Check for updates

Designing the Swedish Crime Harm Index: an Evidence-Based Strategy

Fredrik Kärrholm[1] · Peter Neyroud[2] · John Smaaland[1]

Abstract

Research Question Which method for deriving a Crime Harm Index (Policing 10:171–183, 2016) for Sweden from criminal justice sources offers the best evidence for providing a sensitive indicator of differences in harm levels across offence categories?
Data The number of days of imprisonment for each offence category associated with five different kinds of scales were extracted and compared: consensus by an expert panel of judges, the statutory maximum penalties, statutory minimum penalties, the average of maximum and minimum penalties and the average of actual sentences imposed in a recent time period for each crime type. Unlike the UK, for which the Cambridge Crime Harm Index draws on sentencing guidelines, Sweden has no such guidelines to offer.
Methods The data were compared for sensitivity defined as the difference in length of imprisonment days between high and low severity crimes, as well as other characteristics of the data sources.
Findings Given the available alternatives, the sentencing data average of actual sentences handed down by crime type provided the greatest reliability and sensitivity across the penalties for offences of high and low severity. Applying that method to both crime trends and crime mapping produces substantially different results from counting all crimes with equal weight and can be used by police and others to allocate resources with greater precision in relation to harm prevention.
Conclusions On both empirical and normative grounds, the average sentences in a recent time period for each crime category provides the most sensitive and democratic method for establishing an officially recognized Swedish Crime Harm Index.

Keywords Crime Harm Index · Average sentences · Crime trends · Crime maps

✉ Peter Neyroud
  pwn22@cam.ac.uk

Extended author information available on the last page of the article

🍂 Springer

Introduction

> The only purpose for which power can be rightfully exercised over any member
> of a civilized community, against his will, is to prevent harms to others. (John
> Stuart Mill ([1859] 1978: p. 9).

Different crime types cannot be valued equally. When calculating detection rate, for
example, it is not sensible to equally value murder and shoplifting. Likewise, in the
process of identifying geographical areas that are to be prioritized for patrolling, it is
irrational to value serious assault and verbal harassment equally.

For any police organization, differentiation and prioritization between crimes are
crucial to be cost-effective and ethical (Neyroud 2019). Moreover, we argue that precise
and correct decision-making is dependent on systematic statistical methods. For this
purpose, statistical tools have been developed in several countries for weighting crime
counts according to the severity of different crime types. These tools are often referred
to as Crime Harm Indices (CHI) (Sherman et al. 2016).

However, as Sherman et al. (2016) set out, in presenting the Cambridge Crime Harm
Index, a CHI can be designed in multiple ways. In this study, which centres on the
process of designing a Swedish Crime Harm Index, the possible methods are thor-
oughly explored, balancing the prerequisites of the legal system with the reliability and
practicability of the proposed index. The aim is to create a useable index, as a system to
distinguish degrees of criminality and harm.

What Is a CHI?

While there is a substantial literature on harm in relation to crime, the development of
Indices is a relatively recent phenomenon (Sherman et al. 2016). In creating them,
crimes are integrated into a weighted *index*, based on how much *harm* the different
offences cause. Technically, a CHI has a score derived from the application of a metric,
which weights different offences based on a proxy measure of the harm it causes,
relative to other crimes. In this context, "harm" refers to how harmful a crime type is
perceived to be relative to other types of crimes (Ignatans and Pease 2015; Mitchell
2017; Office for National Statistics 2016).

The argument for using such an index in quantitative analysis has been made
in different ways by several criminologists, amongst them Sherman (2007, 2011,
2013; Sherman et al. 2016), who has repeatedly stated that "all crimes are not
created equal". Sherman argued that an index provides a useful statistical tool for
analysis and decision-making, for example, regarding resource allocation and
crime prevention (Gorsuch 1938; Sellin and Wolfgang, 1964; Wolfgang et al.
1985; Sparrow 2008; Greenfield and Paoli 2013; Neyroud 2017). Sherman et al.
(2016) provides a series of examples of the potential usefulness of a crime harm
index. However, using a crime harm index also raises moral dilemmas about the

fairness of resource allocation that need to be considered and discussed (Sherman 2007).

To comprehend the usefulness of CHI, it is important to understand that a small minority of all crimes—as well as people, places and times—cause a great proportion of the total harm. Sherman has described these overrepresented units as *the power few* (Sherman 2007). This phenomenon is a fundamental prerequisite for the usefulness of taking *crime harm* into account for analysis and decision-making in the police. Sherman's observation about *the power few* correlates with a general trend, characterized by a small number of units, of all possible units, accounting for a majority of the total value. This has been observed in many fields: social, scientific, geophysical, actuarial and more (Sherman 2007; Zipf 2016). The distribution trend is sometimes referred to as *The Pareto Distribution* or *The 80/20 principle* (Koch 2007; Reed 2001). In respect of crime, this phenomenon can sometimes be extreme. For example, in one UK County, an analysis using the Cambridge CHI showed that of all couples who were subjects of callouts for domestic abuse, 1.7% generated 80% of total crime harm (Bland and Ariel 2015).

How Can a CHI Be Created?

To create an operational CHI, harm must be quantitatively assessed, with data that consists of numerical values for different crime types. Over the years, several possible methods have been developed and tested with the purpose of generating the numerical values. Karrholm (2019) reviewed three approaches, which can be broadly categorized as survey-based, cost-based and sentencing data–based approaches. In this article, we limit the discussion to the most practical and readily available source of data across a wide range of offence categories, which is a CHI derived from evidence about sentencing severity measured by days of imprisonment.

An emerging academic consensus suggests that the preferable method to create a CHI is by using the law and sentencing outcomes. The main reason for this is that the democratic process is considered to produce a compromise of conflicting viewpoints of how to assess harm, expressed through the law. Thus, the approach is both practical and defensible (Sherman et al. 2016) However, using the law to create a CHI can be done in several different ways, described below. The choice of method is also dependent on the design of the legal system, which differ between countries. Even within the fairly narrow category of Scandinavian civil law, there are considerable differences between the Nordic countries (Träskmark 2003). Therefore, the creation of a CHI must carefully be considered, with regard to the legal system and other contexts in the country in question.

Sentencing and Prosecutor Guidelines

Using sentencing guidelines, Sherman et al. (2016) have created the Cambridge CHI. Crime types are weighted based on the "starting point" in the England and Wales Sentencing guidelines, which is converted into a defined number of days in prison. The starting point reflects the penalty for a first-time offender without aggravating or mitigating circumstances. The CCHI also excludes proactively detected and previously

unreported offences. One of several benefits to this approach is transparency, as the Cambridge CHI solely relies on official information, already in the public domain. Sherman also suggests a method for comparative analysis of Crime Harm Indices, regarding democracy, reliability, and cost, which we return to below.

Ratcliffe (2015a, b) used a similar method, creating a crime harm index based on the gravity scores in the non-mandatory guidelines for judges in PA, USA. Ratcliffe claimed that this method is independent of the police (thus, not manipulable) and specific enough concerning offence categories. In comparison with the Cambridge CHI, Ratcliffe's Index also meets a three-pronged test designed by Sherman et al. (2016). However, Ratcliffe's scale was relatively short and the differentiation between low and high harm crimes was consequently less clear.

Andersen and Mueller-Johnson (2018) created a Danish Crime Harm Index with a similar method, by using official guidelines of the Director of Public Prosecutions office. There are no official guidelines in Denmark directed to judges.

Expert Estimates of Average Judicial Decisions

In the absence of sentencing guidelines, another recently used method to define crime harm is to estimate average judicial decisions. By asking judges, Rinaldo (2015) created a Swedish Crime Harm Index, consisting of 118 selected crime codes for violent offences in public places. Five judges received a survey of crime types and were asked to assess typical verdicts. For those crimes where the normal penalty was not imprisonment, the judges were asked to convert other sanctions to prison time. By doing this, the judges interpreted the law, and therefore, the method is substantially different from that based on comparing other professionals, students or the general population about their subjective assessments of harm (such as Sellin and Wolfgang 1962 or Wolfgang et al. 1985). In contrast to such subjective assessments, the judges in Rinaldo's study were instructed to make objective assessments of the crime type's penal value.

Sentencing Data

Instead of estimating average judicial decision, court records can provide sentencing data that shows the actual average judicial decisions or the average of sentences actually passed. There are a number of examples of Crime Harm Indices created with this method, amongst them The Canadian Crime Severity Index (Wallace et al. 2009; Statistics Canada 2015), Crime Severity Score for England and Wales (Office of National Statistics 2016), Japan Crime Harm Index (Taira 2018) and Hong Kong Crime Harm Index (Chong 2019). All of these are based on the average length of the sentences for the different crimes. A variation of this method was used by House and Neyroud (2018), who developed a Crime Harm Index for Western Australia, which is based on the average length of sentence for a first-time offender, and Curtis-Ham and Walton (2017), who developed a New Zealand Crime Harm Index, based on estimates of the minimum sentence for a first-time offender using sentencing data.

The argument for focusing on the first-time offender is to obtain an equivalent valuation for each crime, merely based on harm, without adjustment for the offender's criminal history or other circumstances of a particular offence. Sherman et al. (2016)

argue for this method, with homicide as an example: "Whether a first time offender or a serial killer murders someone, the murder creates the same harm to the victims, his or her families, and communities. The actual punishment each offender 'deserves' to receive is a very different question from how much harm the crime has caused." (Sherman et al. 2016: p. 2). This is accurate both on an individual and aggregated level, as previous offending does not affect the quantity of harm inflicted on society (Kleinig 1978).

Penal Code Metrics: Maximum, Minimum and Average of Both

Mitchell (2017) created the California CHI based on maximum sentences for each crime type, accordingly to the California Penal Code. Mitchell argued that this provided a satisfactory reflection of the harmfulness of crimes. Mitchell pointed out that the maximum sentence is not the typical sentence, but the method provides the "bottom line" (or perhaps, more accurately "top line") and consistency needed for analysing the relative harm. Sherman et al. (2016) have stressed that the relevant component is a single metric that provides consistency and the possibility of estimating relative harm reliably. However, Mitchell recognized that the simplicity of this method limits its use to larger datasets.

Meanwhile, House and Neyroud (2018) rejected the use of penal code metrics in Western Australia because they provided insufficient sensitivity to the differences across perceived normative severity of various offence types.

What Does a Good CHI Look Like?

There are several different ways to create and design a CHI. However, the literature on the subject provides guidance for development. Sherman et al. (2016) formulated three criteria, which they consider to be "essential" for every CHI to meet, for the prospect of implementation and utilization. In addition, Curtis-Ham and Walton (2017) have defined two additional, complementary criteria.

The "three-pronged test" (Sherman et al. 2016):

(a)   Democratic

Does the index reflect the resolution of conflicting viewpoints by a democratic process? Crime harm can be defined in different ways and is manifold: crime can cause harm in a form of financial costs, for the victim, physical injury, emotional damage, social harm and destruction of the environment (Ashby, 2018). There is simply no objective and true definition of harm, but through the democratic process, a compromise can be reached. Furthermore, the democratic criteria are relevant because public power, exercised by the public sector, proceeds from the will of the people. Using a democratic approach, it is possible to obtain a definitive answer, in contrast to the ongoing—and seemingly never ending—academic and philosophical discussions about how to estimate or calculate harm. Ratcliffe (2015) points that it is also important to separate the police from defining the harm values in the index, as it could be used to determine the effectiveness of

the police. Making the design of the metric independent of the police guarantees that there will be no suspicion about fixed parameters favouring the police.

(b) Reliability

Does the index provide a reliable measure across units—time, place and people—with the same results for the same levels of harm? The argument for this is that the index must be reliable regarding consistency of measure, as one of its main function as a statistical tool is to compare harm levels between different groups, areas and time periods. For example, if the index is influenced by changes in offender demographics or fluctuating sentencing trends, it cannot be considered reliable.

(c) Affordable

Can the index be implemented without additional funding? The argument is based on the understanding that police forces often have competition for funds or are subject to austerity. Implementing a Crime Harm Index is simply more feasible if it does not cost money or resources, and can be done without creating a new system for recording crime statistics.

Curtis-Ham and Walton's Additional Tests (2017)

(d) Valid

Does the index measure harm objectively? The argument for this is strictly logical, as the index is supposed to be based on harm Therefore, it should not be affected by other variables, such as the area where crime is committed, who is committing the crime or what the offenders' criminal records are. This argument is rational in relation to the purpose of the index, although the judicial system in other respects takes additional variables into account. Validity does also include being precise about crimes types and not using categories that are too broad (see also Ratcliffe 2015).

(e) Operational

Is the index useful regarding its intended purpose? Is it easily understandable, and possible to apply in practice without extensive instruction or training? An estimate is that 80 to 90% of people-dependent innovations in business stop at so-called paper implementations, which means no real implementation (Fixsen et al. 2005). To ensure the use of a CHI, it must be easily operationalized (Pol 2016).

Research Question: Which Data Source for Sweden's CHI?

In choosing a method to create a CHI, the aim should be to fulfil the five criteria: democratic, reliable, affordable, valid and operational. The index should also have sufficient year-to-year reliability for application to ongoing police resource allocation decisions.

Sweden does not have any sentencing guidelines, on which many indices are based, including the Cambridge CHI, though the Swedish Penal Code has strong parallels with

the sentencing guidelines in England and Wales. Within the requirements of the Swedish legal system, there are nevertheless several possible methods to create a Swedish CHI.

As a civil law system, the Swedish law is mainly derived from codifications in a statute passed by legislatures. Therefore, the Swedish Penal Code contains detailed provisions on almost all crimes, which functions as the principal guidance for the court when deciding sanctions (Law Commission 1989; Herring 2004; Swedish Ministry of Justice, 1999a). The jurisdiction of England and Wales does not have such a clear and unambiguous criminal code, hence the development of sentencing guidelines, whereas such a supplement has not been considered necessary to the Swedish Penal Code.

The English-Welsh sentencing guidelines and the Swedish Penal Code provide similar instructing information: three ranges of sentencing (penalty scales) for each crime (two ranges for some crimes in the Swedish Penal Code): one with no aggravation or mitigation, a second with some aggravation, and a third with serious aggravation. Thus, both the sentencing guidelines and the penal code provide "tariffs" for different crimes.

(Herring 2004; Swedish Ministry of Justice, 1999b) Furthermore, the penalty scales in the Swedish Penal Code are primarily determined based upon the crimes' seriousness. Like the English-Welsh sentencing guidelines, ranges of sentencing are explicitly based on crime gravity (the penal value), without reference to, for example, the offender's prior crimes or prior record (Swedish Ministry of Justice 1999). One crucial difference, however, is that the penalty scale, compared with the sentencing guidelines, is more wide-ranging, in order to provide the possibility of very low or very severe penalties in special cases at either end of the range of the specific offence.

Given this framework is in the Swedish legal system, there are three data sets that could be used to create a Swedish CHI:

1. Judges' or prosecutors' estimate of average sentences
2. Sentencing data of average sentences
3. The Penal Code

The research question asks broadly which of these methods would be best for Sweden. The research method for answering this question is drawn partly from the five tests described above (Sherman et al. 2016; Curtis-Ham and Walton 2017). But the most evidence-based dimension of those tests is the reliability of the CHI in distinguishing, with sufficient sensitivity, between the different levels of harm to victims associated with each offence category. Thus, the research question can be refined as follows:

Which data source best meets the five tests for a CHI, with special emphasis on the reliability and sensitivity of the data source?

Data

Alternative 1: Expert Estimates

One alternative is to ask judges or prosecutors to estimate the average judicial decisions. Rinaldo (2015) asked five judges in Sweden, using a questionnaire about the

"expected verdict" for 118 selected crime codes for violent offences in public places. The inter-rater reliability (Pearson correlation) was ranging from $r = 0.966$ to 0.994, indicating high agreement between the five judges. Using this data, Rinaldo created the first Swedish CHI, which has been used successfully to identify harm spots in the city of Gothenburg (Rinaldo 2015).

In Denmark, Andersen (2018) began her research by seeking the views of prosecutors instead of judges. The prosecutors estimated the "normal sentence" by interpreting the guidelines from The Danish Director of Public Prosecutions. The guidelines are advisory and serve as a guide for the prosecutor sentencing demand. She later abandoned this method for creating the Danish Crime Harm Index (DVHI), in favour of applying the official guidelines directly (Andersen and Mueller-Johnson 2018). The guidelines are based on both case-law and the minimum sentence stated in the penal code by the parliament. However, Sweden does not have guidelines of any sort, and therefore, an estimation of typical sentences can be assumed to be done most accurately by the profession responsible for meting out punishment: the judges.

As has been demonstrated by Rinaldo (2015), a CHI can be created simply by asking judges. However, the maintenance of such an index is demanding, if it is to be implemented and used over time for police purposes. Since the law and legal practice is constantly changing, new questionnaires must be sent at least once a year to judges. This is a time-consuming process, and it presents a clear risk of inconsistency from different groups of respondents. Hence, the method's reliability can be questioned, as well as its affordability and effectiveness. Creating an index based on judges' estimate may lead to questions within the police about subjectivity, undermining implementation.

Moreover, the extent to which the index only measures harm is uncertain, as it may be affected by other variables. In Sweden, judges take numerous additional variables into account when deciding sanctions, because of the many aggravating and mitigating circumstances set out in the law. For some crime types, offender characteristics may be taken into account: for example, it is more common that young men commit violent crimes in public places, and age is a mitigating factor according to the Swedish law (Brottsförebyggande rådet 2018).

Alternative 2: Sentencing Data

Another way of creating a Swedish CHI would be to use sentencing data from all judges in all cases. This appears to be preferable to asking a few judges, as it seems likely to be more reliable. However, while judges can estimate the average sentences for first-time offenders (Rinaldo 2015), the sentencing data does not provide that information, only the average sentence for all offenders.

Sentencing statistics are published every year by The Swedish National Council for Crime Prevention. They contain qualitative and detailed sentencing statistics, providing information on the number and average prison sentence for all 350 crime types. However, the average sentence for different crimes is affected by several factors, and three in particular:

Firstly, judicial decisions sometimes include more than one offence, as offenders can be sentenced for more than one crime at the same time. This is a consequence of the investigations' relatively long processing time, repeat offending and waiting times for

court hearings. In these cases, the court imposes a joint sanction for the crimes, but in the *sentencing* statistics, the sanctions are related only to the main crime. Thus, multi-crime sentences will only be counted as a sentence for the most severe crime (i.e. highest penal value). Therefore, the "average *sentence*" for a crime type is not only based entirely on the specific crime type but also on other crime types.

Secondly, the offender's age dramatically affects the sentence, if the offender is younger than 21 years and especially if he is no more than 18 years old. This entails that the data does not only merely reflect the nature of the crime but also the offender. However, offenders younger than 21 are seldom sentenced to prison, which may reduce the impact of this factor (Brottsförebyggande rådet 2018).

Thirdly, prison is in reality an unusual sentence for many crimes in Sweden, as the law states that "in choosing a sanction, the court shall pay special attention to any circumstance or circumstances that argue for the imposition of a less severe punishment than imprisonment" (Swedish Ministry of Justice 1999: p. 139) As a consequence, the average sanction is for some crimes based on only handful or solitary cases. Furthermore, some crimes are overall unusual or seldom lead to convictions, for example, sexual intercourse with a sibling or misuse of authority. This may be more of a theoretical problem as these crimes constitute a very small proportion of the total entire quantitative basis when using a CHI as statistical tool. Furthermore, the "problem" can be minimized by using available sentencing data for several years, calculating an average. For example, the Crime Severity Score for England and Wales uses an average of the latest 5 years of sentencing data (Office for National Statistics 2016).

There is no Swedish dataset that can distinguish *first-time offender*, as was available to House and Neyroud (2018), or the *minimum sentence* for a *first-time offender* (Curtis-Ham and Walton 2017). Focusing on first-time offender gives an equivalent valuation for each crime, merely based on harm, without adjustment for the offender's criminal history or other circumstances of a particular offence, but this is not possible using the available Swedish sentencing data.

Alternatives 3,4 and 5: the Penal Code Maximum, Minimum, or Average

Three more alternatives could create a CHI using the Swedish Penal Code and some complementary special legislation, containing provisions with penalty scale for *all* crimes. This alternative appears to meet the five criteria from Sherman et al. (2016) and Curtis-Ham and Walton (2017), being democratic, reliable, affordable, valid and operational—but the penalty scales are wide-ranging (Table 1).

For petty crimes, the penalty scales often have a maximum punishment that in reality is never used, but it offers the court the possibility of sentencing offenders to severe penalty in special cases. For example, offenders sentenced for inflicting damage have often committed a crime that consisted of painting graffiti on a building or smashing a window. The sentences for such crimes are normally fines. However, the penalty scales provide the option for the court to impose a 2-year prison sentence, which requires unusual and aggravating circumstances, such as where the damage has been done to cultural heritage.

For some crimes, the penalty scale also has a minimum punishment that is rarely used and merely exists for exceptions. For example, stalking is considered a serious

crime in Sweden and often such a crime is associated with other crimes against women and results in prison. However, the penalty scale starts with only 14 days of imprisonment.

It could be argued that, compared with the index created by Rinaldo (2015), one disadvantage might be that an index based on the penal code will consist of broader classifications, compared with the more detailed crime code system used by the Swedish Police Authority. Even so, there are more than 300 crimes in the penal code, which we argue is a sufficient range of offences to cover the vast majority of all crimes.

If choosing this alternative, there is still another choice to be made: should the index be based on the statutory minimum sentence, the statutory maximum or the average of the statutory minimum and maximum sentences? All three alternatives provide a logical and consistent relationship. It can be perceived as counterintuitive, but the primary purpose of a CHI is not to reflect the absolute value of a typical sentence, but rather the appropriate *relationships* (or distances) between the score for each crime type. A CHI is required to be a metric scale, which weights different offences based on a proxy measure of the harm it causes, relative to other crimes. As Sherman et al. (2016) state: "What is relevant is the consistency obtained from a single metric to reliably estimate a harm level in any unit for comparison to the harm level in any other unit" (p 6).

Method: Comparing CHI Scores Across Five Data Sources

The method applied below to the five data sources is to illustrate the application of a CHI based on each source, then critically to assess them each for their internal reliability and sensitivity to different harm metrics. We then choose a single data source and show how it can be applied to Swedish crime trends and maps.

Findings

Without creating comprehensive indices, we have constructed the values for a CHI that are derived from five methods of scoring of a shortened list of the most common offences, as shown below in Table 1.

In Table 2, we have represented the differences between the three scales graphically. The minimum sentence appears to be most similar to sentencing data. One important difference, however, is that more than half of the crimes have the same value regarding minimum sentences. This small variability makes a CHI with such a basis less useful, as its purpose is to function as quantitative method to recognize differences.

Ruling out minimum sentences for this reason, we can then compare the average and maximum sentences. The maximum percent variance is higher for the average sentence whereas the standard deviation is highest for the maximum sentence. These differences are assumed to be negligible.

Figure 1 illustrates the differences presented in Table 2. The ratio between different crimes is nearly consistent between average and maximum sentences. However, for

Table 1 Results of five possible methods of creating a Swedish CHI score (days in prison)

| Offence type	Swedish translation	Judges estimate (Rinaldo 2015)	Sentencing data (Brå 2018)	Minimum sentence	Mean of max and min sentence	Maximum sentence
Murder (homicide)	Mord	6205	3819	3650	5110	6570
Manslaughter	Dråp	6205	1851	2190	2920	3650
Rape	Våldtäkt	1296	770	730	1460	2190
Armed robbery	Rån	657	589	365	1277	2190
Robbery	Rån	375	589	365	1277	2190
Arson	Mordbrand	1752	773	730	1825	2920
Assault (GBH)	Grov misshandel	711	627	365	1368	2190
Assault (ABH)	Misshandel	91	126	14	365	730
Stalking	Olaga förföljelse	No value	240	14	737	1460
Petty theft	Ringa stöld	14	30	14	81	182
Common assault	Ringa misshandel	30	No value	14	81	182
Threatening behaviour	Olaga hot	30	61	14	182	365
Fraud	Bedrägeri	No value	175	14	365	730
Criminal damage	Ringa skadegörelse	14	30	14	81	182
Harassment	Ofredande	30	30	14	182	365
Mean		1245	679	567	1117	1618
Standard deviation		2169	1034	1012	1386	1726
Variance		0.002	0.008	0.004	0.016	0.028
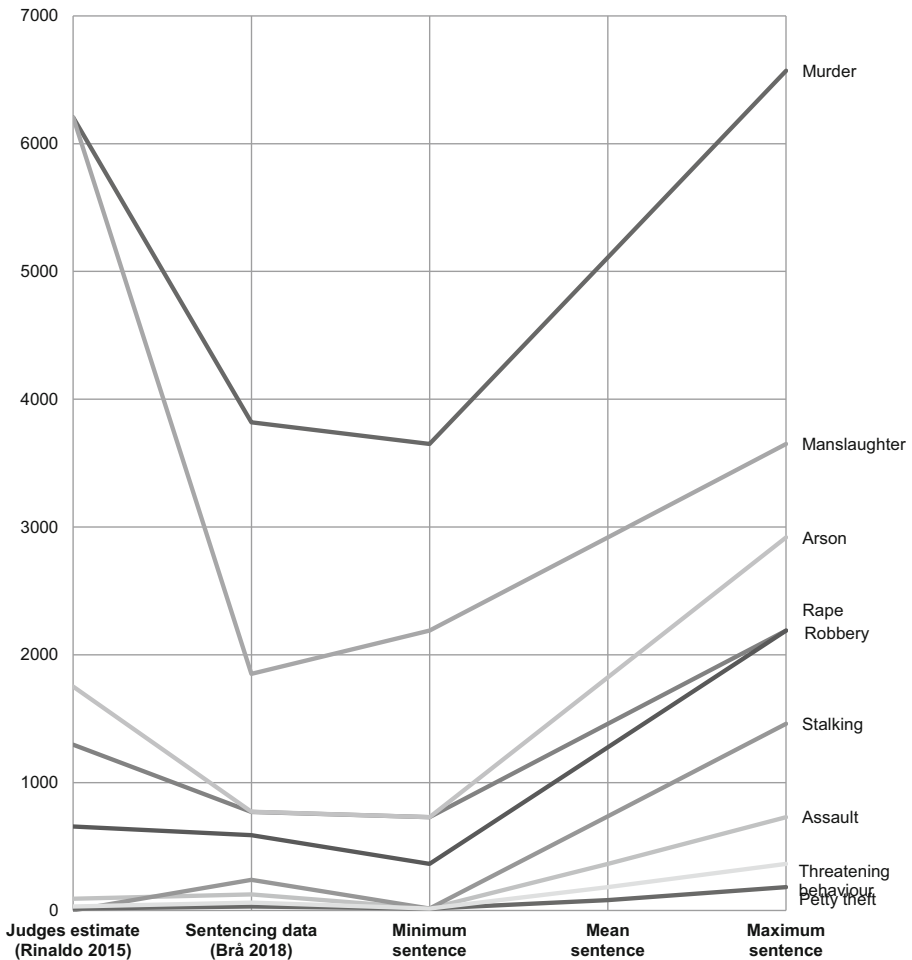
Fig. 1 Outcome of the various possible methods of creating a Swedish CHI—days of prison v offence types

technical reasons, we argue that the maximum sentences are preferable. The maximum sentences were also the method of choice by Mitchell (2017) in creating the California Crime Harm Index. Mitchell argued that: "While the maximum sentence an offender can receive in California is not the typical sentence, it is a consistent starting point for analyzing the relative harm of crimes across places (as well as across offenders and victims)." (p. 107).

Nevertheless, it has been argued that neither the minimum or maximum sentences compile a coherent system. For example, Ashworth (2007) criticize the British maximum sentences for lack of such logic. The maximum sentences have been set by parliament at different times and in different circumstances, without a general plan, on a "piecemeal basis". This is also true for Sweden and the Swedish Penal Code. It has recently been suggested in Sweden to introduce
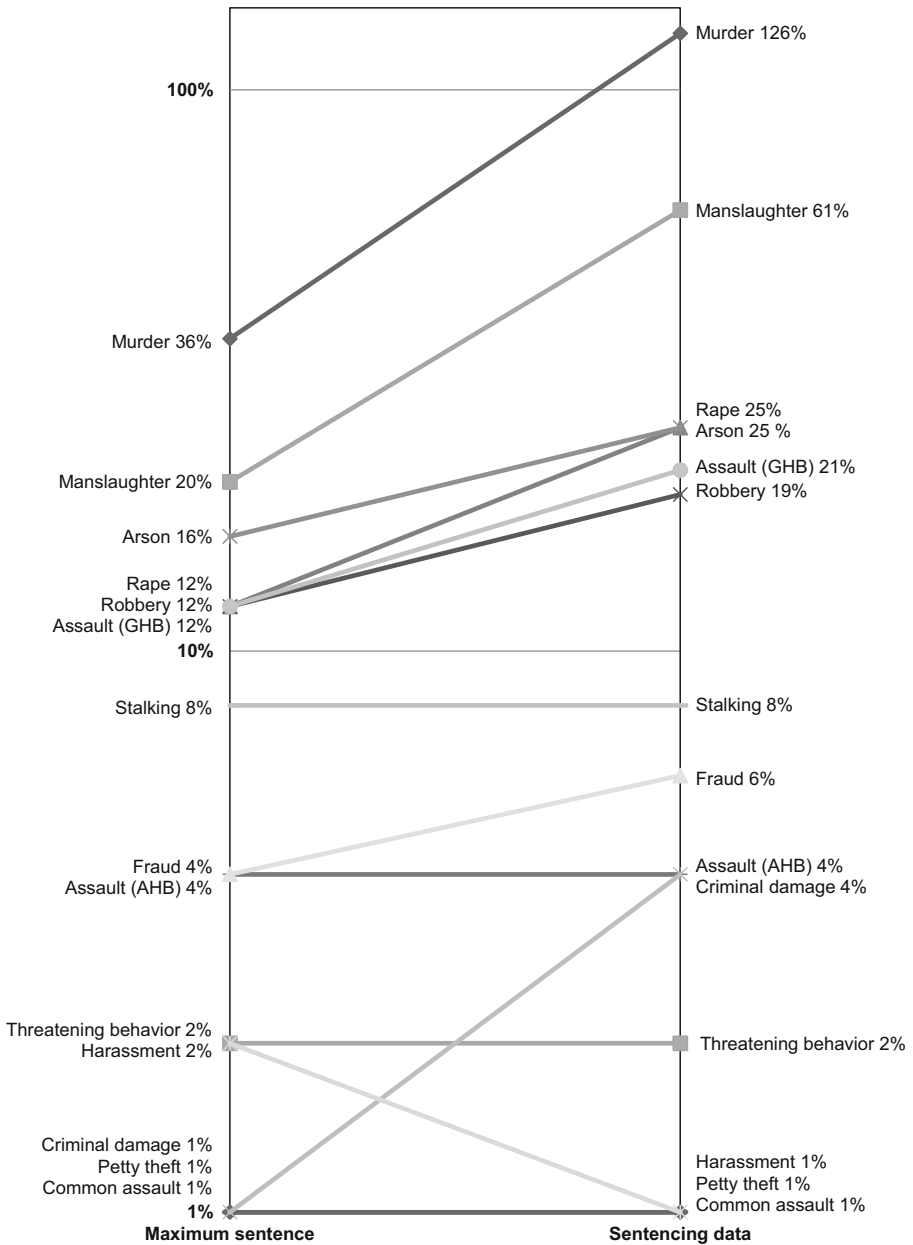
Fig. 2 Percentage difference: maximum sentence and sentencing data (logarithmic scale)

higher maxima for gross possessing of illegal firearm, as a response to the increasing number of gun crimes (Swedish Ministry of Justice 2019). On the other hand, one can argue that even though there is no overall rational plan for

sentences length, the minima and maxima still express the democratic valuation of the seriousness of different crimes.

Comparing judges' estimate and average sentences, there is no consistent ratio, which is apparent looking at Fig. 1. For some crimes, the judges apparently overestimated the typical sentences in comparison with the actual average sentence, according to sentencing data. Likely, the judges underestimate the regularity of mitigating circumstances, causing a lower average in the sentencing data, also close to the minimum sentences. This cannot be explained by the fact that the judges were instructed to estimate average sentences for first-time offenders, whereas the sentencing data includes all offenders. Such an explanation should result in lower values for all crimes, compared with sentencing data, which is not the case. Thus, it is likely that the judges in Rinaldo's study estimated incorrectly the average sentence for first-time offenders. This difference is a strong argument for choosing sentencing data ahead of judges' estimates.

In summary, we argue that in the Swedish context, the sentencing data has clear advantages compared with the judges' estimate. On the other hand, if the penal code should be used instead, maximum sentences appear preferable compared with minimum and average sentences. On balance, compared with our five criteria, of the possible alternatives, the most suitable method to create a CHI appears to be sentencing data or maximum sentences.

In Fig. 2 below, the variations between the maximum sentences and sentencing data are shown. One should note that maximum sentences generate significantly smaller variation width compared with sentencing data (36 v 126%[1]). The differences using the penal code are not realistic—neither for mirroring the relative severity nor the resources necessary for first response or investigation measures. It does not seem defensible to use a model in which a typical assault is equivalent to 10% of a murder in crime harm, nor in any other sense. Therefore, creating a CHI based on penalty scales, maximum, average or minimum, would be unsatisfactory. Notably, in Australia, researchers considered using the maximum sentences but rejected the method because of insufficient variability (Kwan 2016; House and Neyroud 2018).

Considering the above, we determined that the optimal method would be sentencing data. Furthermore, to ensure that the sentencing data is more reliable and less affected by disruptive fluctuations over time, we determined that we should use the average sentences available over at least the last 3 years.

Applications

We summarize the arguments made above in the matrix diagram below (Table 2), which shows our conclusion that the preferable option for creating a Swedish CHI is by using *sentencing data*.

---

[1] The slightly different values in Fig. 2 compared with the index in appendix is because Fig. 2 is based on sentencing data for only 1 year (2018) whereas the actual index is based on an average of the latest 3 years of sentencing data (2016–2018)

Table 2 A matrix diagram of conclusions about possible methods to create a CHI

Possible methods	Democratic	Reliable	Inexpensive	Valid	Operational
Judges estimate	Good	Good	Requires periodic surveys.	Risk for human errors by judges in estimating average.	Variance in test-index: 0,002
Sentencing data Mean	Good	Good	Good	Affected by offenders, aggravating/mitigating circumstances etc.	Variance in test-index: 0,008
Minimum sentences	Good	Good	Good	Good	Half of the crimes have the same minimum penalty.
Mean of Min & Max Sentences	Good	Good	Good	Good	Insufficient spread. Variance in test-index: 0,016
Max sentences	Good	Good	Good	Good	Insufficient spread. Variance in test-index: 0,028

The matrix above shows that the range of harm weighting is insufficient for three of five possible options. The maps below exemplify the importance of this: a CHI with low spread can be used to clearly identify micro harm spots for violent offences in public places (Maps 1 and 2).

Map 1 Crime harm spots, Stockholm City, 2018, regarding public violent crimes, weekend nights (22.00–04.00)
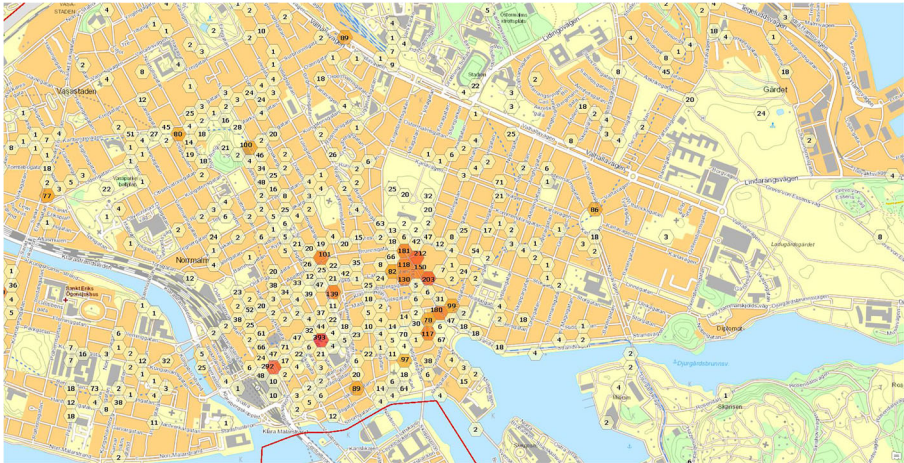
Map 2 Crime harm spots, Stureplan in Stockholm City, 2018, regarding public violent crimes, weekend nights (22.00–04.00)

In order for the maps to clearly show these harm spots, the index values need a sufficient spread. Insufficient spread cannot be the basis of a useful statistical tool for maps pointing out the most harmful places. In other words, with a lower variance, the index would result in a less distinct and discriminating output. Thus, the choice of method is of importance as it will affect the output and the capacity of the CHI to assist prioritization and targeting (Sherman 2013).

Figure 3 shows the immense impact a CHI can have for analysis. Comparison between reported crime and total crime harm, 1989–2018, per capita reveals that crime harm per capita has increased noticeably since 2006: three times more than an unweighted crime count. It is another example of how a CHI can be a valuable contribution to the public debate and political process, and can inspire fuller analysis about the development of crime.
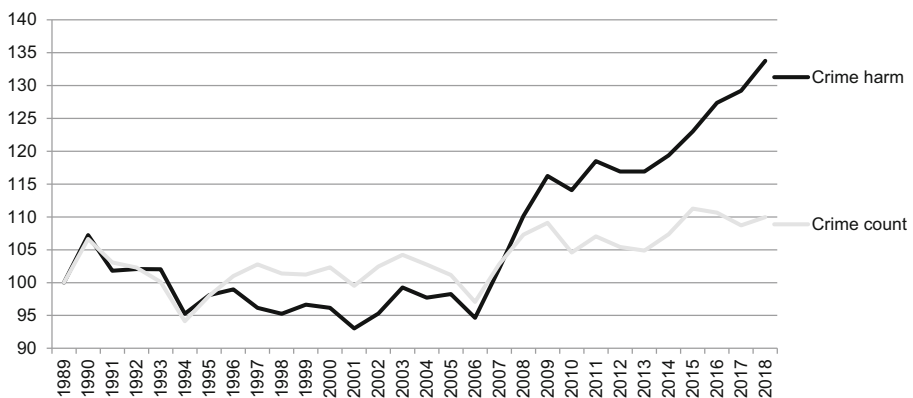

Fig. 3 Comparison between reported crime and total crime harm, 1989–2018, per capita. Crimes that results from self-initiated enforcement are excluded (traffic and narcotic crimes). Source: official statistics from The Swedish National Council for Crime Prevention and Statistics Sweden

Conclusion

Previous research clearly shows that using CHI has a great potential in further ways: distribution of cases between regions and districts, distribution of new police cadets, budget requests to the government, planning shift patterns and targeting of the most harmful offenders. The CHI can improve the precision of analyses supporting decisions about such issues, at national, regional and district levels. The proposed Swedish CHI can also be used by other authorities, political representatives of the Swedish people, the news media and other organizations. Altogether, these considerations emphasize the importance of a carefully conducted, transparent evaluation of possible methods for designing a CHI.

This study suggests that, whenever a CHI is designed, a thorough exploration of the possible methods should be done, in light of the legal system in question. The ambition should be to make a scientifically grounded and considered choice of the most suitable method. On that basis, we recommend to the government of Sweden the adoption of the method of average days of actual sentences, as computed from sentencing data in each crime category for a recent time period as the protocol for defining and revising the Swedish Crime Harm Index.

References

Andersen, H. A., & Mueller-Johnson, K. (2018). The Danish Crime Harm Index: how it works and why it matters. *Cambridge Journal of Evidence-Based Policing, 2*(1–2), 52–69.

Andersen H. (2018). How measurement matters: the creation of a Danish Crime Harm Index and its initial application to national police data. Unpublished M.St. Thesis in Applied Criminology and Police Management, University of Cambridge.

Ashby, P. (2018). Comparing methods for measuring crime harm/severity. *Policing: A Journal of Policy and Practice, 12*(4), 439–454.

Ashworth A. (2007). 'The mechanics of sentencing'. Maguire M., Morgan R. and Reiner R. (Eds.) *The Oxford handbook of criminology*. New York: Oxford University Press.

Bland, M., & Ariel, B. (2015). Targeting escalation in reported domestic abuse: evidence from 36,000 callouts. *International Criminal Justice Review, 25*(1), 30–53.

Brottsförebyggande rådet (2018). Kriminalstatistik 2017. **Available at** <https://www.bra.se/download/18.c4ecee2162e20d258c246fc/1527751812077/Sammanfattning_lagforda_2017.pdf> ().

Chong O. (2019). Developing a Crime Harm Index for Hong Kong: measuring crime severity as well as crime volume. Master thesis University of Cambridge, Institute of Criminology.

Curtis-Ham, S., & Walton, D. (2017). The New Zealand Crime Harm Index: quantifying harm using sentencing data. *Policing, 1*(1), 1–13.

Gorsuch, J. H. (1938). Scale of seriousness of crimes. *Am. Inst. Crim. L. & Criminology, 29*, 245–252.

House, P. D., & Neyroud, P. W. (2018). Developing a crime harm index for Western Australia: the WACHI. *Cambridge Journal of Evidence-Based Policing, 2*(1–2), 70–94.

Ignatans, D., & Pease, K. (2015). Taking crime seriously: playing the weighting game. *Policing, 10*(3), 184–193.

Kleinig, J. (1978). Crime and the concept of harm. *American Philosophical Quarterly, 15*(1), 27–36.

Koch, R. (2007). *The 80/20 principle: the secret of achieving more with less* (2nd ed.). London: Nicholas Brealey.

Kwan, L. (2016). Western Australian maximum sentence values compared with the Cambridge Crime Harm Index. Internship report. Western Australia Police Force.

Law Commission. (1989). Criminal law: a criminal code for England and Wales. In *Report and Draft Criminal Code Bill* (Vol. 1). London: HMSO.

Mill, J.S. [1859], 1974. *On liberty.* Edited by Himmelfarb G. New York: Penguin.

Mitchell, R. J. (2017). The usefulness of a crime harm index: analyzing the Sacramento Hot Spot Experiment using the California Crime Harm Index (CA-CHI). *Journal of Experimental Criminology, 15*(1), 103–113.

Neyroud, P. W. (2017). The ethics of learning by testing: the police, professionalism and researching the police. In M. Cowburn, L. Gelsthorpe, & A. Wahidin (Eds.), *Research ethics in criminology: Dilemmas, issues and solutions.* London: Routledge.

Neyroud, P.W. (2019) De-policing and homicide: applying evidence and ethics to the "de-policing" debate. Criminology and Public Policy, Published online at https://doi.org/10.1111/1745-9133.12425.

Office of National Statistics (2016). Crime Severity Score. https://www.ons.gov. uk/peoplepopulationandcommunity/crimeandjustice/datasets/crimeseverityscoreexperimentalstatistics ().

Paoli, L., & Greenfield, V. (2013). A neglected concept in criminology, a necessary benchmark for crime-control policy. *European Journal of Crime, Criminal Law and Criminal Justice, 21*, 359–377.

Pol, R. (2016). 'Revisiting crime rates as a measure of crime prevention effectiveness: does the 'crime drop' reveal a policy effectiveness 'outcomes' gap?'. Student Pulse 8 (2) Available at: http://www.studentpulse. com/a?id=1350 ().

Ratcliffe, J. H. (2015a). Towards an index for harm-focused policing. *Policing, 9*(2), 164–183.

Ratcliffe, J. H. (2015b). *Harm-focused policing. Issue 19 of Ideas in American Policing.* Washington DC: Police Foundation.

Rinaldo, M. (2015). *Comparing crime hotspots and crime harm-spots in a Swedish City: a descriptive analysis. Master thesis.* Cambridge: University of Cambridge.

Reed, W. J. (2001). The Pareto, Zipf and other power laws. *Economics Letters, 74*(1), 15–19.

Sellin, T. and Wolfgang, M. (1964). The measurement of delinquency. Montclair: Patterson Smith. (Reprinted, 1978, with introduction by S. Turner).

Sherman, L. W. (2007). The power few: experimental criminology and the reduction of harm. *Journal of Experimental Criminology, 3*(4), 299–321.

Sherman, L. (2011). 'Al Capone, the sword of Damocles, 30 and the police-corrections budget ratio.' *Criminology & Public Policy* 10: 195–206.

Sherman, L. (2013). The rise of evidence-based policing: targeting, testing and tracking. *Crime and Justice, 42*, 377–343.

Sherman, L. W., Neyroud, P., & Neyroud, E. (2016). The Cambridge Crime Harm Index: measuring total harm from crime based on sentencing guidelines. *Policing, 10*(3), 171–183.

Sparrow, M. (2008). *The character of harms: operational challenges in control.* Cambridge: Cambridge University Press.

Statistics Canada (2015). The Crime Severity Index. http://www.statcan.gc.ca/pub/85-004-x/2009001/part-partie1-eng.htm ().

Swedish Ministry of Justice (1999a). The Swedish penal code: a summary. Ds 1999:36. Stockholm: Ministry of Justice.

Swedish Ministry of Justice (1999b). *En strängare syn på vapenbrott och smuggling av vapen och explosiva varor.* Ds 2019:14. Stockholm: Ministry of Justice.

Taira, Y. (2018). *Japan Crime Harm Index: the application of the Cambridge Crime Harm Index to official crime statistics in Japan. Master thesis.* University of Cambridge, Institute of Criminology.

Träskmark P. O. (2003). 'Sweden'. L'harmonisation des sanctions pénales en Europe Unité mixte de recherche de droit comparé de Paris 5: 185–198.

Wallace, M., Turner, J., Babyak, C. and Matarazzo, A. (2009). *Measuring crime in Canada: Introducing the Crime Severity Index and improvements to the Uniform Crime Reporting survey.* Statistics Canada Catalogue no. 85-004-X. Ottawa.

Wolfgang, M., Figlio, R. M., Tracy, P. E. and Singer, S. I. (1985). The National Survey of crime severity. Washington, 50 DC: US Department of Justice, Bureau of Justice Statistics.

Zipf, G. K. (2016). *Human behavior and the principle of least effort: an introduction to human ecology.* Ravenio Books.

Fredrik Karrholm is an inspector at the Swedish National Police Commissioner's office whose research on this article was presented for his M.St. thesis in applied criminology and police management at the University of Cambridge.

Peter Neyroud , C.B.E. Q.P.M., MA (Oxon), PhD (Cantab), is Lecturer in Evidence-Based Policing and Director of the senior leader apprenticeship for the Police Executive Programme, University of Cambridge.

John Smaaland is an information technology specialist at the IT Department at the Swedish Police Authority.

Affiliations

Fredrik Kärrholm[1] · Peter Neyroud[2] · John Smaaland[1]

Fredrik Kärrholm
fbk21@cam.ac.uk

[1]   Swedish National Police Authority, Stockholm, Sweden

[2]   Institute of Criminology, University of Cambridge, Cambridge, England