How to Obtain a More Accurate Picture of Crime through Crime Statistics: Proposals and Methods

By Mathias Bug and Kristina Meier

This report aims to show the distribution of crime in Germany. For this, police crime statistics (polizeiliche Kriminalstatistik – PKS) are treated so that they integrate dark figures (unreported crime) of crime types along with their specific grade of burden. The different treatments are based on own recent survey data. Two major trends are confirmed by both treatment methods: First, there is a north-south divide, with the northern regions experiencing a far higher risk of crime. Second, rural/urban differences can be accounted for by the higher levels of everyday crime that affect citizens of towns and cities. These slight differences and changes are more evident and meaningful in state-to-state comparisons rather than looking at Germany as a whole.

To date, police crime statistics (PKS) in Germany, compiled by the Federal Criminal Police Office (Bundeskriminalamt, BKA), have taken the form of a list of the main categories of offenses with relevant figures for each one. In this form, the distribution of offenses across Germany is depicted at federal state level only (as well as by city with populations exceeding 200,000), and the overall picture consists of little more than total raw case numbers.

For some time now, however, both the economic research and international criminological research community have been asking the same question: whether the method applied hitherto — mere totaling of individual criminal offenses reported to the police — enables the social burden from crime to be analyzed at all. In light of this, the present paper will outline possible methods of combating the two main criticisms leveled at police crime statistics in Germany — i.e., the missing

1 This article was written as part of the research project WISIND, an economic security indicator for Germany. It is financed by the German Federal Ministry of Education and Research as part of their “Social Dimensions of Security Research” funding line. The WISIND project and WISIND data were developed in close cooperation between Martin Kroh, Johannes Riekmann, Eric van Um, Nina Wald, and Nathan Fiala. The authors would also like to thank Enrique Fernandez, Martina Kraus, Jan-Lucas Schanze, and Bartosz Walenda for their support throughout this process. Excerpts from this Economic Bulletin report can be found in M. Bug and K. Meier, “Herausforderungen bei der Messung von Kriminalität,” DIW Roundup (2014), accessed December 2, 2014, http://www.diw.de/sixcms/detail.php?id=diw_01.c.466936.de (in German only).


4 An overview of the status quo in research on crime rate assessment can be found in S. Eifler and D. Pollich, eds., “Empirische Forschung über Kriminalität” (2015). A brief overview can be found in Bug and Meier, “Herausforderungen.”
es at administrative district level. This approach allows us to see urban/rural differences, as well as differences between individual rural districts when comparing crime statistics (see Figure 1).

The dark figure correction did not lead to major changes in the measured crime burden.

In order to take account of the very heterogeneous population distribution across the individual German federal states and in Germany as a whole, the considerations here are based on an aggregate of criminal offenses at administrative district level. This approach allows us to see urban/rural differences, as well as differences between individual rural districts when comparing crime statistics (see Figure 1).

Crime Risk Assessment — Gap between Reported and Non-Reported Crime (the “Dark Figure” of Crime)

The central database used to calculate crime rates is the German Police Crime Statistics, which includes data on the number of attempted and actual crimes reported to the police in the given reference period.

The problem with police crime statistics in Germany, however, is — as mentioned above — they only include officially reported criminal acts. Figure 1 shows the aggregate distribution of reported offenses under “Non-Weighted Aggregate (without dark-figure adjustment).” For certain forms of crime, evidence of considerable gaps between reported and unreported offenses exist. Errors on the part of law enforcement agencies also play a role

---


7 Birkel, Viktimisierungssurvey.
Germany, the authors of this report suggest adjusting the official crime statistics by a calculated estimate of the “dark figure of crime.” This approach is based on dark figure studies on victimization experiences conducted among the German population as part of large-scale representative surveys. Even in dark figure studies with large sample sizes, however, the approach used will admittedly involve a certain degree of subjectivity since such studies can ultimately only record whether and how people recall a criminal act, as well as what they are willing to relate in the survey situation.13 Another problem with such surveys is the difference between the legal definitions used in official crime statistics and the common understanding of certain forms of crime that are typical of such surveys. Consequently, adjusting crime statistics to factor in the dark figure is somewhat subjective since the adjustment factor is formed on the basis of the subjective perception of victimization (and not on the basis of police reports or even court findings).

**Crime Risk Assessment — The Problem with Crime Statistics**

To obtain a more accurate picture of the actual offenses committed in the various administrative districts in Germany, the authors of this report suggest adjusting the official crime statistics by a calculated estimate of the “dark figure of crime.” This approach is based on dark figure studies on victimization experiences conducted among the German population as part of large-scale representative surveys. Even in dark figure studies with large sample sizes, however, the approach used will admittedly involve a certain degree of subjectivity since such studies can ultimately only record whether and how people recall a criminal act, as well as what they are willing to relate in the survey situation.13 Another problem with such surveys is the difference between the legal definitions used in official crime statistics and the common understanding of certain forms of crime that are typical of such surveys. Consequently, adjusting crime statistics to factor in the dark figure is somewhat subjective since the adjustment factor is formed on the basis of the subjective perception of victimization (and not on the basis of police reports or even court findings).

**Crime Statistics in Germany**

The following offenses form the data basis of the approach described here: theft (PKS Index ****00 without 440*00), burglary (PKS Index 435*00 and 436*00, as well as 440*00), bodily harm (PKS Index 222000 and 224000), murder and manslaughter (PKS Index 892500).1 The results shown in this article are based on what is referred to as the frequency of offense. These are calculated according to the formula

\[
\text{Absolute number of offenses} \times 100000 = \frac{\text{Number of inhabitants}}{}.
\]

The method used here incorporates all the PKS data collected from 2010 to 2013.

The above-mentioned offenses essentially cover crimes that have a direct impact on the individual and the everyday context and consequently affect the subjective perception of personal security; this type of crime is referred to as everyday crime.2 In the period July through September 2014, as part of the WISIND project, opinion poll company TNS Emnid conducted a representative telephone survey among 12,094 individuals in Germany, who were all asked about their personal experience with crime. Twenty percent of respondents were interviewed on cellphone numbers. The sample is a proportionally representative sample distributed evenly across Germany, with a minimum of 15 participants in each administrative district. A further representative online survey conducted by research institute forsa asked 2,532 people to rank the severity of different types of crime.3 The results of these two surveys form the main basis for further calculations using the PKS data shown below.
Crime Statistics

CRIME STATISTICS

The averages, which serve to incorporate the “dark figure” of crime, are calculated uniformly at national level in order to fully utilize the explanatory power of the entire sample. The resultant distribution of observed offenses is shown in Figure 1 under “Non-Weighted Aggregate” (with dark-figure adjustment).13

The averages, which serve to incorporate the “dark figure” of crime, are calculated uniformly at national level in order to fully utilize the explanatory power of the entire sample. The resultant distribution of observed offenses is shown in Figure 1 under “Non-Weighted Aggregate” (with dark-figure adjustment).13

Table 1

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Murder and manslaughter1</td>
<td>1.8285</td>
<td>0.039</td>
</tr>
<tr>
<td>Burglary</td>
<td>5.565</td>
<td>0.128</td>
</tr>
<tr>
<td>Theft</td>
<td>2.937</td>
<td>0.128</td>
</tr>
<tr>
<td>Bodily harm</td>
<td>4.047</td>
<td>0.721</td>
</tr>
</tbody>
</table>

1 Calculated on base of a Germany wide study about autopsy mistakes (Brinkmann 1997).


Especially the dark figures for burglary and bodily harm indicate a substantial share of unreported crime.

In addition to the problems related to reported versus unreported crime, when it comes to measuring crime itself there is another issue with the impact crime has on society. Adding up the total number of incidents and weighting them equally is unlikely to give a true indication of the actual risk that crime poses. Instead, what is called for is suitable weighting indexes for individual criminal offenses. Indexes of this type have already been published in other countries, such as the US or UK Peace Index.16 Both of these are based on five key indicators, appropriately weighted and aggregated. The shortcoming of this type of weighting, however, is its subjectivity. This is not the only conceivable approach, however. Indeed, various other methods are already being used in criminological research.17 The calculations presented here are essentially derived from the concepts behind these methods. For reasons of comparison, the crime risk indicator is calculated using various weighting methods, which are briefly presented below.

Monetization Weighting

In economic research literature, various monetization approaches exist.18 Here, the costs resulting from different crimes are estimated, taking into account the financial loss (e.g., due to incapacity to work, treatment costs), as well as — where possible — the emotional impact.

These quantifications enable relative degrees of severity to be calculated. Here, the estimated total damage caused by each offense is compared to the most serious offense (homicide) and the quotient used as the weighting factor.

Opinion-Based Weighting Using a Representative Public Opinion Poll

On the basis of a representative online survey among 2,532 respondents, degrees of severity were calculated in order to categorize individual offenses. Respondents were asked

14 In addition, for reasons of demographic change, the low participation of elderly people who are in need of care is becoming increasingly relevant.

15 Birkel, Viktimisierungssurvey 31; Birkel, “Hellfeld versus Dunkelfeld.”


to rank ten different offenses according to their relative degree of severity. This was done in two stages: first, respondents were asked to order the offenses according to severity, starting with the most minor (misdemeanor). The second step was to take the ranking from step one and compare the offenses in pairs. The resultant individual weightings were averaged across the entire sample to produce indicator weights for the respective types of crime.

Data-Based Weighting

The third and least subjective approach is purely data-based and uses Item Response Theory (IRT), which has also been used to assess pupil performance in PISA studies, as well as to measure corruption and democracy indexes.

The basic concept behind this method is that it attempts to estimate a latent variable (in this case, the crime risk level) using the severity of various subindicators (here, items). In other words, using the relevant data, the severity of the offense and its relevance for latent risk is ascertained. IRT is essentially the same as calculating a weighted aggregate from individual offense aggregates. However, unlike the latter, IRT does not need a priori assumptions regarding the severity of offenses, but estimates the severity and relevance based on the data itself using an accepted and reconstructible method. In addition, IRT allows standard errors to be calculated, meaning conclusions can be drawn on the statistical significance of the estimated values. Table 2 shows the weighting that results from the different approaches. The coefficients from the IRT calculation are not to be interpreted in the same way, which is why they are not mentioned explicitly here. They are, however, included in the calculations below.

In all of the weighting methods, the different criminal offenses are placed in the same order of priority. As expected, murder and manslaughter are seen as the most severe. An interesting fact, however, is that homicide comes last in the IRT method with regard to relevance to crime risk. This may be because such extremely rare incidents are likely to be randomly spread across the country, meaning they are an unreliable indicator of the crime risk.

Interpreting the Results

Figures 1 and 2 show the main crime risk by administrative district for the period 2010 to 2013, calculated on the basis of the raw PKS data, the figures adjusted for unrecorded crime, and the weighting methods presented here. To make the results more readily comparable, the indicator values for the given weighting methods were all normalized to lie within the interval [0,1]. All the weighting methods show a similar picture with very few surprises. The difference between rural and urban areas is very apparent, and, as expected, the crime risk level far higher in urban regions. This finding clearly shows the need for a more differentiated approach to measuring the crime risk at district level, since these trends are difficult to evaluate if the results are compared on state-level. With the exception of Munich,urbans also display higher values. It should be noted, however, that the low crime risk seen in Munich is very much in line with the general north-south divide, one of the issues discussed in the PKS yearbooks for the period under observation. A slight change in this north-south divide is evident for the monetization method only, with everyday crime being less problematic in Brandenburg, Saxony, and Westfalen in particular. In some parts of Thuringia, Upper/Middle Franconia, and Upper Bavaria, the monetization method revealed a relatively high crime risk.

<p>| Table 2 |</p>
<table>
<thead>
<tr>
<th>Weights of specific crime types</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Murder and manslaughter</td>
</tr>
<tr>
<td>Burglary</td>
</tr>
<tr>
<td>Theft</td>
</tr>
<tr>
<td>Bodily harm</td>
</tr>
</tbody>
</table>


Both ways of weighting show the same ranking in between crime types.
Similarly, the much-bemoaned higher crime rates in the border regions to Poland could not be observed across all of the weighting methods. Indeed, the results of monetization and opinion-based weighting showed nothing to confirm this.

The similarity between the different methods may seem surprising at first glance, especially given the dissimilar weighting given to the various types of offense (see Table 2). Owing to the distinct frequencies of individual offenses, these differences are almost negligible in the bigger picture. In the monetization weighting method, for example, the ratio of murder/manslaughter to theft is 1:0.0004; in relation to the weighted aggregates, this difference is less significant owing to the high frequency with which offenses such as theft occur (in 2013, a total of 2,379,091 incidents involving theft were reported in Germany, compared to as few as 2,119 cases of murder and manslaughter).

In other words, the results are affected most by those forms of crime that occur most frequently. This makes

---

**Figure 2**

**Burden through everyday crime — three ways of weighting**


All three ways of weighting show similar distributions to the basic police statistics. However, the monetarisation process slightly dissolves the north south slope.
CRIME STATISTICS

Figure 3

Development of the everyday crime burden 2010-2013 (weighted calculation)


The survey based weighting shows — in addition to the dark figure corrected but unweighted version — the strongest changes in the regional burden of everyday crime.

intuitive sense, since it is these offenses that produce situations of constant risk and less so crimes such as murder and manslaughter which, for all their severity, are few and far between.

Crime Development from 2010 to 2013

Since the early 1990s, crime has been on the decline, with individual rates ultimately stabilizing — a positive development, particularly with regard to the fight against crime on the political stage. Figure 3 shows the corresponding changes in crime rates measured between 2010 and 2013. Despite being slight on average, the changes observed are still very pronounced in some regions.

For each administrative district, the figures were calculated as the difference between 2010 and 2013 relative to the average value for the entire period $\frac{\text{No. of cases}_{2013} - \text{No. of cases}_{2010}}{\text{Average no. of cases}_{2010-2013}}$. The results obtained were then divided into five categories. Differences exceeding two standard deviations were classified as a
An interesting finding is the near parallel results for crime development for the adjusted non-weighted crime figures, as well as for the opinion-based weighted crime statistics. In both cases, clusters of increased crime are evident in the regions of Brandenburg/Saxony, Thuringia / Upper/Middle Franconia, and Upper Bavaria as well as in parts of Baden-Württemberg.

Conclusion

The present report looks at two key shortcomings of Police Crime Statistics in Germany. First, an attempt to mitigate the problem with unreported crime (the "dark figure of crime") was made using a victimization survey conducted by DIW. Here, a comparison of the results before and after dark-figure adjustment revealed no major differences in crime distribution. In addition, owing to criticism made of the PKS that it amounts to no more than aggregate non-weighted figures, the next step was to examine the impact of alternative weighting methods on crime risk assessment.

For all of the methods, two crime risk trends can be observed: first, a north-south divide, with the northern regions showing a "clear increase" (2), while a slight increase denotes a difference in the region of 1 to 2 standard deviations (1). By analogy, the categories "clear decrease" (−2) and "slight decrease" (−1) are formed on the basis of negative standard deviations. The category "No change" (0) refers to those values that lie between a negative and a positive standard deviation.

With regard to crime risk development, all of the measurement methods presented here give more or less a similar picture. For most administrative districts, no or very little change was seen throughout the period of observation — official crime statistics (PKS) showed the same for the overall development of crime during the same period. Nevertheless, the three weighting methods displayed minor changes in the robustness of their categorization. While the IRT showed changes for a small number of districts only, the monetization method showed clearer differences in both directions (increase and decrease). The opinion-based weighting method, for its part, showed more districts to have higher crime rates. An interesting finding is the near parallel results for crime development for the adjusted non-weighted crime figures, as well as for the opinion-based weighted crime statistics. In both cases, clusters of increased crime are evident in the regions of Brandenburg/Saxony, Thuringia / Upper/Middle Franconia, and Upper Bavaria as well as in parts of Baden-Württemberg.

Note that, over time, the results of the IRT analysis are not directly comparable with the other two weighting methods, since the weighting parameters are re-estimated from the data provided each year using the IRT method.
regions displaying a higher risk of crime. Second, rural/urban differences, which can be accounted for by the higher levels of everyday crime affecting the population of towns and cities. Although the different weighting methods produce largely similar results, slight differences and changes are evident and are more meaningful in a state-to-state comparison rather than in a country-wide context.

In light of this, it will be all the more interesting to see what picture the crime risk indicators for 2012 and 2013 will paint. In these two years, the forms of crime also include Internet crime, personal threats, and violent extremism. In this context, detailed findings can be expected, particularly with regard to the urban/rural gap. This may even give a better insight into daily commuting between the city and the countryside.

Moreover, with the data findings of the DIW research project presented here, a subjective crime risk indicator representing people’s fear of crime can be developed and compared with objective crime rates.

Mathias Bug is Research Associate in the Department of Development and Security of DIW Berlin | mbug@diw.de

Kristina Meier is Research Associate in the Department of Development and Security of DIW Berlin | kmeier@diw.de

JEL: K14, R19, H56, H77
Keywords: security, crime, indicator, home affairs, police, federalism, inner security, threat, crime statistics, police statistics