INTRODUCTION
1. GENERAL INTRODUCTION

“I saw a number of dogs, maybe eight or nine, German Shepherds, walking around with soldiers amongst the people...They [Serb soldiers] were taking some men away, saying that they wanted to have a chat with them. Maybe they were their neighbors. I don’t know. So they kept walking around a little bit, and they took them away. But in the evening, they started taking people away in greater numbers. At one point, people started to scream, and everybody stood up. Afterwards, we heard rumors that a woman had given birth. And after a few minutes, shortly after that, we heard the same thing again. So we were thinking, “What’s going on here?” But actually it was [Serb soldiers] who kept coming and taking men away. ... Women were screaming, everybody screamed. And a little later, maybe 15 minutes later, you could hear people screaming and moaning from the outside. Yes, they were moaning. Occasionally you would hear a shot, but sometimes not, and then everything would be silent for a while, and then nothing. ... It lasted all night. And there were also crazy women, women who had gone round the bend there out of fear there, those whose nerves were not as strong and poor health. And I heard, I didn’t see, but I heard that there were also people who hanged themselves out of fear.”... Witness XXX was then taken on a bus with other men and detained on the bus in Bratunac. After spending two days detained in a school in Bratunac, having witnessed severe mistreatment of fellow detainees and hearing screams of men taken outside, followed by shots and then silence, Witness XXX was once again placed on a bus and taken to another school in Pilica. During this entire period almost no food or water was provided to the detainees, and beatings and other forms of mistreatment were common... According to Witness XXX, on the second day, the men were told that they would be going to Tuzla. Shortly thereafter, Serb soldiers brought in sheets to be torn up and used to tie the detainees’ hands: “I offered my hands to be tied, but I was surprised, why they were tying us if we were allegedly going to freedom, to Tuzla?” Instead of being taken to freedom, Witness XXX and the other Bosnian Muslim detainees were taken by bus to a hilltop where they heard gunfire and voices. Witness XXX described what he saw next, “I watched as the column of men goes down a path, and I watched to where the dead ones are. I got there and I heard Serb troops cursing, making noise. Bursts of fire simply mowed them down. They all fell to the ground.” Witness XXX was then forced to walk down that same path and to stop between the rows of dead bodies. With their backs turned to their executioners, the men in the column were shot and the men began falling to the ground. Witness XXX also fell and laid among the dead bodies of his fellow detainees as
column after column of men were brought to the spot and executed. When Witness XXX finally was able to stand up and look around, he saw an estimated 1,000 to 1,500 dead people around him. "

The preceding paragraphs are extracts of a testimony of one of the survivors of the so-called Srebrenica massacre - one of the most notorious episodes of the conflict at the territory of the former Yugoslavia. In July 1995 during the take-over of the UN ‘safe area’ of Srebrenica over 7000 persons (mostly Bosnian men and boys) were executed by the Bosnian Serb forces. It was described as the worst crime on the European soil since WW II and classified as genocide – the only act of genocide occurring in Europe since 1945. The toll of human suffering, desperation and the sense of impotence of victims and survivors at Srebrenica could hardly be put into words. It is something unimaginable and hard to grasp. However, it is just one example of international crimes committed during the conflict in the former Yugoslavia in the early 1990s (though arguably one of the most serious) and the crimes committed in the Former Yugoslavia are just a drop in the sea of international crimes committed throughout the world in the 20th century.

International crimes, such as genocide, crimes against humanity and war crimes, are manifestations of large-scale and serious violations of human rights that have been defined as the most serious crimes of international concern - unimaginable atrocities that deeply shock the conscience of humanity. The twentieth century witnessed atrocity on a scale beyond anything chronicled in history: one million killed in the Armenian genocide; tens of millions slaughtered by Stalin's Soviet regime; the calculated extermination of millions of people by Hitler's Nazis; mass killing, torture, and systematic rape in Cambodia, the former Yugoslavia, Rwanda, egregious violence in Sudan or Congo – to name just a few examples. It has been estimated that

---

in the last century approximately 191 million persons lost their lives due to collective violence.\textsuperscript{7} Not only because of their massive and horrendous character, but due to their specific features being the manifestations of collective and structural violence often committed on behalf of a state or with its acquiescence, it is a well-accepted view that international crimes are to be distinguished from ordinary crimes.\textsuperscript{8}

Given the unique features and seriousness of international crimes, individual states as well as the international community have established various mechanisms to address these acts and hold the perpetrators legally accountable. At the international level the International Criminal Tribunal for the Former Yugoslavia (ICTY) and the International Criminal Tribunal for Rwanda (ICTR) can be viewed as primary examples of these efforts. These ad-hoc international criminal tribunals have been established by the United Nations Security Council acting on the basis of Article 39 of the UN Charter as responses to widespread and flagrant human rights abuses perpetrated during the conflicts taking place in the early 1990s at the territory of the Former Yugoslavia and Rwanda, respectively.

The ICTY and ICTR are by far not the only institutions established at the international level to apply international criminal law and punish perpetrators of international crimes. The scholars date the history of international criminal law centuries back, however, the first time an international tribunal held an individual responsible for international crimes came only after the WWII.\textsuperscript{9} In 1945, the International Military Tribunals (IMT) in Nuremberg and Tokyo were established to punish the major Axis war criminals. Due to the stalemate at the international arena during the period of Cold War, there were no similar institutions created, however, there were some ongoing efforts to set up an international criminal court.\textsuperscript{10} After the fall of the Berlin wall ending the period of Cold War, the picture changed significantly. Media coverage of stories of concentration camps in Bosnia, showing pictures which evoked memories of the Holocaust,
and of the wholesale slaughter committed in Rwanda in the early 1990s despite the presence of the UN peacekeepers caused public outcry and led to demands that something should be done about the situation. The UN Security Council reacted by the establishment of the ICTY and ICTR – the first ad-hoc international criminal tribunals.\(^\text{11}\) Shortly thereafter, the permanent International Criminal Court was set up at the Rome Conference in 1998. The increased interest in achieving justice and thus preventing such criminality in future led to the establishment of many other, the so-called- internationalized courts and tribunals such as for Sierra Leone, Kosovo, East Timor, Cambodia, Bosnia & Herzegovina, Iraq or Lebanon.\(^\text{12}\) There are many differences among these institutions - in the applicable law, their jurisdiction or composition. What connects them together, however, is the fact that they are to deal with international crimes such as genocide, crimes against humanity or war crimes.

Therefore, after the fall of the Berlin Wall in 1989 there has been a boom in the field of international criminal justice and new unprecedented institutions were established at the international arena and have been convicting and sentencing perpetrators of international crimes.\(^\text{13}\) However, in the early 1990s there was no ‘international criminal code’ or extensive precedents save for the case law of the IMTs in Nuremberg and Tokyo. The legal framework was very sketchy\(^\text{14}\) and judges at these courts and tribunals had to rely on customary


\(^{14}\) Usually the documents establishing these tribunals contain only very general rules indicating in broad terms limits of their respective jurisdictions. No detailed legal guidance is provided. The only exception is the Rome Statute of the ICC where a rather elaborate legal framework for the functioning of the Court has been set up. Next to the documents establishing various courts and tribunals, a number of international treaties relevant to issues of international crimes (especially substantive matters relating to e.g. crime definitions) was adopted after the WW II and during the Cold War such as Convention on the Prevention and Punishment of the Crime of Genocide adopted on 9 December 1948, in force as of 12 January 1951, four Geneva Conventions adopted on 12 August 1949 and their Additional Protocols I and II adopted on 12 December 1977, Convention against Torture and Other Cruel, Inhuman and Degrading Treatment or Punishment, adopted on 10 December 1984 by the UN GA, in force as of 26 June 1986 etc.
international law or general principles of international law to delimit their jurisdiction, including definitions of crimes, modes of liability or issues of punishment and sentencing.\textsuperscript{15}

The main aim of this thesis is to empirically analyze the international sentencing practice. How do we punish perpetrators of the most serious crimes? As noted by Mettraux it would be hard to identify crimes more difficult to sentence than international crimes.\textsuperscript{16} This is not only due to the atrocity type character of international crimes that are usually entailing horrendous acts committed on a large scale but also due to their collective, systematic nature, often involving state authorities. International crimes are usually offences entailing a large number of victims, committed as part of a large scale campaign and implicating many individuals with very different tasks. These crimes are often state sanctioned and characterized by mass involvement of both military functionaries as well as civilians.\textsuperscript{17} Criminologists have labelled international crimes as ‘crimes of obedience’\textsuperscript{18} and argued that international crimes are a different type of criminality than ordinary crimes. They are crimes of conformism rather than crimes of deviance. In times of collective violence many otherwise law-abiding citizens get involved and commit extremely cruel acts simply by conforming to their environment.\textsuperscript{19} As a direct result of all these factors, i.e. extreme gravity, systematic character and massive participation, it is extremely complex to assign and, for the purposes of sentencing, to ‘measure’ culpability/blameworthiness of individual perpetrators.

This thesis analyzes how this extremely difficult task has been approached by international tribunals. By using statistical methods it tries to disentangle the phenomenon of sentencing international crimes. Ideally an all-encompassing study of international sentencing would include sentences handed down by all the international courts and tribunals established so far, including the Nuremberg and Tokyo Military Tribunals, the ad hoc tribunals and many

\textsuperscript{16} Mettraux., Ibid., at 343.
internationalized jurisdictions. For theoretical and pragmatic considerations, this thesis concentrates only on the sentencing practice of the first two ad hoc international criminal tribunals – the ICTY and ICTR. Both Tribunals have been established in a similar way (by the UN Security Council), they are very similar in their institutional design and function under a very similar applicable legal framework. So far, they have produced the most extensive sentencing practice of all the international(ized) criminal courts and tribunals. They could be perceived as the pioneers of international criminal law and their case law is often referred to and discussed by the other courts and tribunals.²⁰ Many of the institutions listed above such as the International Criminal Court or the Special Tribunal for Lebanon have not handed down any sentences yet. Otherwise, this research excludes sentencing practice of the other international courts and tribunals for four principle reasons:

1) Time lapse, i.e. the IMTs in Nuremberg and Tokyo took place immediately in the aftermath of the WWII under very different circumstances. At the time, the modern system of international law, human rights protection and international criminal law were at their inception. Since then, there have been huge developments in all these fields that have arguably influenced the current system of international criminal justice as it now stands; one example thereof is the abolishment of death penalty at the international level. In contrast, judges in Nuremberg and Tokyo sentenced 11 and 7 accused to death, respectively;

2) Lack of data, i.e. in many cases the number of sentenced individuals is rather limited and it is therefore, difficult to quantitatively analyze trends in the sentencing practice (e.g. at Nuremberg 21 and in Tokyo 25 sentences were rendered; the Special Court for Sierra Leone (“SCSL”) convicted and sentenced only 8 defendants). Furthermore, statistically, adding more legal institutions (more data) would give more power to analyses, however it also adds heterogeneity and thus increases risks for confounding results;

3) Heterogeneity, i.e. there are qualitative differences in jurisdiction, composition and/or applicable law at the internationalized courts, such as the SCSL or the Special Panels for Serious Crimes in East Timor (“SPSC”), compared to the ICTY and ICTR. Arguably all the internationalized courts have a much stronger “national element” (hybrid jurisdiction over international and domestic crimes, mixed composition with international and domestic judges,

prosecutors or defence attorneys; applicability of international and domestic law). Therefore, it is difficult to compare directly the sentencing regimes and practice across these Tribunals. In this respect, the ICTY and ICTR can be perceived of as the only “purely international criminal tribunals”; and

4) Qualitative differences in composition of cases, i.e. at some of these internationalized tribunals, that have already handed down some sentences, the composition of convictions is not so much varied as in the ICTY/ICTR and it is thus difficult to make comparisons within a tribunals and across institutions (e.g. SPSC have convicted over 80 individuals but the vast majority of defendants are very low ranking perpetrators).

Consequently, this thesis is limited to the examination of the sentencing practice of the ICTY and ICTR as the representatives of international sentencing and international criminal justice system. The Tribunals’ sentencing practice has been subjected to heavy criticism in the legal discourse. However, this criticism has never been based on a thorough empirical enquiry into the Tribunals sentencing practice. In contrast to the domestic sentencing research where there is an abundance of studies dealing with ‘the reality of sentencing’, the empirical research on international sentencing is only in its ‘embryonic stage’. Given the fact that the re-instituted international criminal justice system was reinvented only about 20 years ago, this fact is not so surprising. There have been only very few empirical studies dealing with international sentencing and there is a clear information gap in this respect. This is all the more serious because the sentencing legal framework at the Tribunals is very general and judges are provided with large discretionary powers when it comes to sentencing. It is therefore not clear what factors influence sentence determination. In order to disentangle the multiplicity of factors and considerations that play a role in meting out sentences at the ICTY and ICTR, it is necessary to empirically examine the Tribunals sentencing practice. This thesis attempts to fill in this gap by trying to find empirically substantiated answers in relation to the consistency of sentencing at the two ad hoc international criminal tribunals.


The Statutes and Rules of Procedure and Evidence, the written law establishing the Tribunals and governing their functioning, are very vague and contain only very general guidelines with respect to sentencing. The law does not pronounce any sentencing rationales or general principles of sentence determination. It does not contain any penalty structure in the sense of minimum or maximum penalties for individual offences. The crimes under the Tribunals’ jurisdictions are not distinguished on the basis of their inherent gravity and corresponding sentence severity. No list of relevant aggravating and mitigating factors is provided.
Consequently, the central research question of this manuscript is:

*Is international sentencing as practiced by the ICTY and ICTR consistent?*

Sentencing at the ICTY and ICTR is to a very large extent left to the discretion of the judges. This begs the question as to how the judges, in the application of justice, have used this discretion and in so doing, established a predictable and consistent sentencing practice. This is the main question to be investigated by this thesis.

2. **Consistency of Sentencing**

The requirement of consistency of sentencing could be perceived as one of the fundamental principles of any sentencing practice. It stems from the modern internationally recognized human rights principles and has been explicitly endorsed by the ICTY and ICTR case-law. From the inception of both Tribunals it has been emphasised that despite the condemnable and horrendous character of the atrocities committed in Yugoslavia and Rwanda, proceedings before the ICTY and ICTR should, above all, respect the principles of fair trial.  

The right to a fair trial forms one of the cornerstones of the current system of international human rights protection. The principles embodied in the right to a fair trial and other human rights protected by international law call for consistency in sentencing practice. Next to the right to a fair trial with its general requirement of fairness of proceedings, the principle of legality embodied in the phrases *nullum crimen sine lege* and *nulla poena sine lege* and the principle of equality

---

23 It was presumed that only fair proceedings and sentencing could lead to the attainment of the Tribunals’ main objectives as proclaimed in the constitutive Resolutions. See UN Doc S/RES 827 (1993) ; UN Doc S/RES 955 (1994).


before the law\textsuperscript{28} are among the main human rights principles underlying the requirement of consistency in sentencing.

The need for consistent and predictable sentencing has also been recognized by the ICTY itself. In the \textit{Celebici} case the ICTY Appeals Chamber noted: “One of the fundamental elements in any rational and fair system of criminal justice is consistency in punishment. [...] [P]ublic confidence may be eroded if ... [courts] give an appearance of injustice by permitting substantial inconsistencies in the punishment of different offenders, where the circumstances of the different offences and of the offenders being punished are sufficiently similar that the punishments imposed would, in justice, be expected to be also generally similar”\textsuperscript{29}. In the ensuing cases the judges often discussed the requirement of consistent sentencing and it is often repeated that “sentences of like individuals in like cases should be consistent”\textsuperscript{30}. Accordingly, the Appeals Chamber stated that “a sentence may be considered “capricious or excessive if it is out of reasonable proportion with a line of sentences passed in similar circumstances for the same offences”\textsuperscript{31}. It is, however, also emphasised that “the relevance of previous sentences is [...] often limited as a number of elements, relating, \textit{inter alia}, to the number, type and gravity of the crimes committed, the personal circumstances of the convicted person and the presence of mitigating and aggravating circumstances, dictate different results in different cases such that it is frequently impossible to transpose the sentence in one case \textit{mutatis mutandis} to another. This follows from the principle that the determination of the sentence involves the individualisation of the sentence...”\textsuperscript{32}. Consequently, the individualization of sentences is another important sentencing principle that is particularly emphasised in the ICTY and ICTR sentencing practice. Both principles of consistency and of individualization should be respected and it is necessary to strike the right balance between the two for sentencing to be considered fair and legitimate.

This thesis focuses on the concept of consistency of ‘international sentencing’. Theoretically, there are three basic dimensions of the concept of consistency that will be addressed in this thesis: (i) consistency in approach, (ii) consistency in outcome and (iii) systemic consistency.

\textsuperscript{28} The principle of equality generally entails that people in similar situations should be treated equally. Therefore, with respect to sentencing, while imposing a penalty judges should weigh similar relevant factors in similar situations similarly. Rhona K. M. Smith, Textbook on International Human Rights (Blackstone Press. 2003). at 248.

\textsuperscript{29} Judgment, Mucic et al. (IT-96-21-A) Appeals Chamber, 20 February 2001, para. 756.


\textsuperscript{31} Kyocka et al., Ibid., para. 681.

a) Consistency in Approach

First, all sentences should follow the same underlying principles, thus creating a coherent and harmonious system – “consistency in approach”. In this respect, it is relevant whether the ICTY and ICTR judges have developed a consistent narrative, wherein legal factors relevant to sentence determination are emphasised and their relationship to sentence severity discussed in the case-law. Consistency in approach requires that sentences considered appropriate be expressly related in a principled way to sentencing rationales and case facts. Therefore, the sentencing discretion should be exercised in a principled manner - there should be a coherent judicial approach to the exercise of discretion in sentencing; all decisions should be based on common standards – general underlying principles – that are uniformly applied to the facts of each case. The first sub-question of this thesis is thus: Is it possible to identify such general principles of sentence determination and resultant legally relevant sentencing factors that are consistently referred to in the ICTY and ICTR judges’ narrative?

b) Consistency in Outcome

Second, sentences should be consistent to the degree that individual offenders are placed into distinctive groups on the basis of these “legitimate” (legally relevant) set of characteristics/case attributes, i.e. similar offenders (with similar case attributes) receive similar sentences – “consistency in outcome”, predictability of sentences. In this respect, the second sub-question of this thesis examines whether sentence outcomes follow in a predictable manner from the combinations of examined legal factors. In statistical terms, do the legal factors consistently emphasised in the judges’ sentencing narrative account for the observed variation in sentencing? While consistency in approach refers to a principled way of sentence determination, consistency in outcome concerns the actual sentencing outcomes in a sense of empirical comparisons of sentence length across individual cases. Both levels are closely interrelated and it is hardly conceivable that there would be one without the other. Consistency in outcome is arguably dependent on consistency in approach because if there is not a principled way whereby similarities and differences between cases are determined and weighted, it is hardly conceivable that the ensuing sentences are predictable.

c) Systemic Consistency
Finally, for international sentencing to be considered consistent there should not be any differences depending on which tribunal actually decides the case; similar offenders should receive similar sentences across all the Tribunals – “cross-institutional (systemic) consistency”, system coherency. It is one question whether defendants within one Tribunal are sentenced consistently. It is a completely different matter whether there is a consistency across different international tribunals. The international criminal justice system consists of a number of institutions that exercise their mandates rather independently with no formal or institutional links to each other.\footnote{This is, however, not entirely true for the ICTY and ICTR. There are several institutional links between these two institutions such as the common Appeals Chamber, they used to have a common prosecutor and many judges from one Tribunal after a completion of their mandate at one Tribunal were elected to serve on a bench at the other Tribunal. In this sense, the ICTY and ICTR are rather unique.} Systemic consistency encompasses the two already mentioned levels of consistency. With respect to consistency in approach, it is necessary to analyze judicial sentencing reasoning across different courts and tribunals. Since the majority of the rules of international criminal law is unwritten, judges play an important role in determining precisely what the law is. Therefore, despite the subsidiary status of judicial decisions as sources of law in international law, the judicial decisions play a very important role in international criminal law.\footnote{J.C. Nemitz, \textit{The Law of Sentencing in International Criminal Law: The Purposes of Sentencing and the Applicable Method for the Determination of the Sentence}, 4 Yearbook of International Humanitarian Law 87(2001) at 107.} The pronouncements of the ICTY and ICTR judges in their sentencing jurisprudence could thus be perceived as an authoritative declarations of international rules on sentencing of international crimes.\footnote{Cf. Akande, supra note 20, at 53.} With respect to consistency in outcome, sentencing outcomes should be analyzed to see whether, empirically, similar factors play a similar role at different tribunals. Consequently, by analyzing sentencing case law and comparing the sentencing practice of the ICTY and ICTR it is possible to get the first hints as to the systemic consistency of the new international penal regime. Therefore, the final sub-questions of this thesis are: Can we speak of a coherent emerging international sentencing system? Do the ICTY and ICTR judges rely on similar principles in their sentence argumentation? And are the actual sentencing outcomes comparable at both Tribunals?
3. Previous Empirical Research

Most academic attention concerning the international sentencing has been devoted to its legal and normative aspects. The main criticism of the international sentencing practice, in particular of the ICTY and ICTR, stems from this field of research. The international sentencing has been labelled as ‘irrational’, ‘akin to a lottery system’ or ‘confusing, disparate, inconsistent and erratic’ to name just a few examples. However, these criticisms have never been founded on a comprehensive empirical analysis of the Tribunals’ sentencing practice. In order to assess consistency of any sentencing practice, it is necessary to conduct a thorough empirical (and quantitative) analysis. It is impossible to conclusively assess consistency of sentencing on the basis of a qualitative examination of few cases or on the basis of simple case comparisons. In order to detect patterns in sentencing, a quantitative/statistical approach is the best suited. Since sentence determination is arguably influenced by a multiplicity of factors and their combinations, one of the best statistical methods to analyse a sentencing practice is the multiple/multivariate regression analysis. Multiple regression analysis is a statistical technique that attempts to predict from independent variables (predictors) a dependent variable – that is, in case of this thesis: it attempts to predict sentence length from a combination of selected legally

---


37 Henham (2003), Ibid., at 59.


relevant factors. It seeks a combination of predictors that is maximally correlated with the sentence. It then enables assessment of the relative importance of the individual predictors given the effect of other variables in the model. Multiple regression accounts for a combined effect of all predictors included in the model. Whenever an independent variable is found significant, it has a predictive value for sentence length given the effect of the other variables entered into the analysis. As such, the analysis shows the added value of each independent variable for predicting sentence length after the effect of all the other predictors has been taken into account. Therefore, with the help of multiple regression analysis it is possible to reflect the multiplicity of factors influencing sentence length in their combination. Furthermore, especially in case of international sentencing practice it is often the case that a defendant is convicted of a multiplicity of offences but only one total sentence is pronounced. It is usually not clear how each offence and circumstances related to its commission influenced the final sentence. In this respect, transparency of international sentencing is lacking and it is impossible to see contribution of individual factors to sentence length. This fact is also one of the reasons why multiple regression analysis is very well suited as an analytical tool for assessing consistency of international sentencing since it enables to tease out contribution of individual sentencing factors to sentence length.

In contrast to a more extensive body of normative analysis listed above, empirical and quantitative research on the ICTY and ICTR sentencing has been scarce. In 2007 (when the research leading to this thesis started) there had been almost no empirical studies of international sentencing. In the course of following years, however, more and more scholars dedicated their attention to the empirical reality of sentencing of international crimes. Without any exceptions the empirical studies have identified consistent patterns in the ICTY and ICTR sentencing. The first to introduce empirical analysis into the research on international sentencing was Meernik. Together with his colleagues, he has published several studies analysing primarily the early sentencing practice of the ICTY. Their first article in 2001 presented preliminary results of an empirical analysis of the ICTY sentencing. They used multivariate regression analysis with sentence length as the dependent variable and selected legal factors such as category of crimes.

---

40 In statistics, loosely speaking, a result is significant when it is unlikely to occur by chance.
rank of a defendant, mitigating and aggravating factors and one extralegal factor, ethnicity of a defendant, as independent variables. The analysis included cases of 21 defendants. They concluded that the ICTY sentences evidenced fair justice and the ICTY is ‘remarkably effective’: those individuals with the greatest command responsibility were subjected to the severest sentences; those convicted of crimes against humanity were more likely to receive longer sentences, as were those who zealously committed their crimes. This study has been the first study to empirically evaluate the practice of an international criminal tribunal. In 2003, Meernik repeated the analysis presented in 2001 taking a broader focus regarding the influence of extralegal factors on sentence determination at the ICTY. The political and legal models were separately tested for cases of 32 defendants again using the multiple regression analysis. Meernik concluded that the ICTY’s verdicts seem not to be influenced by ‘political factors’ and the estimates of sentence length also provide little support for the notion that ICTY judges are guided by political considerations. In their next study published in 2003, Meernik and King examined “the ICTY judges’ opinions on the determinants of punishment and the data on sentences handed down by the trial chambers in order to understand which factors are the most powerful in explaining sentences”. They tested bivariate relationships (correlations) between sentence length and category of crime, selected aggravating and mitigating factors, the rank of a defendant, level of liability and ethnicity. They concluded that “there is a fair degree of consistency in the sentences conferred on the guilty” and “sentences are premised on those critical factors that the judges are admonished to employ by the ICTY Statute and RPE”. In 2004, Meernik shifted his attention to the ICTR. The principal focus of this purely descriptive

43 Sentence determination should only be based on legally relevant factors. It should never be affected by consideration of the so-called extralegal factors such as ethnicity, race or gender of the offender. Cf. A. Ashworth, *Sentencing, in* The Oxford Handbook in Criminology (R. Morgan and R. Reiner M. Maguire ed. 2007) at 1003. In case of the ICTY and ICTR sentencing the legal factors are those considerations that should play a decisive role in the sentence determination. They are derived from the international law, the ICTY and ICTR Statutes and the Rules of procedure and evidence and include factors such as nature of the crime, degree of responsibility or aggravating and mitigating circumstances. In contrast, the extralegal factors are factors not regulated or permitted by law. They should not influence the sentence decision-making in any way. Examples of these legally irrelevant considerations are e.g. ethnicity of the offender, professional background or personal characteristics of judges or political factors.

44 Meernik & King, supra note 42.


47 Bivariate correlation describes the degree of relationship (association) between two variables, e.g. in this case sentence length and category of crime. It is a measure of the linear relationship between the two variables. By computing bivariate correlations, (possible) effects of other relevant variables on sentence length are not accounted for. See e.g. A. Field, *Discovering Statistics Using SPSS* (SAGE Publications 2nd Edition ed. 2005), at 107.

48 Meernik & King, supra note 46, at 717.
study was an analysis of sentences for the crime of genocide handed down by the ICTR. He analysed all the verdicts and sentences of ICTR trial chambers at the time, i.e. 10 individual verdicts and sentences. After a very brief analysis of the ICTR verdicts, he examined the ICTR sentences in a descriptive manner; noting difficulties of this empirical enquiry given the limited number of cases and prevalence of life sentences. He concluded that verdicts and especially sentences are informed by identifiable and fairly consistent standards. In 2005, Meernik et al. re-shifted their focus on the ICTY again and tried to explain judicial decision making by analyzing the impact of individual, national and international factors using multiple regression to analyze sentencing and binary probit regression to analyze judges’ verdicts. This research aimed at explaining judicial decision making on a personal level and the authors examined the influence on sentence length of many legally irrelevant variables. Again, the conclusion was that the ‘international legal facts’ model (i.e. legal criteria derived from the ICTY Statute, international law and Rules of Procedure and Evidence as predictors) best explained judicial decision making, especially regarding sentencing. Furthermore, ICTY judges generally follow the case facts in light of the guidelines laid down in the ICTY Statute. Finally, in 2007, Meernik analysed the ICTY sentences and the relationship between trial and appeal including appeal modifications of trial sentences. By examining bivariate relationships between sentence length and selected factors, he concluded that the Appeals Chamber left untouched most of the trial decisions on a count-by-count basis, so it would seem there is a fair degree of respect for trial decision-making. When the Appeals Chamber modified a decision, it was mostly by way of a sentence reduction.

More recently in 2009, Doherty and Steinberg published their draft paper examining both the ICTY and ICTR sentencing practice. They sought answers to the question: what factors explain the length of sentences handed down by the ICTY and ICTR? In order to answer this

50 Binary probit analysis is a type of regression used to analyze models where a dependent variable is not numerical but categorical binomial, i.e. made up of categories and it has only two possible outcomes such as e.g. conviction/acquittal or male/female, J. Scott Long, Regression Models for Categorical and Limited Dependent Variables (Sage Publications, Inc. 1997).
question Doherty and Steinberg tested how 20 selected legal and extra-legal factors influence sentence length. They analysed all the data from the ICTY and ICTR cases decided by December 2008 by using ordered probit regression. Despite raising a lot of (often unjustified) criticism against the previous empirical studies they basically confirmed their results – the most important sentencing determinants seem to be the gravity of crime (in a sense of category of crimes and crime magnitude) and position of the offender. Furthermore, they found that several factors the Tribunals claim to matter, such as various mitigating factors, in fact bear no significant relationship to sentence length.

In 2010 in his article on the ICTY sentencing Ewald identified emerging patterns in the ICTY sentencing practice while examining relationship between leadership level of a defendant and sentence severity. Ewald reviewed the existing scholarship dealing with international sentencing noting that the current criticism of international sentencing practice regarding its consistency is based more on an anecdotal evidence rather than empirical investigation. In the empirical part of his study Ewald statistically analysed the ICTY sentencing practice, in particular the relationship between the leadership level of a defendant and sentence severity. First, he provided descriptive overview of the ICTY sentences distinguishing between different groups of cases (i.e. guilty plea versus non guilty plea cases and Art 7(3) versus Art 7(1) cases). Then, Ewald presented results of the regression analysis with sentence length as a dependent variable and a leadership level as a predictor. He identifies ‘a horse shoe’ (U-shape) relationship between sentence severity and leadership level at the ICTY while lower leadership levels “tend to be linked to medium-higher sentences, medium leadership levels are related to mainly under-average sentences, and higher leadership levels are linked to higher sentences”. He concluded that for the majority of the ICTY accused, the impact of leadership level on sentences is rather

---

54 The predictors in their model included category of crimes, ICTY versus ICTR, number of victims, number of incidents, sadism, rank, hands-on perpetrator versus order-giver, superior responsibility versus participation in JCE, ten mitigating factors (duress, duress to a family member, remorse, character of a perpetrator, assistance to victims, guilty plea, cooperation with the OTP, voluntary surrender, promotion of peace, old age, young age), procedural irregularities, education, age, sex and nationality of a defendant, nationality and gender of victims, nationality of judge, prosecutor and defence lawyer and challenging the court’s legitimacy.

55 Ordered probit is a type of a regression analysis where a dependent variable is neither numerical nor categorical binomial but categorical with more possible outcomes. In their article, Steinberg and Doherty treated sentence length as a categorical variable distinguishing among 4 categories: ≤5 years; 6-15 years; 16-50 years; life sentence.


57 Ibid., at 365.

58 Ibid., at 398.
stable. Ewald, however, cautioned that the findings were still explorative and preliminary and show only trends.59

In 2010 Jodoin offered a slightly different perspective and examined judicial decision making at the ICTY and ICTR drawing on models of judicial behavior developed for domestic courts. His main contention was that the ideas and interests of the ICTY and ICTR judges account for variations in their decision-making. His study was thus not concerned with factors explaining sentencing or sentencing consistency as such, but with factors explaining individual judges’ decisions through the application of attitudinal and strategic models. He analyzed the influence of judges’ professional background (i.e. judges occupation before taking up their judicial position (whether a judge was a diplomat, practitioner or an academic), combined with a self-constructed activism score) and defendants’ ethnicity on the sentence severity at the ICTY and ICTR. He showed that judges in both ICTY and ICTR displayed “stable patterns of judicial behavior […] former academics tend to be more international activist, former diplomats tend to be more statist conservative and former judges/practitioners did not appear to tend towards either […]”.60 The results of his empirical analysis suggested that judges within the ICTY and ICTR evinced divergent patterns of judicial behavior in terms of sentencing practices, reflecting dissimilarities in their environments - Hutu tended to receive harsher sentences compared to Serbs or Croats. Jodoin backed up the analysis only with correlational evidence and therefore, does not control for other, legally relevant, factors that may account for the identified variation.

Finally, in 2011 D’Ascoli published a book entirely dedicated to the analysis of the ICTY and ICTR sentencing practice.61 After a rather thorough and well-researched doctrinal analysis of the Tribunals’ sentencing case law, she dedicated one chapter to an empirical study of the ICTY and ICTR sentencing practice. D’Ascoli analyzed all the ICTY and ICTR cases decided up to August 2010. Her conclusions are mostly based on descriptive comparisons (mean sentences, medians, standard deviations or ranges of sentences for different categories of cases) but she also presents correlations between sentence length and examined variables, and a multivariate regression with sentence length as the dependent variable. As independent variables D’Ascoli includes a rather

59 Ibid.
60 S. Jodoin, Understanding the Behaviour of International Courts: An Examination of Decision-making in the ad hoc International Criminal Tribunals, 6 Journal of International Law and INternational Relations (2010).
61 D’Ascoli, supra note 36.
extensive set of case/defendant characteristics. The main conclusions of D’Ascoli are in line with the previous empirical research. She states that “[t]he results achieved throughout the analysis show that there is a certain degree of consistency in the sentencing practice of the ad hoc Tribunals, and that certain factors are more influential than others on sentence length.”

“[F]air standards and consistent criteria are followed by judges of the ICTY and ICTR. There is therefore no doubt that sentences meted out by the ad hoc Tribunals tend to show general patterns of consistency and to satisfy general criteria of legitimacy.”

With respect to the multivariate findings there are, however, several points of criticism that can be raised against her analysis:

(i) arguably, some variables are represented twice in her models (such as guilty plea and mitigating factor-guilty plea) or to a large extent overlapping (e.g. conviction of genocide mostly occurs at the ICTR, therefore it is highly correlated with the type of tribunal variable; type of participation is arguably correlated with variables expressing aggravating factor direct participation or mitigating factor indirect participation; there are also high correlations among some aggravating and mitigating factors (e.g. guilty plea, remorse, cooperation with OTP) etc.). The correlations among various variables raise issues of multicollinearity and among others increase the probability of the Type II errors in the coefficient estimates;

(ii) the sample size is arguably insufficient to make reliable estimations with such an extensive set of predictors (she is testing a set of 37 predictors on several data consisting of 1) all sentences (trial and appeal) issued by both Tribunals (N=201); 2) trial sentences issued by both Tribunals (N=131) 3) only finalized sentences issued by both Tribunals (N=95); 4) all sentences (trial and appeal) issued by the ICTY (N=132) or 5) all sentences (trial and appeal) issued by the ICTR (N=69); and finally

(iii) the assumption of independence of observations, one of the basic assumptions of multivariate regression, is arguably violated in the models based on trial and appeal decisions

---

62 The independent variables include the type of the tribunal (ICTY/ICTR), category of convicted crime, type of participation (distinguishing between two categories: conviction under Art 7/6 (1) and conviction under Art 7/6 (3)), rank, guilty plea and various aggravating (crime gravity, victimisation, trauma, vulnerability of victims, premeditation, cruelty/sadism, willingness, direct participation, superior position, abuse of authority/trust, criminal record, good character, conduct during trial) and mitigating factors (superior order, no prior crimes, family status, old age, young age, poor health, remorse, unwillingness in the commission of crimes, indirect participation, help offered to victims, co-operation with OTP, voluntary surrender, guilty plea, testimony provided, good character, conduct during trial).

63 D’Ascoli, supra note 36, at 259.

64 D’Ascoli, supra note 36, at 260.
combined, since the same cases with similar characteristics decided by a trial chamber and modified on appeal are included twice in the dataset.

The multivariate analysis does not seem to form the main basis for her conclusions that are generally formed on the basis of descriptive and bivariate correlational data.\(^{65}\) However as noted above, in order to assess conclusively the consistency of any sentencing regime and the relevance of individual sentencing factors given other pertinent properties of cases, estimates from multivariate models are necessary. The sentence determination is a complex phenomenon which one cannot unravel with descriptive methods that only aim to summarize the main trends.

### 4. Structure of this Thesis

Despite the increased interest in the empirical reality of international sentencing in the last years, the empirical enquiries into international sentencing have still been relatively uncommon. Basically, the focus of these studies has been limited to one Tribunal at the time. Only two of the above mentioned studies have focused on both, the ICTY and ICTR sentencing practice. The main problematic issue of these studies lies in the fact that the authors assumed that the mechanisms of sentence determination are similar at the ICTY and ICTR. Therefore, they neither examined the models separately for each Tribunal nor tested for possible differences between the Tribunals. Next, none of the above articles focused on the concept of consistency. Consistency of international sentencing was only cursorily mentioned but more as a by-product of the conducted analyses. Therefore, this thesis develops the empirical scholarship further by focusing primarily on consistency of international sentencing as represented by the practice of the ICTY and ICTR. It presents an empirical, quantitative analysis of the Tribunals’ sentencing practice with the aim of identifying any consistent patterns in the actual sentences handed out by the ICTY and ICTR. The sentences handed out by the Tribunals are analyzed separately for each of the Tribunals and then the ICTY and ICTR sentencing practice is combined and compared to detect any possible differences in sentencing mechanisms.

Next to the introduction and conclusion, the thesis is divided into three main parts each dealing with the three different dimensions of consistency: (i) consistency in approach, (ii) consistency in outcome and (iii) systemic consistency.

\(^{65}\) See supra note 46.
The first part consists of Chapter I and introduces to the reader the issue of sentencing at the ICTY and ICTR. It analyses consistency in approach of the ICTY and ICTR sentencing case law. It doctrinally discusses the most controversial aspects of the sentencing narrative of the ICTY and ICTR judges and reviews inconsistencies in the judges’ reasoning already noted by various academics focusing on international punishment. This chapter demonstrates that on a general level, a basic set of sentencing principles is consistently emphasised in all the cases and adhered to by the judges at both Tribunals. There are, however, inconsistencies and disparities across cases relating to particularities, such as what factors are relevant for the gravity assessment and whether a particular mitigating/aggravating factor indeed aggravates/mitigates the sentence in a particular case.

The rest of the thesis concentrates mainly on ‘consistency in outcome’ of international sentencing and presents quantitative analyses of the ICTY and ICTR sentencing practice.

In the second part called ‘Consistency in Outcome’ and consisting of Chapters II and III, the sentencing practice of each of the Tribunals is examined separately. Chapter II focuses on the ICTY sentencing practice. By analysing the case-law up to May 2009, legal factors influencing sentencing decisions at the ICTY are examined. The selection of legal factors is based on the ICTY case law analysis and stems from the general principles judges consistently emphasise in their sentencing argumentation (as identified in Chapter I). The tested variables include the category of convicted crime, the number of guilty counts, rank, modes of liability and number of aggravating and mitigating factors. The extent to which the selected factors predict sentence length is tested in a multiple regression analysis. The analysis suggests that the sentence can be to a large extent predicted by legal criteria.

Chapter III turns the attention to the ICTR. Unlike ICTY defendants, almost all ICTR defendants have been convicted of and sentenced for genocide – arguably the most serious international crime. The third chapter thus examines the sentencing practice of the ICTR and examines the relationship between sentence severity and the primary consideration in sentencing – crime gravity. Due to the low number of cases and the high prevalence of life sentences at the ICTR it is inadvisable to use regression (or any similar statistical method) to analyze ICTR sentencing practice. Therefore, a more exploratory study is performed by using HOMALS - homogeneity analysis: a descriptive method that seeks to identify multivariate associations among selected variables. First, the relevant principles stemming from ICTR case law are
reviewed, followed by an examination of the interrelationship between sentence severity and factors relating to crime gravity, such as category of crime, scale of crime and the form and degree of a defendant’s involvement in the crime. The analysis demonstrates that the ICTR judges appear in most cases to follow the main principles emphasized in their case law, with sentences gradated in line with the increasing seriousness of defendants’ crimes and their culpability.

The third and final part of this thesis focuses on systemic consistency of international sentencing and compares and combines the data from the sentencing practice of both, ICTY and ICTR. It consists of the final two chapters of this manuscript – Chapter IV and V. The sentencing narrative of the ICTY and ICTR judges and its coherence was already reviewed in Chapter I. Consequently, the final two chapters of the thesis compare and analyse sentencing outcomes across the Tribunals. Chapter IV presents descriptive comparisons between the ICTY and ICTR sentences. In this chapter the ICTY and ICTR sentences are juxtaposed and compared in relation to different categories of crime, types of offence, scale of crime, modes of individual liability, “ranking” of defendants and finally, aggravating and mitigating factors. The differences between the Tribunals regarding sentence severity for individual categories are discussed. The analysis shows that the main differences in sentence severity between the ICTY and ICTR are arguably connected to the different case composition at the Tribunals. At the ICTR the majority of defendants are convicted of genocide and also many key figures (members of government and other high ranking figures) and organizers of violence stood trial in Arusha. Other factors that could account for the difference in sentence severity include the relatively limited range of crimes prosecuted by the ICTR and the lack of guilty pleas compared to the ICTY.

In the final, fifth chapter, the consistency of sentencing of both Tribunals is statistically examined. Similar to all the previous chapters, the model in Chapter V is based on legal factors stemming from the case law analysis and general principles the ICTY and ICTR judges consistently rely on in the case law. The multiple regression analysis conducted in this final chapter specifically focuses on possible differences in sentencing mechanisms between the Tribunals. The results suggest that similar, legally relevant patterns have emerged in the sentencing practice of both Tribunals and there do not seem to be any notable differences in the mechanisms underlying sentence determination at the ICTY and ICTR.
The thesis ends with the conclusions summarizing the findings and their implications and offering suggestions for future research. Since recently many scholars have called for international sentencing guidelines to address the alleged inconsistencies in international sentencing, the conclusion also discusses feasibility and necessity of drafting international sentencing guidelines given the results of the conducted analyses and a specific character of international crimes.