

VU Research Portal

Consistency and predictability in international sentencing. ICTY and ICTR case study

Hola, B.; Bijleveld, C.C.J.H.; Smeulers, A.L.

published in

European Journal of Criminology
2012

DOI (link to publisher)

[10.1177/1477370812453112](https://doi.org/10.1177/1477370812453112)

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Hola, B., Bijleveld, C. C. J. H., & Smeulers, A. L. (2012). Consistency and predictability in international sentencing. ICTY and ICTR case study. *European Journal of Criminology*, 9(5), 539-552.
<https://doi.org/10.1177/1477370812453112>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

European Journal of Criminology

<http://euc.sagepub.com/>

Consistency of international sentencing: ICTY and ICTR case study

Barbora Holá, Catrien Bijleveld and Alette Smeulers

European Journal of Criminology 2012 9: 539

DOI: 10.1177/1477370812453112

The online version of this article can be found at:

<http://euc.sagepub.com/content/9/5/539>

Published by:



<http://www.sagepublications.com>

On behalf of:



European Society of Criminology

Additional services and information for *European Journal of Criminology* can be found at:

Email Alerts: <http://euc.sagepub.com/cgi/alerts>

Subscriptions: <http://euc.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://euc.sagepub.com/content/9/5/539.refs.html>

>> [Version of Record](#) - Sep 19, 2012

[What is This?](#)

Downloaded from euc.sagepub.com at Vrije Universiteit 34820 on July 12, 2013

Consistency of international sentencing: ICTY and ICTR case study

European Journal of Criminology
9(5) 539–552

© The Author(s) 2012

Reprints and permission: sagepub.
co.uk/journalsPermissions.nav

DOI: 10.1177/1477370812453112

euc.sagepub.com



Barbora Holá and Catrien Bijleveld

VU University Amsterdam, The Netherlands

Alette Smeulers

Tilburg University, The Netherlands

Abstract

The International Criminal Tribunal for the Former Yugoslavia and the International Criminal Tribunal for Rwanda are the first, post Cold War international criminal tribunals convicting perpetrators of genocide, crimes against humanity and war crimes. Their sentencing practice has been largely criticized as inconsistent. This quantitative study addresses the criticism and empirically investigates the consistency of international sentencing. The extent to which the selected factors predict sentence length is tested in a multiple regression analysis. The analysis suggests that similar, legally relevant patterns have emerged in the sentencing practice of both tribunals. Sentencing in international criminal practice does not appear to be less consistent than sentencing under domestic jurisdictions.

Keywords

Consistency, ICTR, ICTY, sentencing

Introduction

For more than 15 years now the International Criminal Tribunal for the Former Yugoslavia (ICTY) and the International Criminal Tribunal for Rwanda (ICTR), the first (post Cold War) international criminal ad hoc tribunals, have been sentencing perpetrators of international crimes. Together they have convicted more than 100 individuals for involvement in genocide, crimes against humanity and war crimes.

Corresponding author:

Barbora Holá, VU University Amsterdam, Faculty of Law, Department of Criminal Law and Criminology, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.

Email: b.hola@vu.nl

The ICTY and ICTR judges have been vested with almost unfettered discretionary powers in sentence determination (UN, 1993; UN, 1994; ICTY Rules, 1994; ICTR Rules, 1995), rendering sentencing decisions vulnerable to local or individual particularities and differential weighting of legally relevant (or even irrelevant) factors. Indeed, the sentencing practice of the tribunals has been heavily criticized and labelled as confusing, disparate and inconsistent (De Roca, 2008; Drumbl, 2007; Olusanya, 2005). Conversely, the scarce empirical evaluative studies (D'Ascoli, 2011; Doherty and Steinberg, 2009; Hola et al., 2009; Meernik, 2003; Meernik and King, 2003; Meernik et al., 2005) have concluded that there appears to be a fair degree of consistency in international sentencing, although differences in sentencing mechanisms between the tribunals have never been analysed.

This research adds to previous literature by examining consistency of international sentencing across the two international tribunals (ICTY and ICTR), investigating whether the sentencing determinants play a similar or a different role at each of the tribunals, and thus whether international sentencing has evolved into a consistent, homogeneous system.¹

A central concept is consistency of 'international punishment'. It is a widely accepted principle of human rights that consistency of punishment is necessary for the fair operation of any criminal justice system (De Roca, 2008; *Prosecutor v. Mucic et al.*, 2001: §756). Several dimensions of this concept can be distinguished. First, all sentences should follow the same underlying principles: 'consistency in approach'. Second, similar offenders (with similar case attributes) should receive similar sentences: 'consistency in outcome (predictability)'. And finally, for international sentencing to be considered consistent there should be no differences between the tribunals: 'systemic consistency'. For more, see Hola (2012: 10–13).

To empirically explore whether the ICTY and ICTR sentencing practice could be labelled as consistent, this article examines all defendants sentenced by the tribunals until 1 June 2010. It should be kept in mind that international sentencing is 'young', and therefore only limited stability in sentencing practice can be expected (Ewald, 2010). This study is therefore exploratory in its attempt to detect and describe these emerging patterns without any pre-set hypotheses. In the following section, we describe the sentencing law and principles that emerged in the case law of the tribunals. We then present the methodology and report our findings. In the conclusion, we situate the results within a broader discussion on international sentencing and sentencing consistency.

ICTY and ICTR sentencing law

As noted by Mettraux (2005: 343), it is hard to identify crimes that are more difficult to sentence than international crimes. This is owing not merely to the atrocity of international crimes but also to their collective, systematic nature, often implicating many individuals. As a direct result, assigning and, for the purposes of sentencing, 'measuring' the culpability of individual perpetrators are extremely complex.

The Statutes of the tribunals (UN, 1993, 1994) provide only very general instructions regarding sentencing. The regulatory framework for sentencing is almost identical for both tribunals. Applicable penalties are limited to imprisonment and, when determining the terms of imprisonment, judges have recourse to the local courts' practices regarding

prison sentences (Yugoslavian or Rwandese) (UN, 1993: Article 24; UN, 1994: Article 23). Reference to local practice has been construed restrictively by the tribunals (*Prosecutor v. Delic*, 2008: §592; *Prosecutor v. Semanza*, 2005: §377), and judges cite applicable domestic provisions mostly as a formality.

Articles 24/23 of the ICTY/ICTR Statute list two general sentencing factors: the gravity of the offence and the individual circumstances of the convicted person. However, it is unclear what is actually meant by the gravity of a crime or individual circumstances. The Statutes do not contain any sentencing tariff or list of factors relevant to sentencing. Judges often stress that the gravity of the offence is the primary consideration in sentencing (*Prosecutor v. Mucic*, 1998: §1225), which has been interpreted as encompassing two aspects: (i) the magnitude of harm caused by the offender, and (ii) the form and degree of participation of the accused in the crime, that is, the offender's culpability (*Prosecutor v. Milutinovic*, 2009: §1147; D'Ascoli, 2011: 145–55; Hola, 2012: 51–4).

The provisions of the Statutes are supplemented by the Rules of Procedure and Evidence. Only Rule 101 is dedicated to factors relevant to sentencing, limiting the range of applicable sentences – the maximum being life imprisonment. It also instructs judges to take into account any aggravating and/or mitigating circumstances when determining the sentence, although only two potential mitigating factors are explicitly mentioned in law: 'superior orders' and 'substantial cooperation with the Prosecutor' (UN, 1993: Article 7(2); UN, 1994: Article 6(2); Rules, 1994: Rule 101; Rules, 1995: Rule 101). In effect, judges are left to determine what factors justify an increase or reduction in sentence length on a case-by-case basis. Mitigating factors need to be proven 'upon the balance of probabilities' (the burden of proof is placed on a defence to establish a mitigating circumstance) and need not directly relate to the charged offences (*Prosecutor v. Nindabahizi*, 2004: §502). Aggravating factors must, however, be proven 'beyond any reasonable doubt' (the burden of proof is placed on a prosecution to establish an aggravating factor) and only those circumstances directly related to the commission of the offence charged and to the offender him/herself when he/she committed the offence may be considered in aggravation (*Prosecutor v. Haradinaj*, 2008: §489).

The general principles of sentence determination that evolved in the case law outlined above, and further below (see the description of individual independent variables), seem to be mutually influential at the ICTY and the ICTR. Judges from one tribunal often refer to the case law of the other in their general sentencing considerations, thus developing a common ICTY/ICTR narrative (Hola, 2012: 72). The tribunals share a joint Appeals Chamber, which also contributes to the development of a consistent jurisprudence (Van der Wilt, 2008: 270).

Methodology

Our data comprise all defendants convicted by the ICTY and the ICTR up until 1 June 2010 ($N = 111$). All finalized sentences pronounced in individual cases ($N = 94$) and all sentences handed down by trial chambers but still pending on appeal ($N = 17$) are included. The information was obtained from written versions of the judgments published on the ICTY and ICTR web pages.

Dependent variable

Sentences were analysed at the level of a defendant, measured in years. Nineteen defendants have been convicted to life imprisonment. Because life sentences are indeterminate, we have to assign them a numerical value to be able to include them in our quantitative analysis. Given that the maximum determinate sentence ever handed down has been 45 years (*Prosecutor v. Kajelijeli*, 2005) and that the majority of defendants were in their mid-40s (or older) at the time of conviction, we recoded 'life' to 55 years – a number reflecting both the time that an average ICTY/ICTR defendant convicted to life imprisonment would spend in prison as well as the severity of the sentence.

Independent variables

The following data were collected for each defendant: (a) category of convicted crime, (b) number of guilty counts, (c) mode(s) of responsibility, (d) number of mitigating factors, (e) number of aggravating factors, and (f) rank in the military or political hierarchy. The independent variables comprise the two main relevant sentencing considerations according to law: the gravity of the crime (category of crime, number of counts, aggravating factors, mode of liability and rank) and the individual circumstances of the offender (mitigating factors). An overview of the variables is presented in Table 1.

Table 1. Measurement strategies and proportions/means of independent variables ($N = 111$)

| Legal factor | Variable | Measurement | Proportion / Mean (SD) |
|---|---------------------------------|----------------------|------------------------|
| Category of crimes (each category coded as a distinct variable) | Genocide | I = yes | 32% |
| | Crimes against humanity | I = yes | 78% |
| | War crimes | I = yes | 54% |
| Number of counts (numerical) | Number of counts | Actual number (1–31) | 4.23 (4.4) |
| Aggravating factors (AFs) (numerical) | Number of AFs | Actual number (0–10) | 3.17 (2.1) |
| Mitigating factors (MFs) (numerical) | Number of MFs | Actual number (0–12) | 4.42 (3.2) |
| Mode of liability (each category coded as a distinct variable) | Superior responsibility | I = yes | 23% |
| | Hands-on perpetration | I = yes | 42% |
| | Joint criminal enterprise | I = yes | 22% |
| | Planning, instigating, ordering | I = yes | 33% |
| | Aiding and abetting | I = yes | 44% |
| Actual rank (dummy coded with low rank as a reference category) | High rank | I = yes | 29% |
| | Middle rank | I = yes | 49% |

Category of crime. The tribunals have jurisdiction over three categories of international crime: genocide, crimes against humanity, and war crimes. The Statutes do not distinguish between the categories in terms of their objective gravity and related sentence length. It is left to the judges' discretion to determine the gravity of a crime on a case-by-case basis. Theoretically, the gravity of a crime can be determined *in abstracto* and *in concreto*.

Gravity *in abstracto* is based on an analysis, in terms of criminal law, of the subjective and objective elements of the crime. The subjective element refers to a guilty mind (*mens rea*) whereas the objective element refers to a criminal act (*actus reus*). Because the *actus reus* and *mens rea* may differ per offence (for example, murder compared with plunder), the objective gravity of individual offences also varies. Similar arguments could be presented with respect to individual categories of international crime: that is, genocide requires special intent to destroy in whole or in part a specific group of people; crimes against humanity must be committed as part of a widespread or systematic attack against a civilian population with knowledge of such an attack; war crimes could be isolated incidents committed within an armed conflict. Owing to these inherent differences, scholars have advocated a hierarchy of individual categories of international crime for the purposes of sentencing (Carcano, 2002; Danner, 2001; Olusanya, 2004).

On the other hand, gravity *in concreto* depends on the actual harm done and the degree of culpability of the offender. The ICTY and ICTR judges have consistently emphasized the necessity to evaluate gravity in light of the particular circumstances of each case (*Prosecutor v. Cesic*, 2004: §32). However, the question of whether a hierarchy exists among the individual categories of crime based on the comparative analysis of their gravity *in abstracto* has also been discussed many times in case law. In the early cases, judges accepted the idea of a hierarchy among genocide, crimes against humanity and war crimes (*Prosecutor v. Tadic*, 1997: §73; *Prosecutor v. Serushago*, 1999: §15). Later, however, both tribunals adopted the stance that there is no pre-established hierarchy, emphasizing that *all* crimes under their jurisdiction are very serious violations of international humanitarian law (*Prosecutor v. Milosevic*, 2007: §989). In order to determine whether sentence severity differs for genocide, crimes against humanity and war crimes, we include three separate dichotomous variables reflecting these crimes.

Number of counts. With most accused found guilty of multiple counts, the so-called principle of totality implies that the sentence reflects the totality of criminal conduct of a defendant (*Prosecutor v. Mucic et al.*, 2003: §46). Arguably, more incidents or the violation of multiple legal prohibitions should result in a more severe sentence. To test this, we recorded the number of counts a defendant was convicted of.

Aggravating factors. Aggravating factors are those circumstances related to the commission of the offence that increase its criminal severity. Factors taken into account in evaluating the gravity of a particular crime may not be reconsidered as aggravating factors (*Prosecutor v. Lukic & Lukic*, 2009: §1050). Clear guidelines as to what factors are relevant for assessing the gravity of the offence and what circumstances constitute aggravating factors have not yet been developed by the tribunals. The tribunals' judges have accepted a large variety of aggravating factors in the case law. 'Abuse of superior position/authority/influence' are the most frequent aggravating factors cited by the judges.

For a list of all the factors that were coded as aggravating when cited by the judges, see Hola (2012: 77). Given our sample size, we will investigate not the impact of each separate factor but instead their total number.

Mode of liability. The modes of individual liability of which a defendant is convicted are indicative of the manner in which he/she participated in crimes. The form and degree of participation of an accused in a crime is one of the elements constituting the gravity of the crime (*Prosecutor v. Milutinovic*, 2009: §1147). Article 7(6) of the ICTY/ICTR Statute differentiates between superior responsibility and planning, instigating, ordering, committing or otherwise aiding and abetting in crime (Article 7(3)(1) / 6(3)(1) ICTY/ICTR Statute). Participation in a joint criminal enterprise (JCE) as a specific liability mode falling under the category of ‘committing’ was developed at the ICTY (Ohlin, 2007; Van Sliedregt, 2003). Each liability mode entails the proof of specific *actus reus* and *mens rea*. Consequently, sentence severity should differ depending on whether a defendant is convicted as a hands-on perpetrator, as an order-giver, or as ‘an aider’. The law, however, contains no guidance for sentence determination in relation to individual modes of liability. Some fragmentary principles have evolved in case law.

Liability modes can either ‘augment’ or ‘lessen’ the gravity of a crime (*Prosecutor v. Ndindabahizi*, 2007: §122). It seems to be currently accepted that sentences of those convicted on the basis of superior responsibility should be generally lower (*Prosecutor v. Hadzihasanovic & Kubura*, 2006: §2076). Superior responsibility entails liability for omission – the failure to prevent or punish subordinates’ crimes. Otherwise, judges distinguish between direct participation (that is, commission proper, participation in JCE, instigating, ordering, planning) and indirect participation (that is, aiding and abetting) (*Prosecutor v. Semanza*, 2005: §388). Direct participation should generally incur a higher sentence than criminal participation by way of facilitating and permitting the commission of the crimes (*Prosecutor v. Mrksic & Sljivancanin*, 2009: §407). We recorded the mode(s) of liability a defendant was convicted of, that is, committing, participating in JCE, ordering/planning/instigating, aiding and abetting, and superior responsibility, with liability modes coded as separate dichotomous variables.

Actual rank. Another important factor in the assessment of the gravity of the crime is the actual position of a defendant within a state structure. Perpetrators of mass atrocities can be divided into three broad categories: leaders (top hierarchical levels), bureaucrats (occupying middle-ranking positions and implementing orders stemming from the top leadership) and field executors (low-ranking individuals actually committing atrocities). These groups represent an ordinal ranking of moral blameworthiness (Smeulers and Hola, 2010). It is generally accepted in legal as well as criminological discourse that those orchestrating crime at the top leadership levels are the most culpable (Del Ponte, 2004). Judges seem to endorse this approach, emphasizing that ‘the most senior members of a command structure, . . . the leaders and planners of a particular conflict, should bear heavier criminal responsibility than those lower down the scale’ (*Prosecutor v. Musema*, 2001: §383). The sentences should thus be graded along these lines, although there is one caveat to this principle of gradation: the position of the offender is just one and not necessarily the most important consideration in sentencing. In exceptional cases

when crimes are extremely grave, even low-ranking defendants could be subjected to severe sentences (*Prosecutor v. Naletilic & Martinovic*, 2003: §744).

We coded the rank of the offender within the overall state hierarchy as a variable with three categories: high, middle and low. The low-ranking category comprises offenders who held little or no power/influence in the overall circumstances of each conflict, such as camp guards, shift leaders in detention camps, rank and file soldiers, local politicians or others with no extensive powers (25 individuals). The middle-ranking category encompasses defendants who had more extensive de jure or de facto authority to command and/or influence the conduct of others, such as camp commanders, local or more senior army commanders and *conseillers*, *bourgemesters* in Rwanda or leaders of the Interahamwe (54 individuals). Finally, the group of high-ranking offenders consists of regional or national military and political leaders, such as members of regional or national governments, regional political leaders, prefects in Rwanda, members of national government, and military officers above the rank of colonel or military commanders of operational sectors in Rwanda (32 individuals). Rank was coded as a dummy variable, with low rank as the reference category.

Mitigating factors. A wide range of factors, generally a defendant's individual circumstances, have been accepted by the ICTY and ICTR judges in mitigation of a sentence. 'Family circumstances of a defendant' or 'his/her assistance of victims' are among the mitigating factors cited the most at the tribunals. For a list of all factors, see Holá (2012: 79). Owing to the limited number of cases in our dataset, we are not able to test for differences between individual mitigating factors and instead focused on the number of factors.

Multiple regression was used to analyse the data. Given the sample size for the ICTR, it is not feasible to run separate models for the two tribunals. By analysing the data from the ICTY and the ICTR in one model, it is assumed that the manner in which the independent variables predict sentence length is the same for both tribunals. We examined bivariate correlations between sentence length and the predictors separately for each tribunal, in order to assess whether correlations are indeed comparable across tribunals, using Spearman correlation coefficients because the distribution of sentence length is skewed.² However, because this method investigates only similarity in bivariate associations across the tribunals, we also included in our regression models a number of interaction terms of type of tribunal with specific independent variables to investigate whether this conclusion also holds multivariately.³

Results

All 111 cases of individuals sentenced by the ICTY and the ICTR were included in a multiple regression analysis investigating the impact of 12 independent variables on sentence length.⁴ The results of the analysis are given in Table 2.

Six of the tested legal factors contribute significantly to the prediction of sentence length, that is, conviction for genocide, rank of the offender, number of guilty counts, conviction for crimes against humanity, and number of aggravating and mitigating factors. The final model predicts 72 percent of the variability in sentence length.

Table 2. Multiple regression predicting sentence length at the ICTY and ICTR ($N = 111$)

| | B | β | Significance |
|--|------|---------|--------------|
| Conviction for genocide | 16.1 | .45 | .000 |
| Rank (dummy coded – low rank as reference) | | | |
| High rank | 14.5 | .39 | .000 |
| Middle rank | 4.4 | .13 | .069 |
| Conviction for crimes against humanity | 8.1 | .20 | .001 |
| Number of counts | 0.8 | .20 | .002 |
| Number of aggravating factors | 1.5 | .19 | .001 |
| Number of mitigating factors | -0.9 | -.18 | .005 |
| Conviction as hands-on perpetrator | 4.2 | .13 | .077 |
| Conviction as planner, instigator, order-giver | 3.6 | .10 | .122 |
| Conviction as aider and abettor | 2.3 | .07 | .239 |
| Conviction for war crimes | 2.3 | .07 | .344 |
| Conviction as superior | 2.3 | .06 | .308 |
| Conviction as participant in JCE | -0.4 | -.01 | .882 |
| Intercept | -4.5 | | .503 |

Notes: Adj $R^2 = .723$; SE = 8.8168; F test (13, 97) = 23.098, $p < .0001$.

The strongest predictor, based on the standardized coefficients, is a conviction for genocide. When a defendant is convicted of genocide, his/her sentence is increased by 16 years and 1 month compared with those not convicted of genocide, all other things being equal. In contrast, a conviction for crimes against humanity results, *ceteris paribus*, in an increase of 8 years and 1 month in imprisonment. We cannot interpret the relative addition of each category of crimes to sentence length because the categories were entered as separate variables.⁵ However, additional analyses indicate that, compared with individuals convicted solely of war crimes, those convicted of crimes against humanity receive 5.3 years extra and those convicted of genocide are sentenced to an extra 20.1 years, controlling for all the other variables. On the basis of our analyses, it can be argued that, in the tribunals' sentencing practice, genocide is actually treated as 'the crime of crimes' and the most serious international crime if other legally relevant factors are taken into account.⁶ In practice, the sentences handed down by the ICTY and the ICTR are ordered hierarchically – with genocide at the top of the seriousness scale followed by crimes against humanity. A conviction for war crimes did not add significantly to the prediction of sentence length.

The second-strongest predictor of sentence length is the rank of the offender. When a defendant occupies a position at the top level of a state political or military hierarchy, his/her sentence is 14 years and 6 months longer, given all other factors, than that for a low-ranking offender. Consequently, the highest authority figures receive more severe punishment than their followers – those without any influence in the overall conflict. Since the coefficient of middle rank just misses statistical significance, it seems that there is no statistically significant decrease in sentence severity between middle- and low-ranking offenders. It might be that, in case of low-ranking defendants, considerations of the role

of the offender in the overall conflict are overshadowed by considerations of the cruelty and depravity of the committed crimes and the zeal of their executioners. In contrast to their direct, middle-ranking, superiors, the low-ranking offenders are in the majority of cases the actual perpetrators committing brutal atrocities, and judges apparently assign them equal or even more blame than their direct superiors (Ewald, 2010: 398; Holá et al., 2009: 90).

The number of guilty counts stands out as the next important predictor of sentence length. Each guilty count results in nearly 10 extra months in prison, controlling for the other variables. Thus, those involved in more extensive criminal activities are punished more severely than those convicted of separate incidents.

The numbers of aggravating and mitigating factors are the final significant predictors of sentence length in our model. On average, each factor in mitigation reduces sentence length by 11 months and each aggravating factor increases sentence length by an extra 1 year and 6 months imprisonment, given all the other predictors.⁷

The regression model described above assumed that the mechanism of sentence determination was identical at both tribunals. In order to test this assumption we (1) compared bivariate correlations across tribunals; and (2) added interaction terms to our regression model. No significant differences in bivariate associations of sentence length and tested legal factors emerged, except for the association between sentence length and conviction under superior responsibility. In this case, the correlation coefficients differ in strength and in the direction of association between the tribunals. At the ICTY, sentence length is negatively related to a conviction as a superior and is of medium strength ($\rho = -.32$); at the ICTR, the two are only weakly positively associated ($\rho = .24$). The negative association between sentence length and a conviction as a superior is in accordance with the judges' argumentation discussed above. Superior responsibility is based on a superior's omitting to prevent or punish the crimes of his/her subordinates and, therefore, entails lower sentences compared with other modes of liability representing more active participation in crime. The difference between the tribunals in correlation coefficients might be influenced by the fact that superior responsibility has arguably not played such a prominent role at the ICTR as at the ICTY and many defendants were convicted of a combination of superior responsibility with other more active modes of liability, such as perpetration.⁸

Although we established that almost no differences exist in bivariate associations, such a difference might still emerge in a multivariate model. In order to test this, we ran several multiple regression models with interaction terms added (Cohen et al., 2003). As a starting point, we used the model with all significant predictors of sentence length resulting from the regression analysis. Then we separately (one by one) added an interaction term between each significant predictor and the type of tribunal. Given the sample size, not all interaction terms could be added at the same time. In general, our additional analyses indicated that there appear to be no significant differences between the tribunals in the impact of each predictor on sentence length. Only for the number of counts does the effect on sentence length appear to differ depending on the tribunal.⁹ It seems that at the ICTR each additional count results in a larger sentence increase than at the ICTY. This difference is likely related to the fact that ICTR defendants are on average convicted of a lower number of counts compared with ICTY defendants, yet are nonetheless sentenced to severe sentences. At the ICTR, the crime base (factual background) of these

guilty counts generally covers a much broader array of incidents that are usually the most serious violations of international criminal law (such as multiple murders of hundreds of victims) and is thus qualitatively different from a guilty count covering only one incident, often of a less serious nature (such as mistreatment of a victim), as occurs at the ICTY. Formulated differently, at the ICTY, sentence length is more often determined by the accumulation of more counts of a less serious nature. At the ICTR, a smaller number of counts of more serious crimes give rise to lengthier sentences.

Discussion and conclusions

The ICTY and ICTR sentencing practices have been subjected to a lot of criticism. Academics have repeatedly denounced the tribunals' sentencing and labelled it as inconsistent and erratic. However, this criticism was not founded on a comprehensive empirical analysis of the tribunals' sentencing practice. The current study has tried to address this issue by examining the sentences handed down by the ICTY and the ICTR. The primary aim of this empirical enquiry was to explain the dynamics of a number of prominent legal factors behind sentence determination and to evaluate the consistency of international sentencing practice.

The sentencing practice of the ad hoc tribunals emerges as consistent. Regarding consistency in approach, the sentencing jurisprudence appears similar, with judges of both tribunals cross-referencing each other's case law. Consequently, the case law of the ICTY and the ICTR has developed fairly similarly and a set of general sentencing principles is consistently referred to in the case law, such as the primary importance of the gravity of the crime for sentence severity or the principle of gradation of sentence relative to the seniority of a defendant. For a detailed discussion, see Hola (2012: 26–83).

Regarding consistency in outcome, the regression model accounted for 70 percent of the variation in sentence length at the ICTY and the ICTR. It is very difficult to judge the adequacy of this figure, both theoretically and empirically. Theoretically, sentencing should never be 100 percent predictable. In order to be fair, sentences should be not only predictable but also individualized, reflecting the particularities of each case. Therefore, some variation should always remain unexplained to allow for individualization. The explained variation that emerged in our study is comparable to that found in similar quantitative research into sentencing consistency at a domestic level (Anderson and Spohn, 2010: 389; Engen and Gaaney, 2000: 1209). In many countries, concerns over unwarranted disparities in sentencing have led to reforms of sentencing law, introducing, for example, determinate sentencing law or numerical sentencing guidelines that limit the discretionary powers of domestic judges. To assess the success of these reforms, studies evaluating sentencing predictability at a domestic level have been conducted, showing that legally relevant factors are the most important determinants of sentences, explaining as much as 70 percent of the variation in imprisonment length (Hofer et al., 1999). Consequently, it seems that, at the international level, despite the lack of sentencing guidelines and the large discretionary powers of judges, the sentences are as statistically predictable as sentences in domestic legal systems with more detailed legal regulation of sentence determination.

We also analysed the overall consistency of international sentencing practice ('systemic consistency'), examining possible differences across the two international criminal tribunals. It turned out that in this regard too international sentencing emerges as consistent. Looking at the set of empirically most important predictors of sentence length, there are no major differences in the mechanisms behind sentence determination at both tribunals. Empirically consistent (and legally relevant) patterns have emerged in the sentencing practice of both tribunals: those convicted of genocide are subjected to more severe sentences than those convicted of crimes against humanity; high-ranking defendants are sentenced more harshly than low-ranking individuals; more severe sentences are handed out to defendants convicted of multiple crimes; and mitigating and aggravating factors account for sentence mitigation/aggravation.

In the end, we have to address the limitations of our study. As already noted, a vast array of legally relevant factors are at play and it is in practice impossible to examine them all simultaneously in a multivariate analysis with a dataset that is small in a statistical sense – even though it comprises the entire 'population' of decided cases. It would be desirable if a more detailed analysis could be performed and differences in sentence length between, for example, individual underlying offences (such as a conviction for murder compared with a conviction for torture) or individual aggravating and mitigating factors examined. Although our study is one of the first to examine jointly the ICTY and ICTR sentencing practice, the dataset is by necessity unbalanced. More cases have been decided by the ICTY ($N = 71$) than by the ICTR ($N = 40$). Thus, ICTY patterns in a sense will play a larger role in the regression model. It should also be noted that the characteristics of defendants included in our sample do not fully represent the whole population of individuals to be (potentially) convicted by the ICTY and the ICTR. The cases still pending before the tribunals' Trial Chambers represent more complicated and arguably also more serious cases, in the sense of the seniority of the accused and in relation to the crime base/crime gravity, compared with the already finished cases.

Furthermore, there are observable differences between the two tribunals in the composition of cases and the distribution of sentences. The ICTR sentences tend to be much longer than the sentences pronounced by the ICTY. Arguably, the lengthier ICTR sentences are related to the different composition of cases dealt with by the two tribunals. At the ICTR, all defendants have been tried on genocide charges, the majority are subsequently convicted of genocidal killings, and many of them are sentenced to life imprisonment. This is not the case at the ICTY, where charges are more varied, convictions for crimes against humanity/war crimes are prevalent, and life imprisonment is an exceptional sentence. Consequently, qualitatively different crimes are dealt with and quantitatively different sentences are handed down at the two tribunals. These issues – that is, different N , qualitatively different crimes, quantitatively different sentences – make a quantitative analysis of the tribunals' sentencing practices and interpretation of the results complex. Even so, it does appear that the same legal factors are accorded similar importance at both tribunals.

This research constitutes one of the first studies comprehensively testing the international sentencing practice represented by the two ad hoc international criminal tribunals – the ICTY and the ICTR. Despite inherent limitations, we have shown that legally relevant

patterns are evident in international sentencing practice. In short, international sentencing practice is consistent in approach, in outcome and across tribunals. Since the ICTY and the ICTR are still functioning tribunals, pronouncing new verdicts and sentences, it will be possible to carry out a more detailed empirical analysis of international sentencing in the future with more cases. Such research could also include future sentences pronounced by other international courts, such as the International Criminal Court, or by domestic courts dealing with international crimes, in order to assess whether there has indeed emerged one 'common' system of international sentencing.

Notes

1. The article excludes other courts with international elements dealing with international crimes (internationalized courts), such as the Special Court for Sierra Leone or the Special Panels for Serious Crimes in East Timor, for two reasons: (i) theoretically, it is arguable that all the internationalized courts have a much stronger 'national element' and it is thus difficult to make direct comparisons; and (ii) methodologically/pragmatically, it is not possible to make any sensible comparisons because either there have not been so many cases completed by these courts or the composition of convictions is not as varied as in the ICTY/ICTR. For more detailed discussion, see Hola (2012: 8–9).
2. Using Blalock's formula (Blalock, 1972), we tested for any differences between the tribunals.
3. To make sure that the solution is robust against our choice to code life sentences as '55 years', the model was tested; (1) with life sentences recoded to 55 years; (2) with life sentences recoded to 45 and 65 years; and finally (3) with no life sentences included. The set of statistically significant predictors remained the same and the regression coefficients were comparable.
4. The first block included the following 11 variables: genocide, crimes against humanity, war crimes, number of counts, number of aggravating factors, number of mitigating factors, perpetration, aiding/abetting, ordering/instigating/planning, superior responsibility, and participating in JCE. Rank was added to the model in a second block as a dummy coded variable (high, middle; low as a reference category). This led to a statistically significant improvement in the model (F change (2,97) = 16,246, $p < .0001$).
5. This modelling strategy had been chosen to assess whether the categories of crimes are hierarchically ordered in practice. Given that that turned out to be the case, we are able to construct a new categorical variable where each defendant is categorized according to the most serious crime he/she was convicted of. Next, we were able to run an additional analysis with category of crimes coded as a dummy variable, with the category leading to the shortest sentences coded as the reference category.
6. It should also be noted that almost all convictions for genocide have been at the ICTR. At the time of writing, only one defendant has been convicted of genocide at the ICTY (Radislav Krstic).
7. None of the other factors included in the analysis add significantly to the prediction of sentence length. All the statistical assumptions for multiple regression were met.
8. For a more detailed discussion, see Hola (2012: 190).
9. The analysis of interaction terms was, however, technically complicated by multicollinearity. The problem remained even after centring the variables. In order to double-check for possible interactions, we ran several different models and the only interaction term that remained significant was the interaction of number of counts with tribunal. For a more detailed discussion of the methodology, see Hola (2012: 194–6).

References

- Anderson AL and Spohn C (2010) Lawlessness in the federal sentencing process: A test for uniformity and consistency in sentence outcome. *Justice Quarterly* 27(3): 362–393.
- Blalock H (1972) *Social Statistics*, 2nd edn. New York: McGraw-Hill.
- Carcano A (2002) Sentencing and the gravity of the offence in international criminal law. *International and Comparative Law Quarterly* 51(3): 583–609.
- Cohen J, Cohen P, West SG and Aiken LS (2003) *Applied Multiple Regression/Correlation for the Behavioral Sciences*. New Jersey: Lawrence Erlbaum Associates.
- Danner AM (2001) Constructing a hierarchy of crimes in international criminal law sentencing. *Virginia Law Review* 87: 415–501.
- D'Ascoli S (2011) *Sentencing in International Criminal Law, the UN Ad Hoc Tribunals and Future Perspectives for the ICC*. Oxford and Portland, OR: Hart Publishing.
- Del Ponte C (2004) Prosecuting the individuals bearing the highest level of responsibility. *Journal of International Criminal Justice* 2: 516–519.
- De Roca EM (2008) Sentencing and incarceration in the ad hoc tribunals. *Stanford Journal of International Law* 44: 1–62.
- Doherty JW and Steinberg RH (2009) An empirical study of ICTY and ICTR sentencing: Doctrine versus practice. CELS 2009 4th Annual Conference on Empirical Legal Studies Paper, 3 August. URL (accessed 12 June 2012): <http://ssrn.com/paper=1443468>.
- Drumbl MA (2007) *Atrocity, Punishment and International Law*. New York: Cambridge University Press.
- Engen RL and Gainey RR (2000) Modelling the effects of legally relevant and extralegal factors under sentencing guidelines: The rules have changed. *Criminology* 38(4): 1207–1230.
- Ewald U (2010) 'Predictably irrational' – International sentencing and its discourse against the backdrop of preliminary empirical findings on ICTY sentencing practices. *International Criminal Law Review* 10: 365–402.
- Hofer PJ, Blackwell KR and Ruback RB (1999) The effect of the federal sentencing guidelines on inter-judge sentencing disparity. *Journal of Criminal Law & Criminology* 90(1): 239–306.
- Hola B (2012) *International Sentencing, A Game of Russian Roulette or Consistent Practice*. Oisterwijk: Uitgeverij BOXPress.
- Hola B, Smeulers A and Bijleveld C (2009) Is ICTY sentencing predictable? An empirical analysis of ICTY sentencing practice. *Leiden Journal of International Law* 22(01): 79–97.
- Meernik J (2003) Victor's justice or the law? Judging and punishing at the International Criminal Tribunal for the Former Yugoslavia. *Journal of Conflict Resolution* 47(2): 140–162.
- Meernik J and King KL (2003) The sentencing determinants of the International Criminal Tribunal for the Former Yugoslavia: An empirical and doctrinal analysis. *Leiden Journal of International Law* 16(04): 717–750.
- Meernik J, King KL and Dancy G (2005) Judicial decision making and international tribunals: Assessing the impact of individual, national, and international factors. *Social Science Quarterly* 86: 683–703.
- Mettraux G (2005) *International Crimes and the Ad Hoc Tribunals*. New York: Oxford University Press.
- Ohlin JD (2007) Three conceptual problems with the doctrine of joint criminal enterprise. *Journal of International Criminal Justice* 5: 69–90.
- Olusanya OA (2004) Do crimes against humanity deserve a higher sentence than war crimes? *International Criminal Law Review* 4: 431–473.
- Olusanya OA (2005) *Sentencing War Crimes and Crimes against Humanity under the International Criminal Tribunal for the Former Yugoslavia*. Groningen: Europa Law Publishing.

- Smeulers A and Hola B (2010) ICTY and the culpability of different types of perpetrators. In Smeulers A (ed.) *Collective Violence and International Criminal Justice: An Interdisciplinary Approach*. Antwerp–Oxford–Portland: Intersentia, 175–205.
- Van der Wilt H (2008) Equal standards? On the dialectics between national jurisdictions and the International Criminal Court. *International Criminal Law Review* 8: 229–272.
- Van Sliedregt E (2003) *The Criminal Responsibility of Individuals for Violations of International Humanitarian Law*. The Netherlands: T.M.C. Asser Press.

Legal texts

- UN (1993) Statute of the International Tribunal for the Former Yugoslavia. Security Council Resolution 827, UN Doc. S/RES/827, 25 May.
- UN (1994) Statute of the International Tribunal for Rwanda. Security Council Resolution 955, UN Doc. S/RES/955, 8 November.
- ICTY Rules (1994) Rules of Procedure and Evidence of the International Criminal Tribunal for the former Yugoslavia. UN Doc. IT/32/Rev. 45, adopted 11 February 1994 (as amended 8 December 2010).
- ICTR Rules (1995) Rules of Procedure and Evidence of the International Criminal Tribunal for Rwanda. UN Doc. ITR/3/REV.1, adopted 29 June 1995 (as amended 1 October 2009).

Cited case law (in chronological order)

- Prosecutor v. Tadic* (14 July 1997), Case No. IT-94-1-T, Sentencing Judgment.
- Prosecutor v. Mucic et al.* (16 November 1998), Case No. IT-96-21-T, Judgment.
- Prosecutor v. Serushago* (5 February 1999) Case No. ICTR-98-39, Judgment.
- Prosecutor v. Mucic et al.* (20 February 2001), Case No. IT-96-21-A, Judgment.
- Prosecutor v. Musema* (16 November 2001), Case No. ICTR-96-13-A, Judgment.
- Prosecutor v. Naletilic & Martinovic* (31 March 2003), Case No. IT-98-34-T, Judgment.
- Prosecutor v. Mucic et al.* (8 April 2003), Case No. IT-96-21-A, Judgment.
- Prosecutor v. Cesic* (11 March 2004), Case No. IT-95-10/1-T, Judgment.
- Prosecutor v. Nindabahizi* (15 July 2004), Case No. ICTR-01-71-T, Judgment.
- Prosecutor v. Semanza* (20 May 2005), Case No. ICTR-97-20-A, Judgment.
- Prosecutor v. Kajelijeli* (23 May 2005), Case No. ICTR-98-44A, Judgment.
- Prosecutor v. Hadzihasanovic & Kubura* (15 March 2006), Case No. IT-01-47-T, Judgment.
- Prosecutor v. Nindabahizi* (16 January 2007), Case No. ICTR-01-71-A, Judgment.
- Prosecutor v. Milosevic* (12 December 2007), Case No. IT-98-29/1-T, Judgment.
- Prosecutor v. Haradinaj et al.* (3 April 2008), Case No. IT-04-84-T, Judgment.
- Prosecutor v. Delic* (15 September 2008), Case No. IT-04-83-T, Judgment.
- Prosecutor v. Milutinovic et al.* (26 February 2009), Case No. IT-05-87-T, Judgment.
- Prosecutor v. Mrksic & Sljivancanin* (5 May 2009), Case No. IT-95-13/1-A, Judgment.
- Prosecutor v. Lukic & Lukic* (20 July 2009), Case No. IT-98-32/1-T, Judgment.