparity. Several studies, including Waldfogel (1991, 1992), Payne (1997), LaCasse and Payne (1999), reported increases rather than reductions in unwarranted disparity in a number of instances. Waldfogel (1998) also found that a reduction in disproportionality was outweighed by an even larger loss in proportionality. Anderson, King, and Stith (1999), however, considered only the variation that arises from the identity of the sentencing judge, and showed that disparity in this sense decreased during the post-reform period. Mustard (2001) found that departures from guidelines produced much of the sentencing disparity arising from race, gender, and offence and criminal record categories, while Schazenbach (2005) inferred race and gender bias in sentencing related to the race and gender composition of the judiciary. Sarnikar, Sorensen, and Oaxaca (2010) also found race and gender disparity in favour of whites and women after controlling for offence severity and criminal record, and that absent the guidelines, the racial gap would actually decline.

focused on drink driving in DC regions with the highest and lowest imprisonment rates given that data existed for a number of characteristics of offending and offenders, and concluded that about 27 percent of the variation in drink driving imprisonment rates between the two regions was due to differences in offenders and their offending, leaving 73 percent of the difference to "local sentencing policy".[13]

SEU recognized that their conclusions that differences in regional sentencing policies constitute the major determinant of regional imprisonment rate variations do not necessarily carry over to the vast number of remaining offences. Nevertheless, they argued that "... a strong inference that it is likely can be drawn. There is no reason to think that drink driving and burglary are special cases" (p. 21). We return to these conclusions when discussing our empirical results for HSE offences in the following sections.

## 4 HSE Offence Data

Our main dataset consists of charge-level information coded from several sources. The major source was provided to us in an electronic form by the NZ Department of Labour and contained a largely comprehensive list of successful prosecutions for HSE offences since inception of the HSE Act through mid-2007 sentencing dates.[14] This database included case names, citations, defendant names, description of any injuries, accident descriptions, offence and decision dates, and amounts of any fines imposed and reparations awarded, along with case decisions and sentencing notes where available. The Department of Labour also provided a file identifying cases by judge in electronic form[15] and was also particularly helpful in tracking down and supplying both electronic and hard copies of decisions, sentencing notes, and returns on prosecutions that were otherwise unavailable to us. Their summaries of facts were also quite useful in breaking down the elements of each charge. In addition, the Safeguard CourtBase provided succinct summaries of each accident and returns on prosecutions for post-2002 cases (however, only since 2004/5 did the returns on prosecutions begin to include information on the sentencing factors). Returns on prosecutions were very useful where no decision/sentencing note was available for a particular case.[16]

We focus on the period from March 1994 to June 2007. March 1994 was selected as the starting point so that all cases examined could be directly influenced by the availability of the *de Spa* Guidelines.

As indicated in section 1, our main empirical objective is to examine the effects of various sentencing criteria on the liability for employers convicted of offences under the HSE Act. We measure an employer's financial liability by the sum of all fines imposed and reparations awarded in each charge/case.[17] If fines and

[13] This makes the very strong implicit assumption that residual sentencing variation can be identified only with "local sentencing policy." Burglary was also examined, and exhibited a similar pattern to drink driving but showing less unadjusted regional sentencing variation.
[14] Department of Labour, *HSE.xls* (unpublished), Wellington.
[15] Department of Labour, *Cases by Judge.xls* (unpublished), Wellington.
[16] Other legal online databases (Brookers, LexisNexisNZ, and Linxplus) had few additional HSE judicial decisions (and even fewer sentencing notes) available.
[17] We refer to 'financial liability' rather than 'financial penalty' since it is arguable whether reparation should be considered to be a penalty. Mason (2008) argued that reparation is restorative and attempts to

reparations are treated as very close substitutes by employers, their sum is a good measure of financial liability for a particular breach of a statutory duty and, hence, the incentive to comply with the provisions of the HSE Act. Further, until recently (and including the data period of our study), the courts appeared to treat fines and reparations as close substitutes when sentencing under the Sentencing Act in accordance with the 'two-step' approach set out in *Department of Labour v Ferrier Woolscours (Canterbury Ltd)* [2005] DCR 356 in conjunction with the 'totality principle' requiring the overall severity of a sentence to be proportionate to the level of offending. We do not include court costs in our measure of total financial liability for two reasons. First, although we have comprehensive data on fines and reparations, there are many cases with missing information on cost awards. Secondly, there is no indication in the *de Spa* Guidelines that court costs should in a systematic manner depend on the characteristics of the case or the defendant. Therefore, court costs should only appear as a constant term in our models – independent of the level of the explanatory variables we use (and describe below). Our approach was further supported by one of our sensitivity checks (results available on request) where the inclusion of court costs in the total financial liability did not qualitatively affect any of our findings.

With respect to sentencing criteria, the data we code contains detailed information on the characteristics of each charge/case (such as the degrees of harm and culpability, employee breach of duty, and the presence of remedial action) and the defendant (such as the employer's safety record, need for particular deterrence, and financial limitations and size).[18] Using this information (where available), we create proxies for the case characteristics specified in the *de Spa* Guidelines. For each characteristic, we create categories which resemble the categories most often used in the case decisions and/or sentencing notes. In particular, we code the *de Spa* factors as follows:

1. The degree of culpability: we assign each charge/case into one of the following six culpability categories: 'low', 'low-medium', 'medium', 'medium-high', 'high', and 'unknown';

2. The degree of harm resulting: we use four mutually-exclusive categories of harm: 'low or medium', 'high', 'fatal', and 'unknown';

3. The financial circumstances of the offender: we use a binary variable to indicate the presence of a defendant's financial limitations;

4. The attitude of the offender, including remorse, co-operation, and taking remedial action: the presence of remorse, cooperation, and remedial action is

---

right a wrong, and as such is not punitive in nature. Hall (2009) is less emphatic, but also raised doubt (at para I.3.5) as to whether 'true' reparation should be classified as punishment rather than constituting the equivalent of civil damages. Clark (2008), however, noted that the compensatory purpose of reparation does not imply that reparation awards have no punitive or deterrent effect. Whatever the intent of a sentence, our focus is on the bottom line and as such effectively treats 'financial liability' as a financial penalty.

[18] Unfortunately, the quality of information available to us varies and we know relatively little about some of our early charges/cases.

indicated by three separate binary variables – one for each of the expressions of the offender's attitude;

5. Any guilty plea: indicated by a binary variable;

6. The need for deterrence: the need for deterrence is expressed by two binary variables indicating separately the 'need for particular deterrence' and the 'need for general deterrence';

7. Compensation to the victim under s 28 Criminal Justice Act 1985: the HC judges considered that the level of fine should be set prior to deciding if, and how much, of the fine would be awarded to the victim. Therefore, we do not include compensation to the victim as an explanatory variable in our models of the financial liability imposed;

8. The employer's safety record: we use six categories of the defendant's safety record: 'poor', 'previous convictions', 'no previous convictions', 'good', 'great', and 'unknown';

9. The facts of the particular case: in some of our models described below, we include additional characteristics of each case. Namely, we create separate binary variables for the presence of a voluntary payment, employer attendance at a restorative justice conference, and an employee breach of duty.[19] We also express the size of the employer as: 'small', 'medium', 'large', or 'unknown'.[20] Finally, we indicate in which year the offence took place in order to account for a national trend in HSE sentencing including any structural shifts in sentencing after the implementation of the Sentencing Act 2002, the introduction of the HSE Act amendments, and the end of the relatively low-penalty 'honeymoon period'.[21]

To obtain a proxy for the need of general deterrence (in addition to the judge's explicit recognition of such need), we merged our charge-level data with industry-level accident and employment data obtained from two separate sources. Our accident data comes from the Accident Compensation Corporation (ACC) Injury Statistics and consists of annual numbers of new claims for work-related injuries in June-years 1993-2007. Data for 1993 has never been produced and so we use 1992 data instead. In 1992, claims were aggregated by industry at the 3-digit New Zealand Standard Industry Classification (NZSIC) level and starting in 1994, they were aggregated at

[19] It is surprising that employee breach was not included in the *de Spa* list and it is evident that it took some time before employee culpability had much impact on employee penalties. The courts specifically referred to a hierarchy that focused primarily on employer culpability; see *Moore v Department of Labour*, unreported, HC Christchurch, 5 July 2001, A 50/01. Nevertheless, sentences could be lowered for employers (employee culpability being a mitigating factor) even in the early years.

[20] As expected, smaller employers are more likely to be subject to financial limitations but the correlation is far from perfect. While none of the employers identified as 'large' are recorded as having financial limitations, 33% of 'small' employers and 67% of 'medium' employers have financial limitations recorded. Data on employer size, however, is limited.

[21] See the relevant judicial comments by Judge Everitt in *Department of Labour v Asian New Zealand Meat Company Limited and Command Pacific (NZ) Ltd*, unreported, DC Dunedin, 28 April 1995, CRN 4/14001 and 4/14002-4/mm, and Judge Robertson in *Hirepool Auckland Ltd v Department of Labour*, unreported, HC Auckland, 4 February 1997AP 301/96.

the 2-digit Australian and New Zealand Standard Industry Classification (ANZSIC) level. While the two classifications are similar, they are not identical and so we needed to create our own industry classification which encompasses both of the standard classifications and is consistent over time.

In order to calculate ACC accident rates, we divided the number of new ACC claims by total employment (of individuals age 15 and above) from the Household Labour Force Statistics (HLFS). HLFS switched to ANZSIC industry classification in 1997 and so we had to make similar adjustment as with the accident data above.

ACC accident rates were merged to our charge-level data by industry and the decision year and, on average, half-year lags were used. Lagging the accident rates and using industry-level aggregates helps mitigate the possible problem of reverse causality where accident rates are themselves determined by the level of financial liabilities (or a particular financial liability in our case). Importantly, an industry code for the activity (rather than the employer's main industry classification) was used is assigning industry-level accident rates. This approach assumes that general deterrence applies to mitigating hazards related to specific high-risk activities, rather than in selected groups of firms. This is consistent with prevalence as a general sentencing factor (Hall (2009 at para I.5.4)) and with the Department of Labour's Occupational Safety and Health Service (OSH) priority areas.[22]

In one of our sensitivity analyses (results available on request), we further tested for the importance of general deterrence by adding a lag of the industry-level annual number of successful prosecutions calculated from our main dataset. This was done to acknowledge that judges may be more sensitive to cases that reach the prosecution stage than simply to employment accidents. A drawback is that no lag could be constructed for the first year of data, 1994, and so we had to exclude 1994 cases from these estimations. Overall, all of our results proved robust to the addition of the number of successful prosecutions and the new variable itself was statistically insignificant.

Finally, we want to take into account the potential effects of price level changes on the nominal level of financial penalties imposed. To do that, we merged our charge-level dataset with Consumer Price Index data provided by Statistics New Zealand. We tried linear as well as logarithmic specifications. In a linear model, changes in the price level are allowed to affect the level of liabilities. In a logarithmic model, inflation is allowed to have a percentage effect on liabilities. The results were very similar across these different specifications and so, for an ease of interpretation (especially with interacted explanatory variables), we only report CPI coefficients in the tables below.

Our master dataset includes 2,064 charges. Out of those, we primarily focus on s 6 offences that are by far the most common (representing 47% of all charges in our dataset). Section 6 of the HSE Act states that "Every employer shall take all practicable steps to ensure the safety of employees while at work" and section 2A of the HSE Amendment Act qualifies "all practicable steps" as "all reasonably practicable steps". As such, a s 6 offence is a very general offence (unlike most

[22] http://www.osh.dol.govt.nz (accessed 21 April 2009).

criminal offences in New Zealand). We focus on s 6 offences (including the 1% of such offences involving any subsection of s 6) in order to limit ourselves to a coherent set of charges for which similar sentencing criteria (and hence a similar application of the *de Spa* Guidelines) might be expected. In order to achieve further coherence, we only examine charges for an injury (as opposed to an incident), limit ourselves to District Court (as opposed to High Court) cases, and study convictions without a discharge. Limiting ourselves to District Court cases has the added advantage of investigating sentencing practice (and variation) as it 'initially occurs' (as opposed to practice which 'withstands the threat' of a High Court appeal). For all of our charges, we identify which District Court and judge handled the case. With respect to discharges, we have decided not to estimate a two-stage Heckman model (estimating first the probability of discharge and only then the magnitude of the financial liability conditional on no discharge) for two reasons: First, only 12 of the charges satisfying our other sample inclusion criteria involved a discharge and it would be very difficult to estimate the probability of discharge with such a small sample. Second, the *de Spa* criteria only relate to sentencing. As such, these guidelines come into play after (and separate from) a decision regarding conviction. The above restrictions leave us with 775 charge-level observations for which a s 6 charge was either the sole charge or listed in a multi-charge case with a sentence explicitly imposed on the s 6 component. In each of our model specifications (described in section 5 below), we use all observations for which all of the variables of interest (dependent and explanatory) can be constructed.[23]

To check how our model performs in multiple-charge and/or multiple-victim cases (including those without a s 6 offence), we also examine the aggregation of sentences to the case level. This analysis can address the concern that sentencing variability regarding s 6 charges arises in part because judges may attach the whole sentence to a single (often s 6) charge. Our inclusion criteria in the case-level analysis are similar to the charge-level analysis. Namely, we investigate cases which involved at least one injury, were handled by a District Court, and involved convictions without discharge (for each case, we aggregate all charges on which the defendant was convicted). The resulting case-level sample contains 1,100 observations and all observations for which the necessary variables can be constructed are used in each model specification.[24]

***Table 1 about here***

Table 1 presents descriptive statistics for our Section 6 and case-level samples. As expected, the mean total financial liability is higher for cases than for their s 6 components ($11,765 compared to $10,835). The difference is not large but that is consistent with the fact that the mean number of charges per case in our case-level

[23] In charges/cases where information is incomplete, we code some charge/case characteristics as "missing" and include these observations in our regression analysis. While we cannot draw any inference from our results on the "missing" categories, this method enables us to at least preserve and use all the remaining (i.e., reported) characteristics of those charges/cases.

[24] In both the s 6 and the case-level samples, we include the rare s 49 (less than 1%) penalties in addition to the s 50 penalties. We tested the robustness of our results to the exclusion of s 49 penalties and they remained qualitatively unchanged.

dataset is 'only' 1.32.[25] Overall, the case-level sample seems to include slightly less serious offences than the sample limited to s 6 charges. For example, the degree of culpability is lower in the case-level sample and so is the identified need for particular deterrence. Other than that, the two samples look similar.

## 5 Methods

Variation in sentencing can be caused by differences in the case-mix, differences in judicial preferences, or both. Figure 1 illustrates the 'raw' (i.e., not controlling for case-mix differences) inter-court and intra-court district variation of the financial liabilities imposed for s 6 offences in the period from March 1994 to June 2007. This information was obtained from an OLS model regressing the financial liability on District Court binary variables (focusing on District Courts which handled at least 20 charges in our s 6 dataset and using Wellington as the omitted category), year binary variables, and CPI. Aggregating districts with less than 20 charges attenuates small-number variation, preserves the degrees of freedom in our OLS model, and makes the graphical exposition easier. Three data series are presented – the central series represents the difference between the average financial liability imposed in a given district and the average financial liability imposed in Wellington. The other two series are constructed by adjusting the above estimate by its standard deviation (by subtracting/adding the standard deviation from/to the estimate). Hence, the fluctuation of the central series represents variation in sentencing across court district and the distance between the lower and the upper series represents variation within districts. In this setup, Hastings has the lowest liabilities and Nelson the highest liabilities but none of the differences with Wellington are statistically significant.[26]

***Figures 1 and 2 about here***

When controlling for case-mix differences by adding the specific factors from the *de Spa* Guidelines (i.e., all the factors listed above except for 7 and 9) as explanatory variables to our model, the ranking of court districts with respect to the severity of financial penalties changes (Figure 2). In particular, when case characteristics proxied by our *de Spa* variables are held constant, Manukau has the lowest penalties (as compared to Wellington) and Hamilton has the highest. The differences between the financial penalties imposed in Nelson, Waitakere, and Hamilton and the financial penalties imposed in Wellington are now statistically significant at the 90% confidence level suggesting that the three districts may have a judicial preference for higher financial liabilities.[27] However, this result needs to be interpreted with caution because other 'facts of the particular cases' need to be taken into account. We do so formally in our 'full' models described below but here we note that differences in the case-mix across court districts do seem to matter for explaining district variation in sentencing. Interestingly, while controlling for case characteristics increases slightly the difference between the lowest and the highest average penalties

[25] In particular, 77% of cases in our case-level dataset involve a single charge, 18% involve two charges, 3% three charges, 1% four charges, and 1% more than four charges (up to a maximum of seven). Of the single-charge cases, 64% are s 6 (or a subsection thereof) charges.
[26] This may be compared with the unadjusted regional imprisonment rates in SEU (2008) where the highest (lowest) rates were in Napier (Wellington).
[27] Unlike SEU (2007b), we do not identify Napier as an outlier.

imposed (from $6,501 to $6,977), within-district variation is reduced (with the unweighted average standard deviation decreasing from $3,183 to $2,697). In other words, controlling for case characteristics seems to slightly increase sentencing variation *across* court districts but decrease variation *within* districts.[28]

***Figures 3 and 4 about here***

In Figures 3 and 4, we report results of a similar exercise with individual judge effects (rather than District Court effects) on the right-hand side. Specifically, these two figures illustrate inter-judge and intra-judge variation in financial liabilities imposed – both 'raw' (Figure 3) and controlling for differences in case characteristics proxied by our *de Spa* variables (Figure 4). For ease of exposition, we only report coefficients for judges who handled at least 10 charges in our Section 6 dataset ('other judges' represent the omitted category). Out of the 17 judges identified, two have significantly higher 'raw' financial liabilities (at the 95% confidence level) than the rest (Figure 3). However, when case characteristics are controlled for (Figure 4), the ranking of individual judges changes and inter-judge variation is substantially mitigated, consistent with what might be argued to constitute a lower level of 'unwarranted' variation. Once again, differences in the case-mix do seem to matter for explaining sentencing variation. Notably, the inter- and intra- judge variations are: 1. greater than the inter- and intra- district variations, respectively, and 2. both reduced by controlling for case characteristics.[29] Finally, when we regressed the total financial liability on judge dummy variables only, the R-squared of the model was very low at 0.02. Moreover, the marginal explanatory power of judge dummies decreased when case characteristics were added to the model. In particular, the inclusion of judge dummies increased the R-squared just by 0.01 (from 0.61 with case characteristics only to 0.62 with case characteristics and judge dummies).

So what effect do the specific *de Spa* criteria and other case characteristics have on the level of financial liabilities imposed? To study this question, we first estimate a 'baseline' OLS model of the following form:

$$Liability_i = \alpha + \boldsymbol{\beta} X_i + \Sigma \theta_t Time_i + \gamma CPI_i + \varepsilon_i$$

where $i$ indexes charges, *Liability* stands for the sum of fines imposed and reparations awarded, $X$ is a vector of the specific factors from the *de Spa* Guidelines, *Time* are year binary variables, *CPI* is the Consumer Price Index, and $\varepsilon$ is a normally distributed error term. We calculate robust standard errors to correct for heteroskedasticity. In a second model, called the 'full' model, we add several other 'facts of the particular cases' (voluntary payment, restorative justice conference, employee breach of duty, and employer size) and District Court binary variables to the right-hand side. We also allow for interactive effects of our explanatory variables. We do not include sentence starting points in our 'full' model for two reasons: First, starting points themselves can be a function of some of our other explanatory variables and as such represent an intermediate outcome rather than a pure input variable. Also, we cannot effectively include starting points due to data limitations – only 4% of our s 6 charges have the existence (and not necessarily the level) of a

[28] In comparison, in the SEU (2008) study of selected crimes, case- mix adjustment lowers rather than raises district variation in sentencing, while individual judge variations are not examined.
[29] Result 2 is also found in the SEU (2007a, 2007b) studies on imprisonment rates.

starting point recorded. Finally, we estimate a model of the effects of the factors identified as key in the U.S. Guidelines and regress the total financial liabilities on the degree of harm, employer's safety record, year binary variables, and CPI only.

As a robustness check, we repeat all of the above estimations with judge individual effects (instead of District Court effects) on the right-hand side and the results are qualitatively unchanged. We also test the sensitivity of our results to using the natural logarithm of the financial liability as the dependent variable. Again, our main results remain unchanged. In addition to estimating all of the above models on the coherent s 6 sample (Table 2), we also check how our models perform in multiple-charge and/or multiple-victim cases (Table 3). Finally, we allow the effects of case characteristics to differ before the implementation of the Sentencing Act 2002 (i.e., prior to 30 June 2002) and after the introduction of the HSE Act amendments (i.e., after 1 May 2003) by estimating separate models for the two time periods (Table 4). We have too few observations between July 2002 and April 2003 to be able to draw reliable inferences about the interim period.[30]

## 6 Results

Column 1 of Table 2 reports our 'baseline' results for s 6 offences. Many of the *de Spa* criteria have the expected effect on financial penalties and are highly statistically significant. The financial liability increases significantly with the degrees of culpability and harm,[31] and with the need for particular deterrence. For example, the financial liabilities are on average $18,908 higher for fatal accidents than for accidents where harm is classified as 'high' and a 'high' degree of culpability (compared to 'medium' culpability) increases financial liabilities by $11,462. The most significant mitigating factor seems to be the defendant's financial limitations, reducing financial liabilities by $7,006, on average.

***Table 2 about here***

Two puzzling results emerge in the baseline model: First, 'great' safety record is associated with higher financial liabilities, *ceteris paribus*. However, this result is not very robust across our model specifications and so we do not have much confidence in it. Second, a puzzling result occurs with respect to the financial implications of remorse. In particular, our 'baseline' Section 6 model indicates that the presence of remorse *increases* financial liabilities by $7,757, on average! We have several hypotheses about such an effect and we plan to investigate them rigorously in future work. For example, remorse could be seen as a signal of guilt in the absence of a guilty plea, or judges do discount fines for remorse but raise reparation awards by even larger amounts when remorse is expressed. Alternatively, it is possible that our measures of the 'seriousness' of an offence derived from case judgements and sentencing notes are too broad and the presence of remorse indicates that the defendant is likely to be at a high end of our culpability and/or harm range. To shed some light on this latter mechanism, we study interactive effects of remorse and the

[30] This is also true for the pre-*de Spa* period.

[31] It might be noted that the two most important factors identified by the Sentencing Establishment Unit of the Law Commission when framing proposed formal criminal sentencing guidelines are statistically significant in our study of HSE offences.

degrees of culpability and harm (as well as differential effects of remorse by District Court) in our 'full' model (column 2 of Table 2).

In the 'full' model, the effect of remorse *per se* becomes insignificant but remorse significantly increases financial liabilities in cases with fatal harm.[32] The liability for remorse seems to be by far the largest (and the most statistically significant) in Waitakere. With respect to the other *de Spa* factors, a high degree of harm has an independent positive effect on financial liabilities and so does the need for particular deterrence. Interestingly, being a small employer brings an additional sizeable ($15,658) and statistically significant reduction of liabilities. Being a medium size employer reduces liabilities by $9,718. A puzzling result in the full model is that employer attendance at a restorative justice conference greatly increases financial penalties (by $60,997). As shown below, this effect is much reduced in our case-level analysis but it remains positive, statistically significant, and relatively large (at $26,308). Our hypotheses here resemble those for remorse but our scope for exploring them further is limited by the small number of restorative justice conferences in our data. In particular, only 6 charges in our s 6 dataset and 15 cases in our case-level dataset involved a restorative justice conference.[33]

Column 3 of Table 2 reports the effects of the factors considered in the U.S. Guidelines. Here, the financial liability imposed increases substantially with the degree of harm and weakly also with a record of previous convictions (whose coefficient is statistically significant at the 90% but not the 95% confidence level). Interestingly, the R-squared of this simple model is relatively high, 0.37 as compared to 0.64 in the 'full' model, meaning that the simple model is still capable of explaining a relatively large fraction (37%) of the variation in financial liabilities.

As noted earlier, in the District Court *de Spa* case a fine of $6,500 was imposed on the employer and this was later increased to $15,000 by the High Court which noted that an initial fine of as much as $20,000 would have been unlikely to be overturned on appeal. What financial liability would our s 6 models 'retrospectively predict' for the *de Spa* case? Applying the estimated parameters from our 'baseline' model to the factors in the *de Spa* case as evaluated by the District Court and the High Court, we predict total financial liabilities of $22,370 and $21,878, respectively. These are very close to the value deemed appropriate by the High Court. The 'full' model reduces these estimates slightly to $20,564 for the District Court and $17,918 for the High Court while the model based on the U.S. Guidelines predicts $20,888 for both.

***Table 3 about here***

Our case-level models (Table 3) contain the same explanatory variables as the Section 6 models with the addition of one factor – the number of charges involved. As expected, this variable is positive (at $2,500-3,500) and highly statistically significant across all of our model specifications. The effects of the *de Spa* case characteristics in the 'baseline' model (column 1) are qualitatively very similar to those observed in the

[32] In an OLS model, the overall effect of remorse on the total financial liability imposed can be calculated by summing the applicable remorse-related coefficients.

[33] To our knowledge, restorative justice conferences were unavailable in the early years of our study period and that may be the reason why they were not explicitly mentioned in the *de Spa* Guidelines.

s 6 analysis above and are of the same order of magnitude. An interesting exception is the effect of remorse which becomes much smaller ($3,137) and insignificant. In the 'full' model (column 2), however, remorse still substantially and statistically significantly increases financial penalties in cases with fatal harm. The liability for remorse also remains high in Waitakere. All other significant effects in the 'full' model mimic those from our s 6 analysis with the only exception (mentioned above) of a substantially smaller liability for a restorative justice conference. Finally, the model limited to the U.S. variables on the right-hand side (column 3) also supports results from the s 6 analysis above. In addition, 'poor' safety record and a safety record with previous convictions significantly increase case-level liabilities.

Even though most of the statistically significant coefficients in the case-level analysis closely resemble those from the Section 6 models, the 'retrospectively predicted' financial liabilities for *de Spa* are somewhat higher. This is an intuitively appealing result since the *de Spa* case involved 3 charges (Sections 6, 7, and 10) and so the overall liability should be larger than a liability for the s 6 component.[34] More specifically, applying the estimated parameters from our 'baseline' model to the *de Spa* case as evaluated by the District Court and the High Court, we now predict total financial liabilities of $30,128 and $29,719, respectively. The 'full' model reduces these estimates to $23,748 for the District Court and $22,100 for the High Court while the model based on the U.S. Guidelines predicts $31,528.

***Table 4 about here***

The first two columns of Table 4 present results of our 'baseline' Section 6 analysis stratified into two time periods: before the implementation of the Sentencing Act 2002 (i.e., prior to 30 June 2002) and after the introduction of the HSE Act amendments (i.e., after 1 May 2003).[35] The most obvious pattern here is that larger financial amounts are at stake in the more recent period. So, while high culpability, high harm, and the need for particular deterrence significantly increase the financial liability in both periods, the effects are quantitatively much larger in the recent period. For example, the liabilities for 'high' culpability were $5,748 prior to June 2002 and $15,886 after May 2003. Similarly, a discount was awarded on the basis of the defendant's financial limitations in both periods but it was much larger in the later period ($16,106 compared to $2,769). Interestingly, the puzzling liabilities for remorse and a 'great' safety record significant in the full sample are only observed in the later period. Also, while the earlier period (like the entire period) exhibits no clear time trend, there is a significant increase in the financial liabilities imposed post-2003. On the other hand, the earlier period (and not the later period) exhibits statistically significant liabilities for a safety record with previous convictions. The results from the case-level analysis stratified by time period (the last two columns of Table 4) are qualitatively similar but the presence of remorse and the employer's safety record do not reach statistical significance in either period. Interestingly, the number of charges was a significant predictor of the financial liability imposed only in the earlier period.

[34] Unlike in many subsequent District Court cases, the sentence in the High Court *de Spa* case was not attached to any specific charge or charges.

[35] It is important to note that the later period contains much richer data as it largely corresponds with the period for which the Department of Labour provides electronic records of decisions, sentencing notes, summaries of facts, and returns on prosecutions.

## 7 Conclusions and Policy Implications

This study focuses on explaining the determinants of sentencing variation for health and safety in employment offences in New Zealand through mid-2007 mainly in terms of the criteria listed in the guideline judgment of the *de Spa* appeal case and subsequently codified in the Sentencing Act. Unlike many other studies (particularly those examining the U.S. Federal Sentencing Guidelines for criminal offences), however, we do not put too much emphasis on explaining sentencing variation *per se* since an effective comparison of pre- and post-*de Spa* Guideline sentences is thwarted by the paucity of pre-*de Spa* observations. Instead, our primary focus is on what we can explain. For example, we show how financial liability for employers increases with the degrees of culpability and harm, and with the need for particular deterrence. Interestingly, a model limited to some of these basic criteria (the degree of harm and the employer's safety record) similar to those used in the U.S. Guidelines has a relatively high explanatory power. The most significant mitigating factors seem to be the defendant's financial limitations and small employer size. Other variables, such as a 'guilty' plea, cooperation, or the need for general deterrence do not seem to play a significant role. We resist any temptation to identify residual variation in sentencing with variation in judicial preferences. The residual variation has three sources. First, a mix of factors for which we cannot account more precisely (e.g., our widespread use of binary rather than continuous variables). Second, omitted explanatory variables; notwithstanding our attempt to be comprehensive in respect of those variables frequently mentioned by the judiciary, idiosyncratic circumstances of cases are not included. Third, variation in judicial preferences. All three sources are likely to play a role. Further, some period-specific unexplained sentencing variability is likely to be affected by major legislative shifts such as major changes in the cap on fines and the introduction of uncapped, insurable reparations that have primacy in the sentencing process but without any specific sentencing guidelines developed to date.

Given that the *de Spa* Guidelines (including their codification) provide no guidance with respect to recommended magnitudes of the effects of their specific sentencing criteria, we demonstrate the average effects of these criteria, including a number of factors not observable until trial. For policy purposes, the question is whether society as a whole approves of financial liabilities for employers of the magnitudes we report. If not, perhaps tighter guidelines including constraints on adjustments made once additional information (e.g., on aggravating or mitigating circumstances) becomes available at trial are warranted. Even if total financial liabilities are deemed appropriate, there may be some concern by society at the contribution of some variables to sentences; e.g., the apparent 'double discounts' for small employer size and employer judgment-proofness. If, however, the outcomes are deemed satisfactory, the model could be used, in practice as well as principle, to determine a sentencing endpoint (rather than the more conventional starting point). Judges would then know that they were sentencing in a representative manner. We would recommend the use of the model estimated for the post-HSE amendment period since a number of coefficients are substantially larger after May 2003 than previously (e.g., the liability for 'high' harm has almost tripled and the discount for defendant's financial limitations has increased more than five times). Implementation would require continuous updating of data used in the regressions to take account of issues such as inflation, appeal judgments relevant to sentencing, further HSE or Sentencing Act amendments, trends in sentencing, and the like. For example, major modifications to

the *de Spa* Guidelines were introduced by the appeal judgment (involving three cases) in *Department of Labour v Hanham & Philp Contractors Limited* [2008] 6 NZELR 79. Financial liabilities in the form of both fines and reparations were typically increased very significantly while substantial sentencing starting points (depending on employer culpability) were also established.

A related issue is that our models generally retrospectively predict the High Court's position in *de Spa* well. By and large, judges have followed the *de Spa* Guidelines. Consequently, sentences prescribed by our model would seem unlikely to be successfully appealed unless extraordinary circumstances prevailed.

There are limitations on what judges report in their decisions and sentencing notes, and on what is reported in the Department of Labour's returns on prosecution. As noted earlier, much information on explanatory variables, especially in the earlier years, is missing. Broad categories used by judges may affect the interpretation of some of the other explanatory variables employed; e.g., the presence of remorse may be correlated with accident seriousness and employer culpability not picked up in the broad measures widely used. Importantly, however, we have much the same information as does the judiciary as a whole regarding the basis of other judges' decisions.

Further research on HSE sentencing currently being undertaken involves collecting detailed data on evidence of remorse in individual cases. Also, High Court sentencing decisions are being investigated to check the 'appeal-proofness' of existing decisions by re-estimating the models with sentence appeal decisions substituted for their District Court counterparts. Sentencing in cases involving incidents (separate from, or in addition to accidents) is under examination, as is sentencing of employees when they (separately, or jointly with employers) breach the care standard. Further, we are examining the composition of employers' financial liability, since much greater variability can be expected in the ratio of fines to total financial liability than in total liability itself as composition has clearly exhibited marked structural shifts following the implementation of the Sentencing Act, HSE Act amendments, and subsequent changes in court interpretations of the Sentencing Act with respect to the substitutability of reparations (especially those that are insured) for fines. We are also analysing post-December 2008 data to determine the response of the District Courts to the Guidelines as modified in the High Court decision in *Hanham and Philp*.

**Figure 1.**

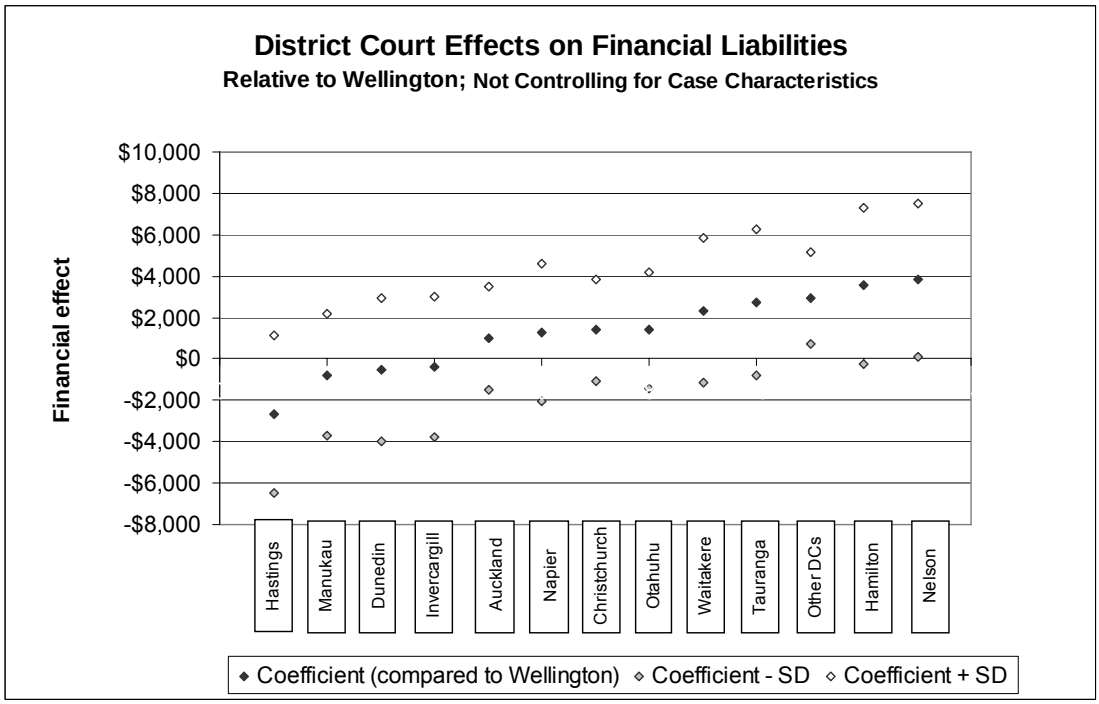


**Figure 2.**

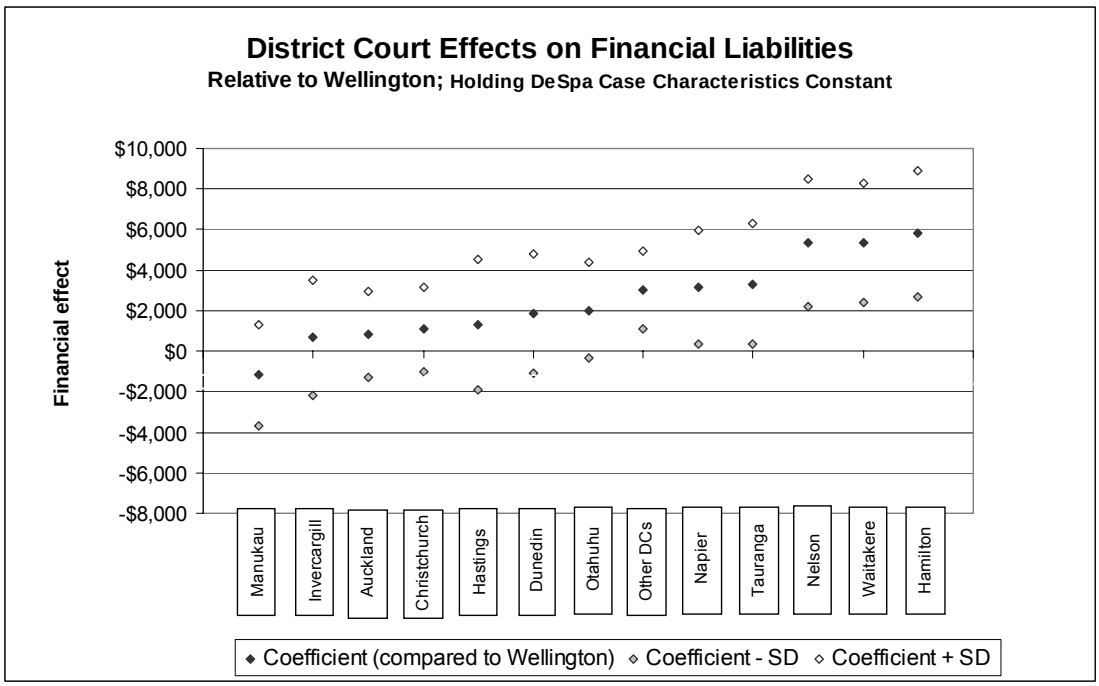

**Figure 3.**

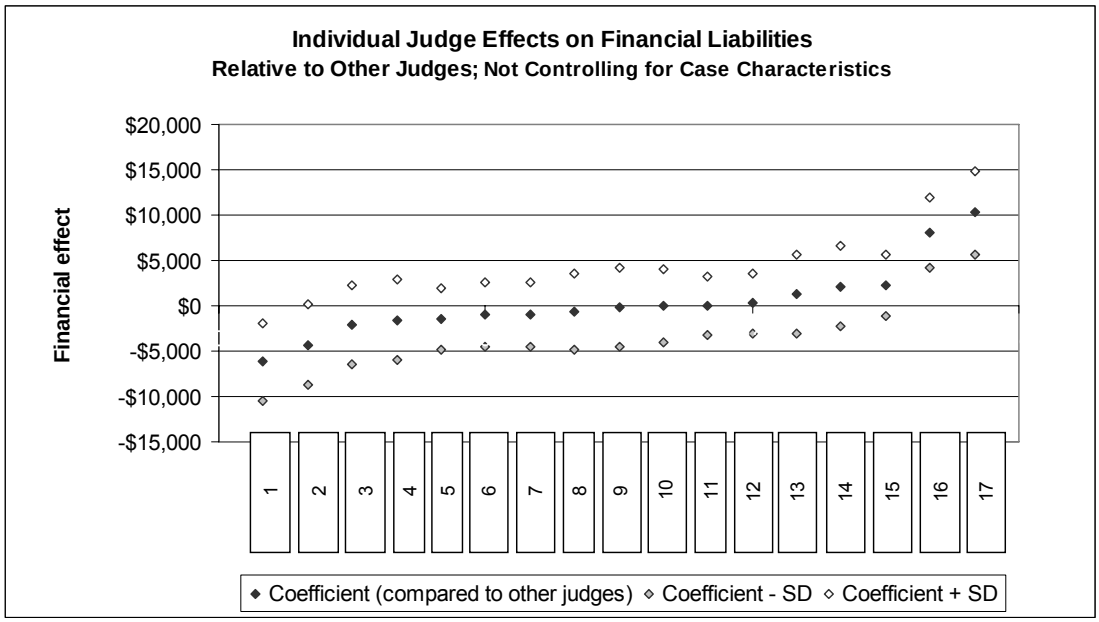


**Figure 4.**

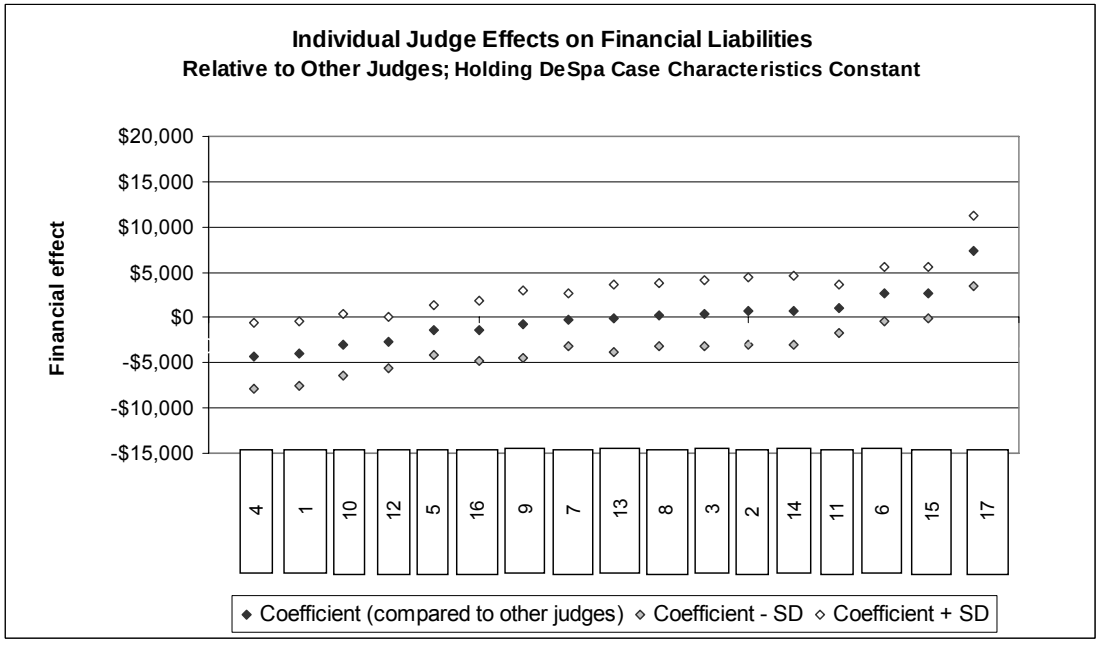

**Table 1. Descriptive Statistics: Section 6 Charges and Cases**

| **Variable** | | **Section 6 sample** (775 obs.) | **Case-level sample** (1,100 obs.) |
|---|---|---|---|
| Financial liability (mean in $) | | 10,835 | 11,765 |
| Degree of culpability (%) | Low | 0.65 | 1.18 |
| | Low-medium | 3.87 | 4.64 |
| | Medium | 10.97 | 9.82 |
| | Medium-high | 5.16 | 4.45 |
| | High | 5.16 | 4.91 |
| | Unknown | 74.19 | 75.00 |
| Degree of harm (%) | Low/medium | 6.71 | 7.27 |
| | High | 82.45 | 81.27 |
| | Fatal | 10.19 | 10.91 |
| | Unknown | 0.77 | 0.55 |
| Defendant's financial limitations (%) | | 3.61 | 5.73 |
| Remorse (%) | | 26.45 | 24.55 |
| Cooperation (%) | | 26.84 | 24.91 |
| Remedial action (%) | | 29.94 | 26.18 |
| Guilty plea (%) | | 48.90 | 45.00 |
| Need for particular deterrence (%) | | 7.23 | 5.91 |
| Need for general deterrence (%) | | 30.06 | 27.18 |
| Safety record (%) | Poor | 1.03 | 0.64 |
| | Previous convictions | 7.48 | 6.18 |
| | No previous convictions | 11.87 | 10.91 |
| | Good | 9.68 | 9.18 |
| | Great | 2.58 | 2.73 |
| | Unknown | 67.35 | 70.36 |
| Voluntary payment made (%) | | 6.06 | 4.91 |
| Restorative justice conference (%) | | 0.77 | 1.36 |
| Employee breach of duty (%) | | 8.26 | 8.09 |
| Size of employer (%) | Small | 2.71 | 3.91 |
| | Medium | 0.39 | 0.36 |
| | Large | 10.97 | 8.28 |
| | Unknown | 85.94 | 87.45 |
| Consumer Price Index (mean) | | 875 | 872 |
| Accident Compensation Corporation (ACC) accident rate (mean of accidents/1,000 employees) | | 30.90 | 31.70 |
| Number of charges per case (mean) | | - | 1.32 |

Specific case characteristics listed in the *de Spa* Guidelines are shaded.

**Table 2. Determinants of the Financial Liability; s 6 Charges**

| Explanatory variable | | 'Baseline' model (765 obs.) | 'Full' model (747 obs.) | U.S. guidelines (765 obs.) |
|---|---|---|---|---|
| Degree of culpability (compared to medium) | Low | -1,002 | -3,964* | - |
| | | (3,007) | (2,154) | |
| | Low-medium | -5,274** | -5,375* | - |
| | | (2,320) | (2,994) | |
| | Medium-high | 14,059*** | 5,192 | - |
| | | (4,361) | (3,436) | |
| | High | 11,462** | 3,014 | - |
| | | (4,754) | (2,539) | |
| | Unknown | 707 | -2,226 | - |
| | | (1,787) | (1,751) | |
| Degree of harm (compared to high) | Low/medium | -2,959** | -4,535*** | -5,353*** |
| | | (1,297) | (1,132) | (1,160) |
| | Fatal | 18,908*** | 8,458*** | 19,910*** |
| | | (3,249) | (1,772) | (3,371) |
| | Unknown | 6,383 | 6,219 | 6,030 |
| | | (4,354) | (4,061) | (4,283) |
| Defendant's financial limitations | | -7,006** | -5,077* | - |
| | | (3,016) | (2,827) | |
| Remorse | | 7,757** | -3,177 | - |
| | | (3,310) | (3,883) | |
| Cooperation | | -7,104 | -5,678 | - |
| | | (5,252) | (3,478) | |
| Remedial action | | 1,130 | 382 | - |
| | | (2,819) | (2,155) | |
| Guilty plea | | -3,789 | -2,115 | - |
| | | (3,087) | (2,084) | |
| Need for particular deterrence | | 9,025** | 8,055** | - |
| | | (4,178) | (3,347) | |
| Need for general deterrence | | 215 | -442 | - |
| | | (1,589) | (1,648) | |
| Safety record (compared to no previous convictions) | Poor | -5,765 | -4,579 | 3,800 |
| | | (4,663) | (4,454) | (3,330) |
| | Previous convictions | 2,325 | 1,177 | 5,702* |
| | | (3,300) | (2,935) | (3,018) |
| | Good | -66 | -2,323 | 1,071 |
| | | (3,025) | (2,559) | (3,419) |
| | Great | 5,368** | 5,140* | 3,223 |
| | | (2,612) | (2,732) | (2,835) |
| | Unknown | 207 | -742 | 234 |
| | | (3,009) | (2,690) | (2,412) |
| Voluntary payment made | | - | 1,841 | - |
| | | | (3,563) | |
| Restorative justice conference | | - | 60,997** | - |
| | | | (25,877) | |
| Employee breach of duty | | - | -1,468 | - |
| | | | (1,580) | |
| Size of employer (compared to large) | Small | - | -15,658*** | - |
| | | | (5,523) | |
| | Medium | - | -9,718** | - |
| | | | (4,892) | |
| | Unknown | - | -6,312*** | - |
| | | | (2,268) | |

| | | | | |
|---|---|---|---|---|
| Degree of culpability (compared to medium) × ***remorse*** | Low | - | 3,349 (8,777) | - |
| | Low-medium | - | 3,047 (4,504) | - |
| | Medium-high | - | 13,040* (7,616) | - |
| | High | - | 6,540 (5,860) | - |
| | Unknown | - | 4,429 (3,363) | - |
| Degree of harm (compared to high) × ***remorse*** | Low/medium | - | 2,054 (3,774) | - |
| | Fatal | - | 18,733*** (5,051) | - |
| Consumer Price Index (CPI) | | 38 (76) | 52 (72) | 85 (79) |
| Accident Compensation Corporation (ACC) accident rate (accidents/1,000 employees) | | - | 22 (26) | - |
| **R-squared** | | 0.48 | 0.64 | 0.37 |

Specific case characteristics listed in the *de Spa* Guidelines are shaded. All models include year binary variables and the 'full' model also includes District Court binary variables and their interactions with 'remorse'. ***, **, and * denote statistical significance at the 99%, 95%, and 90% confidence levels, respectively. Standard errors corrected for heteroskedasticity are reported in parentheses.

**Table 3. Determinants of the Financial Liability; Case-Level Analysis**

| Explanatory variable | | 'Baseline' model (1,099 obs.) | 'Full' model (1,075 obs.) | U.S. guidelines (1,099 obs.) |
|---|---|---|---|---|
| Degree of culpability (compared to medium) | Low | -8,736*** (3,286) | -10,593*** (3,989) | - |
| | Low-medium | -744 (2,440) | -3,481 (2,239) | - |
| | Medium-high | 15,716*** (4,684) | 4,575 (3,143) | - |
| | High | 11,325*** (4,274) | 3,458 (2,781) | - |
| | Unknown | 1,958 (1,546) | -1,538 (1,477) | - |
| Degree of harm (compared to high) | Low/medium | -3,294*** (1,083) | -4,439*** (1,085) | -4,736*** (881) |
| | Fatal | 19,008*** (2,535) | 10,882*** (1,484) | 19,886*** (2,785) |
| | Unknown | 5,373 (3,935) | 4,705 (3,881) | 5,283 (3,882) |
| Defendant's financial limitations | | -7,038*** (2,460) | -5,618** (2,276) | - |
| Remorse | | 3,137 (2,622) | -4,768 (4,291) | - |
| Cooperation | | -4,390 (4,127) | -4,813 (3,355) | - |
| Remedial action | | 2,824 (2,423) | 2,633 (2,149) | - |
| Guilty plea | | -2,136 (2,170) | -1,298 (2,007) | - |
| Need for particular deterrence | | 7,086* (4,131) | 7,001* (3,561) | - |
| Need for general deterrence | | 1,542 (1,356) | 639 (1,420) | - |
| Safety record (compared to no previous convictions) | Poor | -2,675 (5,286) | -986 (6,237) | 7,104** (3,315) |
| | Previous convictions | 2,883 (3,316) | 298 (3,191) | 6,417** (3,093) |
| | Good | 334 (3,020) | -2,581 (2,919) | 1,439 (3,379) |
| | Great | -1,607 (2,699) | -3,737 (2,803) | -1,104 (2,619) |
| | Unknown | -60 (2,591) | -608 (2,414) | -744 (2,244) |
| Voluntary payment made | | - | 2,815 (3,833) | - |
| Restorative justice conference | | - | 26,308** (11,745) | - |
| Employee breach of duty | | - | -947 (1,499) | - |
| Size of employer (compared to large) | Small | - | -17,077*** (4,085) | - |
| | Medium | - | -107 (12,280) | - |
| | Unknown | - | -9,570*** (2,736) | - |

| | | | | |
|---|---|---|---|---|
| Degree of culpability (compared to medium) × ***remorse*** | Low | - | 6,731 | - |
| | | | (7,521) | |
| | Low-medium | - | 5,014 | - |
| | | | (3,755) | |
| | Medium-high | - | 13,057* | - |
| | | | (7,749) | |
| | High | - | 5,623 | - |
| | | | (6,143) | |
| | Unknown | - | 3,826 | - |
| | | | (2,717) | |
| Degree of harm (compared to high) × ***remorse*** | Low/medium | - | -619 | - |
| | | | (4,055) | |
| | Fatal | - | 12,864*** | - |
| | | | (4,451) | |
| Number of charges | | 2,768*** | 2,484*** | 3,524*** |
| | | (818) | (879) | (759) |
| Consumer Price Index (CPI) | | 8 | 31 | 32 |
| | | (65) | (61) | (67) |
| Accident Compensation Corporation (ACC) accident rate (accidents/1,000 employees) | | - | -9 | - |
| | | | (24) | |
| **R-squared** | | 0.44 | 0.54 | 0.36 |

Specific case characteristics listed in the *de Spa* Guidelines are shaded. All models include year binary variables and the 'full' model also includes District Court binary variables and their interactions with 'remorse'. ***, **, and * denote statistical significance at the 99%, 95%, and 90% confidence levels, respectively. Standard errors corrected for heteroskedasticity are reported in parentheses.

**Table 4. Determinants of the Financial Liability; Period-Specific Results**

| Explanatory variable | | s6 Charges; 1 Mar 1994 - 30 Jun 2002 (573 obs.) | s6 Charges; 1 May 2003 – 30 Jun 2007 (160 obs.) | Case-Level; 1 Mar 1994 - 30 Jun 2002 (824 obs.) | Case-Level; 1 May 2003 – 30 Jun 2007 (221 obs.) |
|---|---|---|---|---|---|
| Degree of culpability (compared to medium) | Low | -2,766* | 6,617 | -1,828 | -15,947** |
| | | (1,416) | (5,546) | (2,425) | (7,570) |
| | Low-medium | -1,505 | -9,354** | -3,125* | -4,750 |
| | | (1,132) | (4,361) | (1,760) | (4,025) |
| | Medium-high | 3,196 | 16,365** | 4,027* | 23,519*** |
| | | (1,945) | (6,344) | (2,053) | (7,051) |
| | High | 5,748*** | 15,886 | 8,103*** | 8,851 |
| | | (1,805) | (12,386) | (2,161) | (9,970) |
| | Unknown | 426 | -1,748 | 324 | 267 |
| | | (983) | (3,132) | (1,045) | (2,821) |
| Degree of harm (compared to high) | Low/medium | -2,731*** | -6,692* | -3,285*** | -2,529 |
| | | (544) | (3,896) | (671) | (3,262) |
| | Fatal | 7,647*** | 54,461*** | 9,351*** | 50,789*** |
| | | (1,244) | (8,310) | (1,096) | (7,300) |
| | Unknown | 7,137 | - | 5,930 | - |
| | | (5,221) | | (4,544) | |
| Defendant's financial limitations | | -2,769** | -16,106** | -2,765 | -13,235*** |
| | | (1,343) | (7,212) | (1,791) | (4,575) |
| Remorse | | -912 | 12,279** | -236 | 6,054 |
| | | (1,317) | (5,663) | (1,305) | (4,806) |
| Cooperation | | 858 | -12,413 | -463 | -9,940 |
| | | (1,438) | (8,375) | (1,555) | (6,913) |
| Remedial action | | 2,354* | 4,261 | 2,757** | 4,468 |
| | | (1,229) | (7,571) | (1,199) | (5,447) |
| Guilty plea | | 946 | -19,416 | 1,119 | -9,341 |
| | | (829) | (12,635) | (975) | (9,318) |
| Need for particular deterrence | | 3,792** | 11,877** | 7,186** | 9,477** |
| | | (1,534) | (4,849) | (3,208) | (4,765) |
| Need for general deterrence | | -32 | -1,509 | 345 | -2,118 |
| | | (1,089) | (3,803) | (1,133) | (2,970) |
| Safety record (compared to no previous convictions) | Poor | -2,529 | - | 1,616 | - |
| | | (2,845) | | (4,913) | |
| | Previous convictions | 4,927*** | 3,079 | 3,296 | 2,890 |
| | | (1,715) | (3,686) | (2,000) | (4,782) |
| | Good | 1,878 | -2,457 | 1,821 | -2,101 |
| | | (1,377) | (5,029) | (1,795) | (4,674) |
| | Great | 3,788* | 15,301*** | 3,126 | -540 |
| | | (1,987) | (5,153) | (2,127) | (4,433) |
| | Unknown | 3,111** | -2,966 | 2,192 | -5,153 |
| | | (1,258) | (5,000) | (1,603) | (4,339) |
| Number of charges | | - | - | 3,646*** | -988 |
| | | | | (571) | (2,907) |
| Consumer Price Index (CPI) | | 34 | -73 | 42 | -81 |
| | | (35) | (178) | (38) | (152) |
| **R-squared** | | 0.37 | 0.65 | 0.44 | 0.53 |

Specific case characteristics listed in the *de Spa* Guidelines are shaded. All models include year binary variables. ***, **, and * denote statistical significance at the 99%, 95%, and 90% confidence levels, respectively. Standard errors corrected for heteroskedasticity are reported in parentheses.

***References***

Access Economics, 2006. The Economic and Social Costs of Occupational Disease and Injury in New Zealand. NOHSAC Technical Report 4, Wellington. Available at http:www.nohsac.govt.nz/EconomicandSocialCostsofOccupationalDiseaseandInjury.shtml

Anderson, J., Kling, J., Stith, K., 1999. Measuring inter-judge sentencing disparity: before and after the Federal Sentencing Guidelines. Journal of Law and Economics 42, 271-307.

Bar-Gill, O., Gazal Ayal, O., 2006. Plea bargains only for the guilty. Journal of Law and Economics 49, 353-364.

Clark, A., 2008. Reparation and sentencing. New Zealand Law Journal, 437-439.

Gordon, P., Woodfield, A., 2006. Incentive effects of the changing structure of penalties in New Zealand's Health and Safety in Employment Act. Employment Law Bulletin, 111-115.

Hall, G.G., 2009. Hall's Sentencing (NZ). LexisNexis, http://www.lexisnexis.co.nz (accessed 5 March 2009).

Hughes, J., 2008. Reparation orders and the HSE Act. Employment Law Bulletin, 121-127.

LaCasse, C., Payne, A., 1999. Federal sentencing guidelines and mandatory minimum sentences:do defendants bargain in the shadow of the judge? Journal of Law and Economics 42, 245-269.

Mason, G., 2008. Reparation in sentencing under HSE. New Zealand Law Journal, 33- 42.

Miceli, T., 1991. Optimal criminal procedure: fairness and deterrence. International Review of Law and Economics 11, 3-10.

Miceli, T., 2008. Criminal sentencing guidelines and judicial discretion. Contemporary Economic Policy 26, 207-215.

Mustard, D.B., 2008. Racial, ethnic, and gender disparities in sentencing: evidence from the U.S. federal courts. Journal of Law and Economics 44, 285-314.

New Zealand Department of Labour: Occupational Safety and Health Service (OSH), 2009. http://www.osh.dol.govt.nz (accessed 21 April 2009).

New Zealand Law Commission, 2006. Sentencing Guidelines and Parole Reform. Report 94, Wellington, Law Commission.

Payne, A., 1997. Does inter-judge disparity really matter? An analysis of the effects of sentencing reforms in three federal court districts. International Review of Law and Economics 17, 337-366.

Polinsky, A.M., Shavell, S., 2000. The fairness of sanctions: some implications for optimal enforcement policy. American Law and Economics Review 2, 223-237.

Reinganum, J., 1998. Plea bargaining and prosecutorial discretion. American Economic Review 78, 713-728.

Reinganum, J. Sentencing guidelines, judicial discretion, and plea bargaining. Rand Journal of Economics 31, 62-81.

Sarnikar, S., Sorensen, T., Oaxaca, R.L., 2010. Do you receive a lighter prison sentence because you are a woman or white? An economic analysis of federal criminal sentencing guidelines. (mimeo).

Schazenbach, M., 2005. Racial and sex disparities in prison sentences: the effect of district-level judicial demographics. Journal of Legal Studies 34, 57-92.

Sentencing Establishment Unit, 2007a. Summarised numbers & percentages on type of sentence by court region for 2004-2006. Wellington, Law Commission.

Sentencing Establishment Unit, 2007b. Update on imprisonment rates by court region for 2004-2006. Wellington, Law Commission.

Sentencing Establishment Unit, 2008. Regional variation in District Court sentencing: to what extent is it a consequence of judicial choice or a consequence of variations in offending? Wellington, Law Commission.

Shavell, S., 2007. Optimal discretion in the application of rules. American Law and Economics Review 9, 175-194.

Taylor Duignan Barry Ltd., 2006. Regional Analysis of Variations in District Court Sentencing. NZLC MP0, Wellington, Law Commission.

Waldfogel, J., 1991. Aggregate inter-judge disparity in sentencing: evidence from three districts. Federal Sentencing Reporter 4, 151-154.

Waldfogel, J., 1992. Inter-judge disparity in federal sentencing; evidence from three federal districts 1984-90. Unpublished manuscript. New Haven, Conn., Yale University, Department of Economics.

Waldfogel, J., 1998. Does inter-judge disparity justify empirically based sentencing guidelines? International Review of Law and Economics 18, 293-304.