

Crime in Germany



REPORT by Mathias Bug, Martin Kroh and Kristina Meier

Regional Crime Rates and Fear of Crime— WISIND Findings

167

INTERVIEW with Martin Kroh

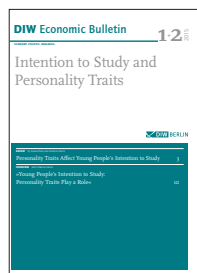
»Regional Crime Rates and Fear of Crime: Distinct North-South Divide— Unified Measurement Methods Needed«

177

REPORT by Johannes Rieckmann and Jan-Lucas Schanze

Perceptions of Personal Security in Social Media and Search Engines— a Realistic Reflection of Actual Crime Rates?

178



The DIW Economic Bulletin contains selected articles and interviews from the DIW Wochenbericht in English. As the institute's flagship publication, the DIW Wochenbericht provides an independent view on the economic development in Germany and the world, addressing the media as well as leaders in politics, business and society.

The DIW Economic Bulletin is published weekly and available as a free download from DIW Berlin's website.

THE NEWSLETTER FROM THE INSTITUTE



The DIW Newsletter in English provides the latest news, publications and events from the institute every two weeks. Furthermore we offer 'New Issue Alerts' for the DIW Economic Bulletin and the DIW Roundup.

>> Subscribe to DIW Newsletter in English at: www.diw.de/en/newsletter

Publishers

Prof. Dr. Pio Baake
Prof. Dr. Tomaso Duso
Dr. Ferdinand Fichtner
Prof. Marcel Fratzscher, Ph.D.
Prof. Dr. Peter Haan
Prof. Dr. Claudia Kemfert
Dr. Kati Krähnert
Prof. Karsten Neuhoff, Ph.D.
Prof. Dr. Jürgen Schupp
Prof. Dr. C. Katharina Spieß
Prof. Dr. Gert G. Wagner

Reviewer

Dr. Anika Rasner
Dr. David Richter

Editors in chief

Sabine Fiedler
Dr. Kurt Geppert

Editorial staff

Renate Bogdanovic
Andreas Harasser
Sebastian Kollmann
Dr. Claudia Lambert
Dr. WolfPeter Schill

Translation

HLTW Übersetzungen GbR
team@hlw.de

Layout and Composition

eScriptum GmbH & Co KG, Berlin

Press office

Renate Bogdanovic
Tel. +49-30-89789-249
presse@diw.de

Sale and distribution

DIW Berlin

Reprint and further distribution – including extracts – with complete reference and consignment of a specimen copy to DIW Berlin's Communication Department (kundenservice@diw.berlin) only.
Printed on 100 % recycled paper.

NEXT WEEK IN DIW ECONOMIC BULLETIN

Uncharted Territory: Large-scale Asset Purchases by the European Central Banks

Regional Crime Rates and Fear of Crime — WISIND Findings

By Mathias Bug, Martin Kroh and Kristina Meier

Many people are afraid of falling prey to crime. The present report investigates the extent to which this fear is in line with the actual regional crime rates. This analysis is based on data from a comprehensive database on the fear of crime, combined with police crime statistics (specifically, adjusted crime statistics which factor in the “dark figure” of unreported crime). No evidence was found to support the (occasionally voiced) contention that the fear of falling prey to crime is irrational in many cases and not representative of the actual level of safety within a given region. In fact, our data shows a clear statistical correlation between regional crime rates and the fear of crime, both of which are more pronounced in the north of Germany than in the south, for instance. The inclusion of cybercrime in crime statistics, however, has meant that the former, higher crime rates and greater fear of crime often recorded in urban areas as opposed to rural regions are no longer as pronounced.

The¹ fear of becoming the victim of a crime is a widely known phenomenon, which, as psychological studies show, will often impair the quality of life of the person affected.² The fear of crime and perceived or subjective safety, however, are occasionally discussed in connection with the irrationality of diffuse fears, media hysteria, or general social insecurity. In fact, studies demonstrate that certain groups — different age groups, for example — miscalculate the probability of their falling prey to crime.³ Similar presumptions are made regarding regional differences: the fear of crime is believed to be high even in regions where in fact there is very little crime.

When statistical data on regional crime rates and fear of crime are compiled, two important questions arise: first, what types of crime are to be included and, second, what is the relative weighting between the different crime categories, the latter occasionally being referred to as the severity of the offense. In regional crime statistics, pick-

¹ The report was compiled as part of the research project *A System of Economic Indicators to Measure Security and Security Provision in Germany* (WISIND), which is financed by the German Ministry for Education and Research as part of the *Social Dimensions of Security Research* funding program. The idea behind the WISIND project and the generation of WISIND-specific data was jointly developed by Johannes Rieckmann, Eric van Um, and Nina Wald. The authors would also like to thank Enrique Fernandez, Martina Kraus, Jan-Lucas Schanze, and Bartosz Walenda for their assistance throughout this process. The present report is based on DIW Economic Bulletin 3/2015, in which the different methods used here are described in detail using a newly-generated crime indicator for the years 2010–2013. See M. Bug and K. Meier, “How to Obtain a More Accurate Picture of Crime Risk through Crime Statistics – Proposals and Methods,” DIW Economic Bulletin 3 (2015) http://www.diw.de/documents/publikationen/73/diw_01.c.494746.de/diw_econ_bull_2015-03-1.pdf, accessed on February 19, 2015.

² U. Dörmann and M. Remmers, “Sicherheitsgefühl und Kriminalitätsbewertung,” in BKA Polizei + Forschung, vol. 1 (2000): 1–2. F. DuBow, E. McCabe, and G. Kaplan, “Reactions to Crime: A Critical Review of the Literature,” unpublished report (Center for Urban Affairs, Northwestern University, 1979): 93–99 based on J. Garofalo, “The Fear of Crime: Causes and Consequences,” *Journal of Criminal Law and Criminology*, vol. 72, issue 2 summer, article 20 (1981): 852 f.

³ D. Ziegler, D. Kudlacek, and T. Fischer, “Zur Wahrnehmung und Definition von Sicherheit durch die Bevölkerung. Erkenntnisse und Konsequenzen aus der kriminologisch-sozialwissenschaftlichen Forschung,” *Sicherheit, Forschungsforum Öffentliche Sicherheit papers* (Freie Universität Berlin, 2012): 25, http://www.sicherheitforschung.de/schriftenreihe/sr_v_v/sr_5.pdf, accessed on November 11, 2014.

pocketing, which is a minor offense, is ascribed less significance than robbery, for instance.

The present report is limited to forms of crime that have a direct impact on individuals and consequently excludes white-collar crimes, for example. While the assignment of specific weighting to individual offenses is instrumental in evaluating overall crime rates and the fear of crime, there are many weighting methods which are equally legitimate. We have therefore chosen to use four different approaches, the results of which will then be compared.⁴

Objective Crime Statistics and Subjective Fear of Crime

Here, objective regional crime is defined as offenses committed, the victims being citizens of the said region. The following offenses are included in the objective crime statistics: burglary, theft, cybercrime, criminal threat, bodily harm, homicide, and politically motivated crime. These offenses were specifically selected to enable the indicator to be used to measure crime that directly affects the individual. The study is based on data from the German police crime statistics (*polizeiliche Kriminalstatistik, PKS*) and, to include politically motivated crime, the Annual Reports of the intelligence services of the German *Länder*. Both databases contain recorded crime only, meaning they do not paint a fully accurate picture of the actual crimes committed. To ensure that the “dark figure” of crime (i.e., the number of offenses that are not officially reported) is also factored in, the figures for the individual offenses taken from the *PKS* and the protection of the constitution reports are adjusted by an offense-specific unreported crime factor (see Box 1).

The subjective fear of crime refers to the fear among a given regional population of becoming the victim of a criminal offense. To enable objective crime statistics to be reliably compared with the subjective fear of crime, the offenses selected to assess the fear of crime largely conform to the aforementioned forms of crime used in objective crime statistics.

Unlike objective crime rates, which are often based on the *PKS*, no reliable database is available for the subjective fear of crime. However, to ensure that findings on the fear of crime are both generalizable and cover as broad an empirical basis as possible, existing survey data is used, as well as a study of our own which was conducted as part of the WISIND project and examines victimization and the fear of crime among the population; in addition, data on

private expenditure on security⁵ and data from social networks were employed.⁶ Although each individual means of evaluating the fear of crime has its own inherent problems — for example, the case numbers in population surveys are often insufficient to allow us to draw meaningful regional conclusions, and information taken from social networks reflects the fear of crime for a specific population group only — the sheer variety of the data sources is intended to offset the individual shortcomings.

Discussions in relevant research communities have long since centered on survey-based analysis of the fear of crime.⁷ Owing to the abstract nature of fear, however, no consensus has been reached to date on how to evaluate this fear in surveys. One common question is that of the individual’s feeling of safety during a late-night walk through one’s neighborhood. (The approximate wording is as follows: “How safe do you feel when walking through your neighborhood alone at night?”) Despite the criticism leveled at this indicator for the fear of crime,⁸ it continues to be used for this purpose, not least because it allows different studies to be easily compared. Since it is incorporated into the Germany-wide WISIND survey, this question is also included in the present study. Another data source used is the Socio-Economic Panel (SOEP) study, a large panel study of households in Germany, which also includes a general question on the fear of crime (“How worried are you about the following areas?: [...] The trends observed in crime in Germany.”)^{9,10}

⁵ C. Gummer et al., „BIGS-Studie Sicherheitswirtschaft 2012“ (2013). Y. Gruchmann, et al., „BIGS-Studie Sicherheitswirtschaft 2013“ (2014).

⁶ in this economic bulletin See J. Rieckmann and J.-L. Schanze, “Perceptions of Personal Safety in Social Media and Search Engines – a Realistic Reflection of Actual Crime Rates” (2015).

⁷ For an overview see Ziegler, Kudlacek, and Fischer, “Zur Wahrnehmung.”

⁸ A discussion of relevant literature can be found in M. Bug and E. van Um, “Herausforderungen bei der Messung von Kriminalitätsfurcht,” DIW Roundup, no. 49 (2015). www.diw.de/documents/publikationen/73/diw_01.c.491381.de/diw_roundup_49_de.pdf, accessed on February 19, 2015. Also see M. Noack, “Probleme bei der Reliabilitäts- und Stabilitätseinschätzung für allgemeine Kriminalitätsfurchtindikatoren,” in “Empirische Forschung über Kriminalität,” eds. S. Eifler and D. Pollich (Wiesbaden: Springer VS, 2015): 249–274. For an international discussion, see K. F. Ferraro and R. L. LaGrange, “The measurement of fear of crime,” *Sociological Inquiry* 57 (1987): 70–101.

⁹ TNS Infratest Sozialforschung, “2014. SOEP 2014 – Erhebungsinstrumente 2014 (Welle 31) des sozio-oekonomischen Panels: Haushaltsfragebogen, Altstichproben,” SOEP Survey Papers 236, series A (Berlin: DIW/SOEP, 2014). Valid answers follow a three-point scale from very concerned, somewhat concerned, to not concerned at all.

¹⁰ Parallel to the usage of the standard indicator, many long term oriented studies are based on the SOEP questionnaire: C. Krekel and M. L. Poprawe, “The Effect of Local Crime on Well-Being: Evidence for Germany,” SOEP Papers 678 (Berlin: DIW/SOEP, 2014). J. Dittmann, “Entwicklung der Kriminalitätseinstellungen in Deutschland – eine Zeitreihenanalyse anhand allgemeiner Bevölkerungsumfragen,” DIW Discussion Papers, no. 468 (2005). http://www.diw.de/sixcms/detail.php?id=diw_02.c.230702.de, accessed on November 23, 2014. J. Dittmann, “Wahrnehmung der Kriminalität im Zeitverlauf,” (Bundeszentrale für politische Bildung, 2013a): <http://www.bpb.de/nachschlagen/datenreport-2013/oeffentliche-sicherheit-und-strafverfolgung/173837/wahrnehmung-der-kriminalitaet-im-zeitverlauf>, accessed on November 24, 2014.

⁴ For a more detailed explanation of the chosen weighting methods, see Bug and Meier, “How to Obtain a More Accurate Picture of Crime through Crime Statistics – Proposals and Methods,” DIW Economic Bulletin, no. 3 (2015).

Box 1

Objective Crime Rates

The offenses under consideration are broken down as follows using *PKS* (police crime statistics) codes:

Theft (PKS code ****00 excluding 440*00), burglary (PKS code 435*00 and 436*00), assault (PKS code 222000 and 224000), verbal threats or similar (PKS code 232300, 673000, 232200, and 232400), cybercrime (PKS code 980100 via the Internet), and murder and manslaughter (PKS code 892500).¹ The findings presented in the present article are based on the frequency of the offenses. These are calculated according to the formula for each region:

$$\frac{\text{Absolute no. of offenses} \times 100,000}{\text{No. of inhabitants}}$$

Since police statistics only include reported crimes, the problem of unreported crimes can only be taken into account by calculating offense-specific correction factors. To estimate the share of crimes that went unreported, a nationwide victimization survey was conducted.²

The table shows calculated correction factors based on offense-specific estimates of figures for unreported crimes (averaged between 2012 and 2013). While the correction of these figures for property crimes and assault is comparatively low, there is a noticeable factor for unreported crimes relating to burglary. What is most striking are the figures for unreported cybercrimes. Only those cases that cause actual damage are considered, e.g., from a virus attack, and not in cases where

the issue was recognized and resolved by anti-virus software. The estimated number of unknown cases of cybercrime is pertinent since some of these damages are borne directly by banks, insurance companies, and other service providers or the low level of damages incurred reduce the willingness of victims to come forward. The generally perceived low clearance rates for these offenses and the lack of knowledge about cybercrime reporting procedures is likely to reduce incentives to report these crimes to the police.

Table

Means of Dark field Estimates for 2012 and 2013

	Mean	Standard Deviation
Murder and Manslaughter ¹	1.8285	
Burglary	5.565	0.039
Theft	2.937	0.128
Bodily Harm	4.047	0.721
Threat	28.911	0.339
Crime via the Internet	247.151	1.848

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

Source: Bundeskriminalamt (2012, 2013): *Polizeiliche Kriminalstatistik*. Bug, M.; Kroh, M.; Meier, K.; Rieckmann, J.; van Um, E.; Wald, N. (2015): *WISIND-data*

© DIW Berlin 2015

One officially known case of bodily harm represents four actual cases of bodily harm (3 out of 4 cases are not reported to the police).

¹ Bundeskriminalamt, "Polizeikriminalstatistik" (2013).

² See Bug and Meier, "Crime Statistics."

Besides these general questions on individuals' feeling of safety and their concerns about crime and the development of crime on the whole, the WISIND survey also examines risk perception with regard to crimes against property, physical violence, and cybercrime (see Box 2).

To cover the different dimensions of the fear of crime as far as possible, the analysis of the region-specific fear of crime presented here is not only based on survey findings but also includes crime-related behavioral patterns which can be seen as an indirect expression of the fear of crime. Thus, the conative dimension of the fear of crime — i.e., protective or avoidance behavior — can be factored into the indicator. This behavior includes private expenditure on security equipment, which is examined in a survey of German security firms conducted

by the Brandenburg Institute for Society and Security (*Brandenburgisches Institut für Gesellschaft und Sicherheit, BIGS*) annually since 2012.¹¹

Finally, user-generated communication with reference to crime in social networks was used for the indicator — a new and innovative approach. Here, over a period of four months, an extensive list of search engine keywords was used to pinpoint relevant posts, which were then assigned to the specified location data; these indicator data were then factored into the analysis of regional fear

¹¹ C. Gummer et al., "BIGS-Studie Sicherheitswirtschaft 2012" (2014). Y. Gruchmann, et al., "BIGS-Studie Sicherheitswirtschaft 2013" (2014).

Box 2

Survey Data on Subjective Fear of Crime (WISIND Survey)

As part of the WISIND project, a representative telephone survey was conducted from July through September 2014 by TNS Emnid in which 12,094 people in Germany were asked about their experiences and perception of crime. One-fifth of respondents were surveyed on cell phones. The nationwide sample was drawn proportional to the regional population and includes at least 15 respondents per city or regional district. The following questions are relevant indicators of subjective fear of crime, or risk perceptions to be more precise:

Questions on the various types of crime:¹

B.1: "If you think back over the last 12 months, how worried were you that someone would steal something from you—whether you were at home or elsewhere—without you being threatened with violence or violence being used against you?"

¹ Possible replies follow the logic of a four, point scale and are the following: "very probable/probable/improbable/very improbable" or "very worried/somewhat worried/not really worried/not worried at all".

B.2: "How likely do you think it is that something will be stolen from you in the next 12 months—whether you are at home or elsewhere—without you being threatened with violence or violence being used against you?"

B.8: "If you think back over the last 12 months, how concerned were you that, for whatever reason, you would be the victim of assault?"

B.9: "How likely do you think it is that you will be the victim of assault in the next 12 months?"

B.17: "How concerned were you in the last 12 months that you would be the victim of cybercrime?"

B.18: "How likely do you think it is that you will be the victim of cybercrime in the next 12 months?"

The analyses below use the respective averages of offense-specific questions about concerns and probability to evaluate risk perception.

of crime.¹² In particular, public content such as Facebook and Twitter were used, as were various online forums.

Weighting Criminal Offenses

To investigate crime rates and the fear of crime in general, the simplest method is to add up the relative frequencies of individual offenses in a given region by offense category. Accordingly, the fear of crime can be calculated as the sum of the relative frequency of expressions of concern among the population across the individual offense categories. One possible criticism of this method is that it does not take into account the relative severity of the offense, meaning a robbery is regarded in the same light as pickpocketing. To evaluate the significance weighting of individual offenses, three alternative weighting methods were used in addition to equal weighting.

The personal perception of the severity of different criminal offenses was investigated using an online survey among 2,532¹³ respondents asked to rank the different offense categories successively by severity. The result was a weighting scale which was originally based on paired comparisons of rank ordered offense categories. The same survey was conducted among 207 experts (predominantly from the field of security research with a small number from the security business) as well.

In addition to these opinion-based methods, the significance weighted values for individual offenses were also estimated using item response theory.¹⁴ The basic idea behind this statistical method is that a factor which is not observed directly (in our case, regional crime rates or the regional fear of crime) is expressed in the indicators observed (in our case, the actual crimes committed or the expression of fear per offense category). Here,

¹² For more details, see Rieckmann and Schanze, "Perceptions of Personal Safety in Social Media and Search Engines – a Realistic Reflection of Actual Crime Rates" (2015) in this issue of DIW Economic Bulletin.

¹³ To avoid systematic deviations between the survey data and the population, the sample also included respondents who do not use the Internet. For more details, see Bug and Meier, "Crime Statistics."

¹⁴ F.M. Lord, M.R. Novick, and A. Birnbaum, *Statistical theories of mental test scores* (Reading, MA: Addison-Wesley, 1968); G. Rasch, *Probabilistic models for some intelligence and attainment tests*, Danish Institute for Educational Research, Copenhagen, expanded edition with foreword and afterword by B.D. Wright (University of Chicago Press, 1960/1980). P.F. Lazarsfeld and N.W. Henry, *Latent Structure Analysis* (Boston: Houghton Mifflin, 1968).

the relative frequency of the relevant offense is taken into account, as are statistical relationships between offenses. If this is found to be particularly strong,—if, for instance, a certain offense is committed particularly often in an area where other crimes can also be found—this offense is deemed to be particularly effective (having good “forecast quality”) as an indicator of general crime and fear of crime. What is known as the relevance parameter (shown in Table 1) denotes this forecast quality. Provided the relevant data are available, the weighting method used for the objective crime rate indicator is the same as that for the subjective fear of crime.¹⁵

Table 1 shows the significance weighted values resulting from the methods used to weight offenses. The results are relatively uniform and intuitive. Physical violence is regarded as being more severe than crimes against property, while verbal threats and suchlike were perceived least serious. An interesting finding was the high weighting assigned to politically motivated crimes, which was the same across all the weighting methods used and would seem to imply that the political system in Germany is highly respected and valued.

The statistical correlations found between the individual offenses—as seen in the item response theory results—suggest that murder or manslaughter is not only a very rare offense (frequency parameter) but also displays a weak association to the frequency of other offenses (low relevance parameter), i. e., in a regional comparison, murder is a rather incidental occurrence. By way of contrast, threatening behavior, theft, and cybercrime are offenses that systematically occur in areas which are particularly strongly affected by crime, which is why such offenses feature more often in determining general crime rates in a given region.

By analogy, Table 2 shows the weighting of criminal offenses used to assess the subjective fear of crime. As explained above, the significance weighted values taken from the objective indicator are averaged according to broader offense categories. For sub-indicators which cannot be allocated to a specific offense, the average is used across all the weighted values. Table 2 also depicts the relevance parameters of the individual offenses.

¹⁵ Owing to the aforementioned, broader crime categories crimes against property, physical violence, and cybercrime, different relative degrees of severity are approximated as the relevant averages (for crimes against property, for instance, the weighting is calculated as follows: $(\text{Weight_Burglary} + \text{Weight_Theft})/2$. Please note that, when calculating the weighting of the broad crime categories, the weighting of other offenses which are included in the online survey, but which were not used for the actual crime indicator is also factored in. For example, the opinion-based weight for physical violence also contains a weight for rape. Due to database problems (in particular areas where police crime statistics show a very high dark figure), this offense is not part of the objective threat indicator.

Table 1

Crime Weights

	Population	Experts	IRT (1/frequency)	IRT (Relevance)
Murder and Manslaughter	0.9055	0.9585	1 (fix)	0.000079
Politically motivated crime	0.1012	0.1067	0.053	0.035
Bodily Harm	0.0476	0.0661	0.001	1 (fix)
Internet Crime	0.0263	0.0339	0.0003	13.224
Burglary	0.0193	0.0224	0.0017	1.406
Theft	0.0114	0.0112	0.0004	4.416
Threat	0.0089	0.0193	0.0002	6.654

Source: Bundeskriminalamt (2013): *Polizeiliche Kriminalstatistik*; Bug, M.; Kroh, M.; Meier, K.; Rieckmann, J.; van Um, E.; Wald, N. (2015): *WISIND-data files: Weighting; Calculations by DIW Berlin*.

© DIW Berlin 2015

Security experts as well as the general population give crime via the internet a weight that is worse than burglary.

Table 2

Weighting of fear regarding different crime categories

	Population	Experts	IRT (Relevance)
Assault	0.247	0.267	
risk_assault			1.46
Manslaughter/Assault			0.922
Property crime	0.015	0.017	
risk_property			1.127
Burglary/Theft			0.896
Assault/Property Crime (standard indicator)	0.181	0.195	
standard indicator			1 (fix)
Cybercrime	0.026	0.034	
risk_Internet Crime			0.6
Unspecific (mean of all crime categories)	0.132	0.146	
SOEP (Worries about development of crime)			0.118
BIGS (Private spending on security)			0.476

Source: Bug, M.; Meier, K.; Kroh, M.; Rieckmann, J.; van Um, E.; Wald, N. (2015): *WISIND-Datafiles Crime weighting/social Networks*. Gruchmann, Y. et al. (2014): *BIGS-Studie Sicherheitswirtschaft Juni 2014*, Gummer, C. et al. (2013): *BIGS-Studie Sicherheitswirtschaft Mai 2013*. SOEPv30; Calculations by DIW Berlin.

© DIW Berlin 2015

The fear regarding Internet crime and property crime show the lowest weights. This follows the feasibility of financial compensation of this crime category.

es, which display similar results to Table and are to be equally interpreted.¹⁶

¹⁶ Table 2 disregards the parameters for relative frequency (IRT model) per offense. When using item response theory (IRT) to analyze the regional fear of crime, the average regional fear of crime per offense is divided into a maximum of three groups with ascending values. The relevant IRT model for ordinal data depicts the relative frequency for each of these (max. three) categories per offense and the inclusion of these parameters would make presenting the summary of figures in Table 2 unnecessarily complicated.

Figure 1

Regional Crime Indicator 2013 (full version)

Source: Bundeskriminalamt (2012, 2013): *Polizeiliche Kriminalstatistik*. Bug, M.; Kroh, M.; Meier, K.; Rieckmann, J.; van Um, E.; Wald, N. (2015): *WISIND-data files: Crime Survey/Crime weighting*. Calculations by DIW Berlin.

© DIW Berlin 2015

The dark-figure adjustment weakens differences between urban and rural areas. Higher crime rates in the north stay obvious.

Regional Differences

Generating a reliable description of regional differences in the fear of crime requires a sufficient number of respondents per regional unit. Although the WISIND sample has at least 15 respondents from each of the 402 administrative districts in Germany, in order to shore up the robustness of the reported findings, the following description divides the country into 60 regions based on the 402 urban and rural districts. The objective of

dividing up these areas is to create regional units with at least one million residents. These units are based on state- and government districts as well as police boundaries (police headquarters).¹⁷ In order to allow a direct comparison between subjective fear of crime and ob-

¹⁷ In some cases, historical/cultural borders were used for regionalizing the federal states. Some cities were removed from their regions to allow them to be considered separately.

jective crime rates, the crime rate is also aggregated in these regions—in principle, data on individual counties could be reported based on criminal statistics compiled by the police.¹⁸ Crime rates and fear of crime are all normalized to the interval $[0-1]$, where 1 represents the highest crime rates and/or fear of crime.

Figure 1 shows the results of the indicator for objective crime rates for 2013 at regional level. All approaches to crime weighting produce comparable results: what is particularly striking is the distinct north-south divide, with higher crime rates in the north. As expected, the major cities (except Munich) stand out, as do the conglomerates of the Rhineland and the Ruhr.

A comparison of the crime rate maps with and without estimated numbers of unknown cases in Figure 1 indicates a stronger regional differentiation due to the adjustment for unreported crime. This adjustment highlights, for example, a higher crime rate in large parts of Baden-Württemberg, Thuringia, Saxony-Anhalt, and Brandenburg.

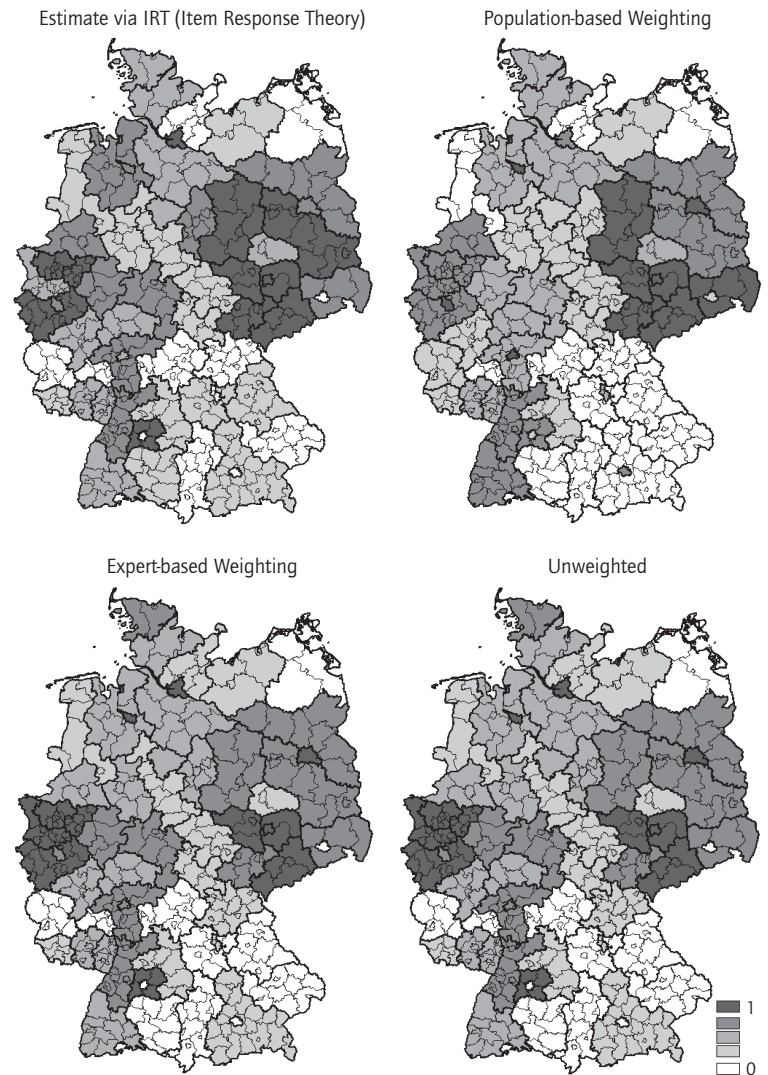
The different methods of crime weighting lead to thoroughly comparable results—as noted in previous studies.¹⁹ In contrast to equal weighting of crimes, opinion-, expert-, and statistics-based (IRT) weighting particularly in the urban areas of North-Rhine Westphalia and in the region Oldenburg indicate higher crime rates—compared to similar regions. The relatively high crime rates in Mecklenburg-Western Pomerania using all forms of weighting can be explained by a disproportionately high volume of cybercrime in these areas in 2013. This effect disappears in the alternative calculation of crime indicators without this form of crime (see Box 2).

Figure 2 shows the results for the indicator of subjective fear of crime compared to the objective crime rates. The frequently made statement that fear of crime is high in regions where the actual threat is low can only be confirmed in individual cases here (see Figure 3). Examples of this are predominantly in parts of Swabia, primarily the counties around Stuttgart. In most regions, however, there is a more or less reflective association between risk perception and measured crime. This correlation is particularly clear in the overall regional distribution shown in Figure 3.

One striking discrepancy between fear of crime and actual crime can be observed in Mecklenburg-Western Po-

Figure 2

Regional Indicator Fear of Crime 2013/14



Source: Bug, M.; Kroh, M.; Meier, K.; Rieckmann, J.; van Um, E.; Wald, N. (2015): WISIND-data files: Crime weighting/Social Networks. Gruchmann, Y. et al. (2014): BIGS-Studie Sicherheitswirtschaft Juni 2014, Gummer, C. et al. (2013): BIGS-Studie Sicherheitswirtschaft Mai 2013. SOEPv30; Calculations by DIW Berlin.

© DIW Berlin 2015

There are slightly higher rates of fear of crime in the north. Between urban and rural areas, no consistent differences are detectable.

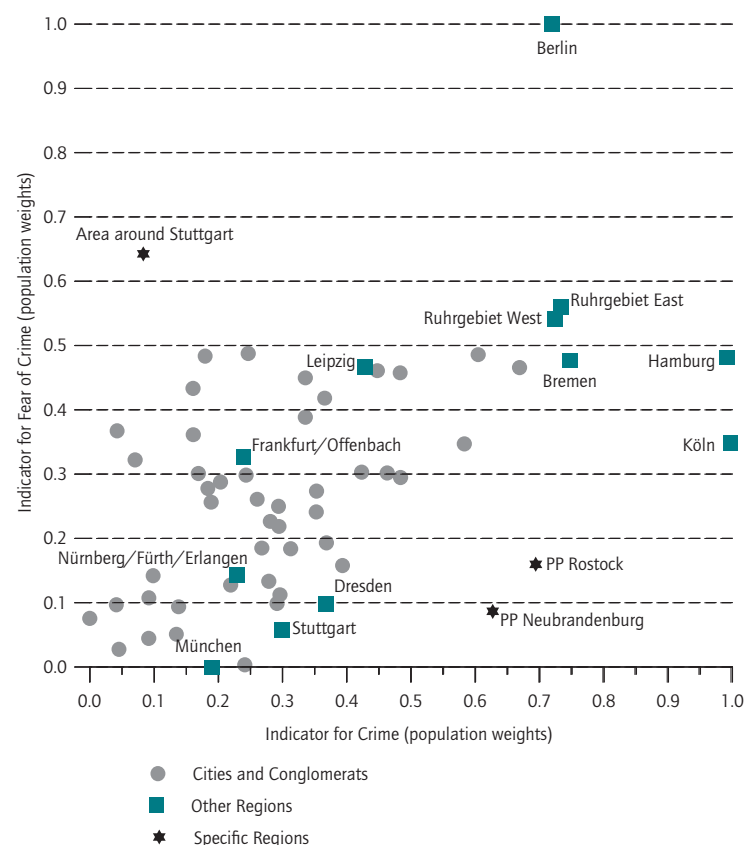
¹⁸ Although WISIND survey data are included in regional crime rate calculations to adjust police crime data for the estimated number of unknown cases, this number is uniform throughout Germany.

¹⁹ Bug and Meier, "Crime Statistics."

merania. A closer look at criminal activities there suggests that the relatively high crime estimates for this region are largely driven by Internet crime. This type of crime is growing very rapidly in two ways. The development of reported Internet crime reveals high growth rates (although the number of official cases is still relatively low compared to more "well-known" offenses).

Figure 3

WISIND-Indicator of Crime and Fear of Crime



1 Bevölkerungsgewichtet.

Source: Bundeskriminalamt (2013): Polizeiliche Kriminalstatistik; Bug, M.; Meier, K.; Kroh, M.; Rieckmann, J.; van Um, E.; Wald, N. (2015): WISIND-Datafiles Crime weighting/Social Networks. Gruchmann, Y. et al. (2014): BIGS-Studie Sicherheitswirtschaft Juni 2014, Gummer, C. et al. (2013): BIGS-Studie Sicherheitswirtschaft Mai 2013. SOEPv30; Calculations by DIW Berlin.

© DIW Berlin 2015

Berliners seem to have the most fear of crime, while Munich's population appears to be relaxed in a context of low crime.

However, by what means damage is being caused (and how this can be registered in PKS) is changing dynamically. Accordingly, victimization studies, particularly in the field of Internet crime, indicate very high numbers of unknown cases. We estimate that only 1 in 247 criminal acts is recorded in official police crime statistics, which as mentioned above, is partly due to the low level of damages paid (see Box 3).²⁰ The analysis of the

²⁰ These analyses provide a regionally constant, unreported crime factor. If, for example, the recording and combatting of cybercrime indicates regional differences, this leads to an under- or overestimation of crime rates in some regions. It is not impossible, for example, that a greater number of actual cybercrimes in Mecklenburg-Western Pomerania indicates a greater willingness of citizens to report such crimes and not necessarily that the crime rate is in fact higher.

Box 3

Taking Account of Cybercrime in Calculating the Crime Rate

The unreported crime factor of 247 represents a significant intervention in the raw data of the police crime statistics – although the official crime figures for 2013 are relatively low despite steady increases in recent years. Moreover, another particular characteristic is that extreme case-number outliers may occur more often due to the nature of cybercrime. This can, for instance, lead to a significant distortion of the frequency numbers due to extensive cases with multiple victims – especially in Länder where each victim finds its way as a case into the police statistics. As an example, crime statistics in the town of Delmenhorst were considerably inflated in 2012 and 2013 due to charges of fraud against a company based there.¹ The extremely wide-ranging adjustment of unreported crime figures then has a critical impact on the estimates for Delmenhorst, when included in a regionalized representation. Another problem is, of course, the location of the offenses. The Delmenhorst example indicates that this town is the scene of the crime, but the defrauded customers are likely to be spread throughout the country and beyond. Consequently, estimates of the local cybercrime rate are subject to greater uncertainty than the “classic” forms of crime. In order to estimate the impact of cybercrime on the findings, a second version of the objective crime rate was calculated without cybercrime for comparison purposes (see Figure Box).

Compared to an analysis based on all forms of crime, two key differences can be identified: a more distinct north-south divide puts the regions of Baden-Württemberg, excluding Stuttgart, in the lowest crime group. Similarly, the estimated crime rate remains high in North Rhine-Westphalia and is now focused in particular on the Ruhr area and Cologne. A second difference is the somewhat greater divide between rural and urban areas which is especially prevalent in Leipzig and Dresden and also clearly visible in Munich and Hanover. This is not surprising since the leveling effect of the city/country is no longer relevant for cybercrime.

¹ Presseportal.de/ots (April 4, 2014), POL-DEL, “Vorstellung der Polizeilichen Kriminalstatistik 2013,” <http://www.presseportal.de/polizeipresse/pm/68438/2705988/pol-del-vorstellung-der-polizeilichen-kriminalstatistik-2013-rckgang-der-registrierten-straftaten>, accessed on January 3, 2015.

Figure

Regional Crime Indicator 2013 (full version without Cybercrime)



Source: Bundeskriminalamt (2012, 2013): Polizeiliche Kriminalstatistik; Bug, M.; Kroh, M.; Meier, K.; Rieckmann, J.; van Um, E.; Wald, N. (2015): WISIND-Datensätze: Kriminalitätsbefragung/ Kriminalitätsgewichtung/ Soziale Netzwerke; Calculations by DIW Berlin.

When not taking Internet crime into account, the dark-figure adjusted crime rates still show higher crime rates in urban areas.

WISIND survey in an article by Rieckmann and Kraus shows what specific forms of cybercrime were committed in the summer/fall of 2014.²¹ This makes it quite clear that the data basis for estimations of unreported crime needs to be improved.

Conclusion

The correlation between fear of crime and actual crime rates can only be satisfactorily examined if reliable measurements for both figures are available. For this purpose, the present report proposes several approaches, all of which produce similar results, thus giving them a certain robustness. One core innovation lies with the integration of key communications data on crime in social networks. The report by Rieckmann/Schanze in this issue of *DIW Economic Bulletin* gives an insight into which issues are of particular interest and where in Germany the exchange of information about crime is particularly vigorous. The article by Bug, Kraus and Walend²² analyzes the findings of the broader WISIND approach to measuring fear of crime compared to the standard question in many public opinion polls asking about feelings of safety when walking in a particular neighborhood at night.

²¹ Rieckmann, J., Kraus, M., „Tatort Internet: Kriminalität verursacht Bürgern Schaden in Milliardenhöhe“ DIW Wochenbericht 12 2015 (2015)

²² Bug, M., Kraus, M., Walenda, B. „Analoge und digitale Unsicherheiten: Eine neue Perspektive auf Kriminalitätsfurcht“ DIW Wochenbericht 12 2015 (2015)

Mathias Bug is Research Associate of the Department for Development and Security at DIW Berlin

Martin Kroh is Deputy Head of the German Socio- Economic Panel Study at DIW Berlin

JEL: K14, R19, H56, H79, H80

Keywords: Security, crime, indicator, fear, fear of crime, home affairs, police, inner security, threat, crime statistics, police statistics

A further review of the different approaches should be conducted once larger datasets with very high numbers of respondents are released for adjusting the figures for unreported crime since the present study is only able to calculate these figures for groups of crimes on a national scale. Greater numbers of cases will allow us to apply the figures on a regional scale.²³

The common hypothesis that people are particularly anxious in areas where there is actually no significant threat was not confirmed by our study. Rather, a more or less realistic assessment of the threat of crime can be observed in most regions in Germany. The article by van Um, Huch and Bug²⁴ outlines how the local media—as an intermediary between local crime and individual perception of crime—treats crime issues in Germany and which categories of crime it targets. This allows us, for the first time, to observe in more detail any minor distortions in crime reporting.

²³ The data set from the BaSiD project with 35,000 respondents from 2012 will allow the figures for unreported crimes to be calculated on a regional scale, Birkel, C., Guzy, N., Hummelsheim, D., Oberwittler, D., Pritsch, J. (2014): Der Deutsche Viktimisierungssurvey 2012. 77–79. http://www.bka.de/DE/Presse/Pressemitteilungen/Presse2014/141208__Viktimisierungssurvey2012.html?__nnn=true accessed on December 8 2014.

In addition, the states of Lower Saxony, Schleswig-Holstein, and Mecklenburg-Western Pomerania have made considerable efforts to research the figures for unreported crimes with similar numbers of respondents as the BaSiD project.

²⁴ Van Um, E., Huch, M., Bug, M. „Lokale Kriminalitätsberichterstattung: Abbild oder Zerrspiegel von Kriminalität“ DIW Wochenbericht 12 2015 (2015)

Kristina Meier is Research Associate of the Department for Development and Security at DIW Berlin



Prof. Dr. Martin Kroh is Deputy Head of the Socio-Economic Panel Study at DIW Berlin

SIX QUESTIONS TO MARTIN KROH

»Regional Crime Rates and Fear of Crime: Distinct North-South Divide—Unified Measurement Methods Needed«

1. Professor Kroh, there are a number of ways of measuring a region's crime rate. What are the differences between the various methods? As a general rule, crime rates are measured based on police crime statistics which include reported crimes differentiated by date and region. This is actually a very good basis. However, if the aim is to determine a region's crime rate, police crime statistics also have certain weaknesses. First, there is the issue of "light and dark" crime figures whereby some offenses go unreported. The second issue is how to weight the different crimes in relation to one another. Murder has quite a different significance to pickpocketing, for instance.
2. How do the findings differ? There are various ways of weighting criminal offenses relative to one another. One method might be, for example, on the basis of the amount of damages insurers pay out to quantify specific offenses. Here, damages for murder would of course be many times higher than those resulting from a pickpocketing offense. If we apply this weighting method we can see, for instance, that a rural region where a murder happened to be committed would be evaluated as having a very high crime rate. Other methods would not assign as much weight to this murder but would take greater account of the relationship between different criminal offenses.
3. What is the correlation between the objective crime rate and subjective fear of crime? In this project, we have tried to use similar methods to measure both the subjective fear of crime and objective crime rates. The findings demonstrate that fear and the objective crime rate are fairly clearly in line with one another. Thus, people have a relatively accurate perception of the level of crime in their local area.
4. What regional differences can be seen from the project's findings? In northern Germany, the fear of crime is considerably higher than in the south of the country which also corresponds with the actual crime rates and, as expected, the level of fear is somewhat higher in urban than in rural areas. However, there are also regions where the fear of crime is higher than the actual crime rate or vice versa. For example, Cologne is a city where the fear of crime is comparatively low while the real threat is relatively high. In the Stuttgart area, on the other hand, at least according to our findings, the fear of crime is greater than the actual threat.
5. Did your project also consider more abstract crimes such as white-collar or Internet crime? Our study focuses in particular on the threat to individuals and so does not take white-collar crimes into consideration. However, the analysis does include cybercrime at least where the victim is a member of the public rather than a company or bank. When measuring crime, Internet crime is an important issue since the number of unreported offenses here is significant. Very few cybercrimes are reported, perhaps because the level of damages is so low. The urban/rural divide is also considerably less pronounced here since victims and perpetrators do not necessarily have to be located in the same region.
1. Which measurement method do you feel is the most informative? As a methodologist, I think that the statistical method we describe in our report is the most robust. However, in my view, ultimately the most important thing is to agree on one method. Currently, any statements made on the development of crime in Germany are based on police criminal statistics. However, it would be more useful to agree on a method that also takes unreported offenses and the severity of a crime into consideration so that differences or changes in crime rates can be better documented.

Interview by Erich Wittenberg

Perceptions of Personal Security in Social Media and Search Engines— a Realistic Reflection of Actual Crime Rates?

By Johannes Rieckmann and Jan-Lucas Schanze

The most common method of measuring subjective fear of crime in the general population has traditionally been through surveys. With the spread of digital technologies, however, data from social media and search engines could now help researchers learn more about people's subjective perceptions of certain types of crime. The present article will show that although the analysis of data from social media and search engines is not suitable as an indicator of actual crime levels in Germany, it can certainly be a cost-effective supplement to traditional methods of collecting data on perceived crime levels.

The number of Internet users in Germany—around 80 percent of the entire population now use the Internet at least occasionally¹—is steadily increasing. This makes the Internet an ever more attractive source of data for crime researchers. Traditionally, the main source of data about the impact of crime on the public has been the police crime statistics (*Polizeiliche Kriminalstatistik, PKS*) of the Federal Criminal Police Office (*Bundeskriminalamt*), supported by occasional and until now largely regional surveys on the “dark figure” of crime (unreported crime) and fear of crime. Online data analysis, on the other hand, could provide deeper insight into the perception of crime; in the long term, it could even become an alternative and independent source of data for crime researchers. The present article² will evaluate whether data collected on the Internet are suitable for supplementing—or even functioning as a cost-effective substitute for—traditional studies on the fear of crime. It will examine the search engine behavior of Internet users in Germany as well as the statements they make on various social media platforms. It is expressly not aimed at forecasting crime trends for the purposes of preventing and fighting crime (predictive policing). Rather, the user-oriented media analysis presented here is intended as an initial descriptive characterization of German perceptions of personal security. The findings

¹ B. Van Eimeren and B. Frees, “Ergebnisse der ARD/ZDF-Onlinestudie 2014. 79 Prozent der Deutschen online – Zuwachs bei mobiler Internetnutzung und Bewegtbild,” *Media Perspektiven*, no. 7–8 (2014). A 12,000-person survey conducted by the WISIND project concluded that the figure was approximately 75 percent.

² The report was compiled as part of the research project An Economic Security Indicator for Germany (WISIND), which is backed by the German Ministry for Education and Research as part of the Social Dimensions of Security Research funding program. The idea behind the WISIND project and the generation of WISIND-specific data was jointly developed by Martin Kroh, Mathias Bug, Kristina Meier, Johannes Rieckmann, Eric van Um, and Nina Wald, together with the staff of the Brandenburg Institute for Society and Security (Brandenburgisches Institut für Gesellschaft und Sicherheit, BIGS). The authors would also like to thank Enrique Fernandez, Martina Kraus and Bartosz Walenda for their assistance throughout this process.

will then be compared with fear of crime as measured by the WISIND project.³

Social Media and Search Engines as Data Sources

The present article uses two types of data. First it draws on data on the number of social media posts about certain types of crime, specifically posts made on Facebook, Twitter, discussion forums, and blogs, as well as comments made on YouTube videos. The purpose of these posts is to express perceptions and opinions, and to communicate; their primary function is therefore expressive.

The second source of data used in the article comes from Germany's leading search engine, Google. Search engines primarily have an exploratory function, as their purpose is to help users acquire information. Google has created the Google Trends platform for exporting data on the occurrence of search terms for given time periods and locations. Google Trends has been used in several scientific studies to observe data generated in the lead-up to various phenomena⁴ and it is used here as a source of search engine data.

Motivation of Internet Users

While search engines are primarily used to acquire information, social media is used mainly to exchange information and views. The use of both, search engines and social media, implies that the user is personally affected, either directly or indirectly. It can be assumed that a large part of the population uses the Internet to gather information on what they perceive to be threatening events or circumstances; that some of them communicate this information on social media platforms; that they are also interested in taking precautions (for example, using alarm systems or pepper spray); and that, in the process, data is generated that creates a picture of public threat perception which is independent of surveys. However, any hypotheses regarding the specific motivations of users are speculative because while tracking data on the Internet shows what users are doing, it does not show why they are doing it.⁵

Collecting Data from Social Media Platforms

DIW Berlin subcontracted data collection from social media platforms to Beck et al. Services GmbH, a company specializing in capturing data from such sources. The data were collected over a period of four and a half months, from June 12 to October 31, 2014. The data used for the analysis comprise all posts and profile details that were shared with the public by users on the media platforms mentioned above; all content was therefore publicly accessible. German-language posts on the five platforms were automatically searched every day using a Web crawler and a list of search terms for ten different crime categories.⁶ The selection of terms, the inclusion of alternative spellings and the use of so-called killer terms (terms that cause content to be excluded) ensured that the results were highly accurate and the data free of irrelevant content. For the crime category "bodily harm," for example, all newly entered text fragments with the keywords *Schlägerei* (brawl), *verprügelt* (beaten up), and *Körperverletzung* (assault or physical injury) were counted. Content with the keyword *Gewalt* (violence) that was connected to places outside of Germany (Ukraine or Iraq, for example) or to abstract concepts, like *höhere Gewalt* (force majeure), were identified as irrelevant and excluded.

Only publicly available content was examined for the presence of search terms. From the start, it was technically impossible to record any messages or content on profile pages that could not be viewed by the general public. In order to be able to map the data by region, any place names mentioned in the text fragments were recorded, and searches were made for mentions of place names by users. For example, the mention of Berlin in someone's profile ("lives in Berlin") can help determine the location of an Internet post on bodily harm as well as the occurrence of a place name in the post itself ("bodily harm in Berlin").⁷ The results of an automated identification of the tone of posts (positive, negative, or neutral) proved to be of only limited use for empirical analysis and are therefore not addressed in this article. No further data were collected on the Internet users, such as age, gender, or other characteristics. User-generated data were anonymized, fully safeguarding user privacy;

³ See also the article by M. Bug, M. Kraus, and B. Walenda in this issue of DIW Economic Bulletin.

⁴ One prominent example is Google Flu Trends (GFT) which was created to predict spikes in flu activity in the US. When people who may be infected with the flu use a search engine to look for (or "google") flu symptoms, the frequency of these queries within a particular region or time is detected by Google Trends and used to infer imminent surges in infection rates, even before patients go to the doctor. For a critique of GFT, see D. Lazer et al., "The Parable of Google Flu: Traps in Big Data Analysis," *Science* 343 (2014): 1203-1205.

⁵ M. Mahrt and M. Scharnow, "The Value of Big Data in Digital Media Research," *Journal of Broadcasting & Electronic Media* 57, no. 1 (2013): 20-33.

⁶ Categories were created for the following types of crime: internet crime, theft and burglary, drug-related crime, bodily harm, organized crime, politically motivated crime, robbery, religious fundamentalism, sex crimes, and homicide.

⁷ Place names in a post tend to be a more reliable indicator of where the incident under discussion occurred.

Box

Problems of Data Collection on the Internet

The analysis of data from social media, often referred to as “big data,” has been the focus of increasing criticism, with charges that the approach taken by some research projects is purely data-driven and not grounded in theory.¹ The mere availability and size of new data sources is not a sufficient argument for the necessity of research. On the contrary, research must always identify problems and qualitative issues and, if possible, establish connections to “traditional” data collection.

The first and most obvious drawback to working with data generated on the Internet is that the data are not representative. Although the group of Internet users in Germany has now become very large numerically, it presents an image that is systematically distorted in favor of younger generations. In the “over-60s” generation, fewer than half of all Germans use the Internet. This deviation in the population is even greater when it comes to social media usage. Three-quarters of respondents under the age of 30 use social media, but only five percent of Internet users over 70 do so.²

In the case of Google Trends, it is not so much the pool of data itself as the way it is presented which is problematic. Google is a commercial enterprise and not a professional supplier of data for research or scientific purposes, and this is reflected in the lack of transparency in the internal processes used for data generation in Google Trends. The way the results are presented is

unfavorable for scientific analyses: they are shown not in terms of absolute numbers of searches but as relative search volumes, always in relation to a maximum value which is set at 100. According to Google, this normalization of the search volume is a result of expressing the search term as a fraction of all searches in a region, making it possible to compare results with other regions. If the total number of searches for a certain term is less than a threshold defined by Google but not publicly stated, the results are either not shown at all (search volume = 0), or they are shown only for longer time intervals (on a monthly instead of a weekly basis). This complex procedure is not made public by Google. Another major problem is that non-verifiable changes may be made to Google algorithms over time, making it more difficult to replicate data.³ Changes in search volumes, particularly when they are observed over long periods of time,⁴ are not always reliable evidence of an actual change in search behavior.

An additional distortion results when certain topics experience a temporary increase in media attention, which artificially drives up searches—a problem that “Google Flu Trends” also struggles with.⁵ On the other hand, attempts to influence the search behavior of Internet users, for example because of the commercial interests of companies that want more clicks on their own pages and hope to sell products, are relatively unlikely in the field of internal security and crime—at least for the search terms used in this study.

1 M. Welker and A. Klob, “Soziale Medien als Gegenstand und Instrument sozialwissenschaftlicher Forschung,” in *Soziale Medien – Gegenstand und Instrument der Forschung*, ed. C. König, M. Stahl, and E. Wiegand (Wiesbaden: Springer VS, 2014).

2 Van Eimeren and Frees, *Ergebnisse der ARD/ZDF-Onlinestudie*: 380, 387.

3 D. Lazer et al., “The Parable of Google Flu,” 1205.

4 Google Trends provides data going back to 2004.

5 D. Lazer et al., “The Parable of Google Flu,” 1204.

the sample does not contain any personal data or data that could be used to identify persons.⁸

Collecting Search Engine Data (Google Trends)

The analysis of search engine usage was restricted to the market leader, Google. The Google search engine was used for 95 percent of all searches in Germany in

December 2014.⁹ The use of search engines among the 55.6 million Internet users in Germany is widespread at 82 percent, according to the 2014 ARD/ZDF online study.¹⁰ Analyzing data using Google Trends is therefore tantamount to collecting almost all search engine queries in Germany. The queries are aggregated by Google at different geographic levels and are not localized by the researchers themselves.¹¹

8 On data protection in the analysis of social media data, see R. Tscherwinka, “Soziale Medien – Gegenstand und Instrument der Forschung – Rechtliche Aspekte,” in *Soziale Medien – Gegenstand und Instrument der Forschung*, ed. C. König, M. Stahl, and E. Wiegand (Wiesbaden: Springer VS, 2014).

9 Internet source: <http://de.statista.com/statistik/daten/studie/167841/umfrage/marktanteile-ausgewaehlter-suchmaschinen-in-deutschland/>

10 Van Eimeren and Frees, “Ergebnisse der ARD/ZDF-Onlinestudie,” 387.

11 There is no detailed information on the exact localization mechanisms used by Google. Presumably this is based on users’ IP addresses (i.e., the addresses of the users making queries).

Collecting data using Google Trends is a much simpler process than capturing data from social media. The user interface has a number of features that make it possible to compare results for different search terms, countries or time periods. Users can obtain regional information about Germany at the state level and filter results by choosing from among categories. The data sample used for this article was filtered by the category “Law and Government” to exclude potentially irrelevant results¹² stemming from the use of ambiguous search terms (see box).

Findings

A comparison of all posts recorded on all social media platforms during the collection period shows that some issues occupied users’ attention significantly more than others (see Table 1). There were almost 300,000 posts on theft and burglary. These were followed closely by posts on religious fundamentalism; posts on homicide, bodily harm and drug-related crime trailed far behind. A comparison of the crime categories that users were interested in with PKS and the findings of the WISIND study on unreported crime supports the supposition that social networks cannot, by themselves, be used as a reliable means of inferring real regional incidence levels. The number of times one crime category was mentioned relative to other crime categories does not reflect the actual incidence level of this crime in Germany. It does, however provide an indication of general sentiment and perceived risk. For example, media coverage of the Hogesa demonstrations in Cologne at the end of October 2014, and of the rioting that accompanied them, led to a sharp rise across Germany in the occurrence of terms that were counted as keywords for the risk categories “religious fundamentalism” and “politically motivated crime.” (Hogesa is an abbreviation of the German phrase meaning “Hooligans against Salafis.”)

Of the 1.2 million posts collected for use in the present study, so far it has been possible to localize at least 18 percent in over 7,300 localities down to the municipal level and assign them to administrative districts in Germany. The results were adjusted for the numbers of inhabitants in the respective areas and converted to posts per 100,000 people.¹³

Table 1

Post frequencies by offences in social media¹

Offence	All posts	Among those: Localized	Proportion of localized offences
Theft and burglary	293,038	57,275	19.5
Religious fundamentalism	288,643	42,672	14.8
Bodily harm	180,133	37,011	20.5
Homicides	147,627	21,604	14.6
Drug related offences	109,260	16,155	14.8
Robbery	107,366	31,253	29.1
Internet crime	53,280	2,639	5.0
Sexual offences	43,217	3,770	8.7
Politically motivated crime	40,384	8,947	22.2
Organized crime	20,478	3,362	16.4
Sum	1,283,426	224,688	17.5

¹ June 12 through Oct. 31 2014.

Source: Bug, M.; Kroh, M.; Meier, K.; Rieckmann, J.; van Um, E.; Wald, N. (2015): WISIND-datasets: social media/Google Trends. Calculations by DIW Berlin.

© DIW Berlin 2015

The number one topics getting the most attention by social media users are theft and burglary as well as religious fundamentalism.

The degree to which content could be localized varied according to crime category. Users writing about robbery, political crimes, and burglary and theft mentioned place names more frequently in their posts than users writing about internet crime or sex crimes. Intuitively this can be explained by the fact that the physical location¹⁴ is largely irrelevant for internet crime, and that specifying locations thus has little value in discussions about this type of crime. In the case of sex crimes, factors such as feelings of shame, frequent lack of witnesses, a sense of respect for the victim, and a desire to protect them presumably account for posts being less frequently associated with physical locations. These examples indicate a systemic imbalance in the social media data that poses a problem for any comparison of localized posts from social media platforms and actual regional crime statistics: some crimes can be assigned to a region more easily than others.

When analyzing social media, it is important to select multiple networks and sources, as selecting a single source further restricts the variance of the data.¹⁵ Most of the posts in the data sample are from Facebook, the most widely used social network in Germa-

¹² Irrelevant results might include references to literature and film (for example, “Murder on the Orient Express”) when searching for murder, or results about illnesses such as Ebola when searching for virus (as in computer virus).

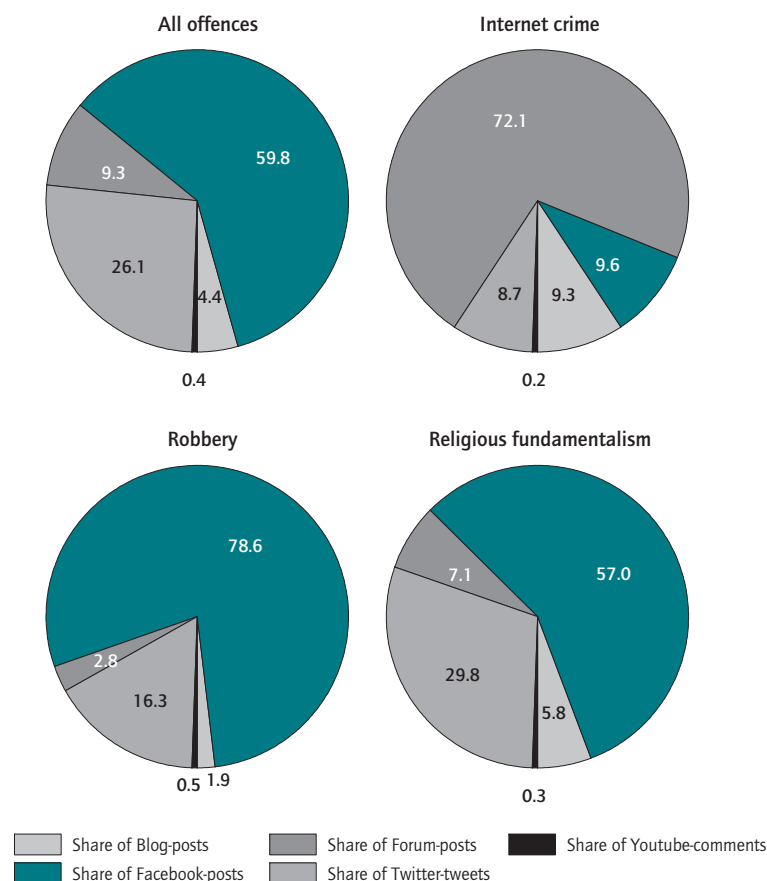
¹³ Otherwise a distorted picture of densely populated urban areas would have been created. For example, more than 45,000 posts were assigned just to Berlin.

¹⁴ The place where the crime is committed is usually far from where the resulting damage occurs.

¹⁵ Mahrt and Scharnow, “The Value of Big Data in Digital Media Research,” 25.

Figure 1

Distribution of social media sources for different types of offences in percent



Source: Bug, M.; Kroh, M.; Meier, K.; Rieckmann, J.; van Um, E.; Wald, N. (2015): WISIND-datasets: social media/Google Trends. Calculations by DIW Berlin.

© DIW Berlin 2015

The relative frequency of posts regarding different offences varies over sources:
 For instance, robbery is written about mostly on Facebook.

ny.¹⁶ This is followed by tweets and posts on discussion forums and blogs (second to fourth place) with YouTube comments accounting for only a very small part of the sample. It is worth noting that different media platforms were associated with different categories of crime. The vast majority of posts on internet crime come from discussion forums in which users can get advice on how to protect themselves from internet crime or recover from any losses they have already incurred (see Figure 1). Posts on robbery are discussed and shared on Facebook with above-average frequency. Fundamentalism—considered to be one of

¹⁶ See also K. Busemann and C. Gscheidle, "Dabei sein ist alles – zur Nutzung privater Communities. Ergebnisse der ZDF-Studie Community 2011," Media Perspektiven, no. 7–8 (2012): 380–390.

the most important issues in the period for which the sample was collected—is mentioned with equal frequency in blogs and discussion forums but more frequently on Twitter and Facebook.

Comparing Social Media with Google Trends

How big is the difference between expressive and exploratory information behavior on the Internet? A Google Trends ranking of the frequency of certain search terms and a comparison with word frequencies analyzed in the sample of social media data for the period from June to October 2014 can shed some light on this (see Table 2). For each data source, the most frequent term was assigned a value of 100 and the other terms were assigned values in proportion to this maximum value. In Google Trends, queries containing the term "murder" clearly predominated; other terms ranked far lower. The rest of the ranking is similar: "theft," "burglary," and "bodily harm" rank near the very top, both in Google Trends data and on social media. Evidently Germans rarely searched for terms that were associated with religious fundamentalism (for example, "Islamist" or "Salafi") in this period—the most obvious difference between Google search and social media.

Table 2

Comparison of ordinal rankings of frequencies in social media and search engines¹

Scale from 1 to 100

Offence	Google Trends	Social Media
Homicides (murder)	100	50
Theft and burglary	26	100
Bodily harm	26	61
Drug related crimes (drugs)	18	37
Robbery	9	37
Internet crime (virus)	9	18
Sexual offences (rape)	9	15
Religious fundamentalism (Islamist/Salafi)	5	99
Organized crime (human trafficking)	2	7
Politically motivated crime (Right wing & Left wing extremism)	2	14

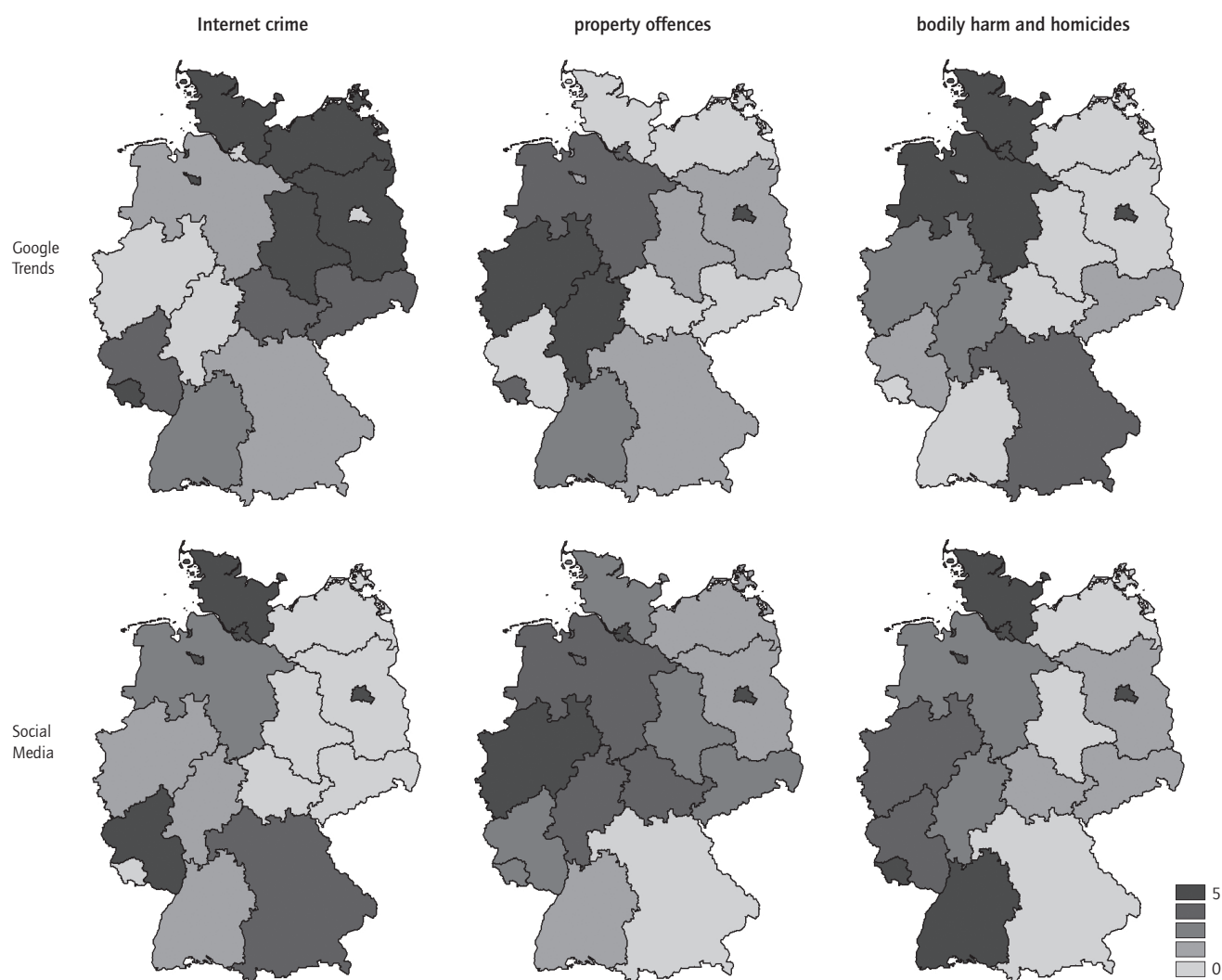
¹ June through October 2014

Source: Bug, M.; Kroh, M.; Meier, K.; Rieckmann, J.; van Um, E.; Wald, N. (2015): WISIND-datasets: social media/Google Trends. Calculations by DIW Berlin.

© DIW Berlin 2015

Terms related to theft, burglary and bodily harm are ranked quite high in Google Trends as well as in social media.

Figure 2

Spatial distribution of Google Trends search queries (above) and posts in social media (below) in quintiles

Source: Bug, M.; Kroh, M.; Meier, K.; Rieckmann, J.; van Um, E.; Wald, N. (2015): WISIND-datasets: social media/Google Trends. Calculations by DIW Berlin.

© DIW Berlin 2015

Terms related to Internet crime yield opposing results in Google Trends and social media: in East Germany Internet crime terms display higher frequencies in Google, in social media they are mostly mentioned in West Germany.

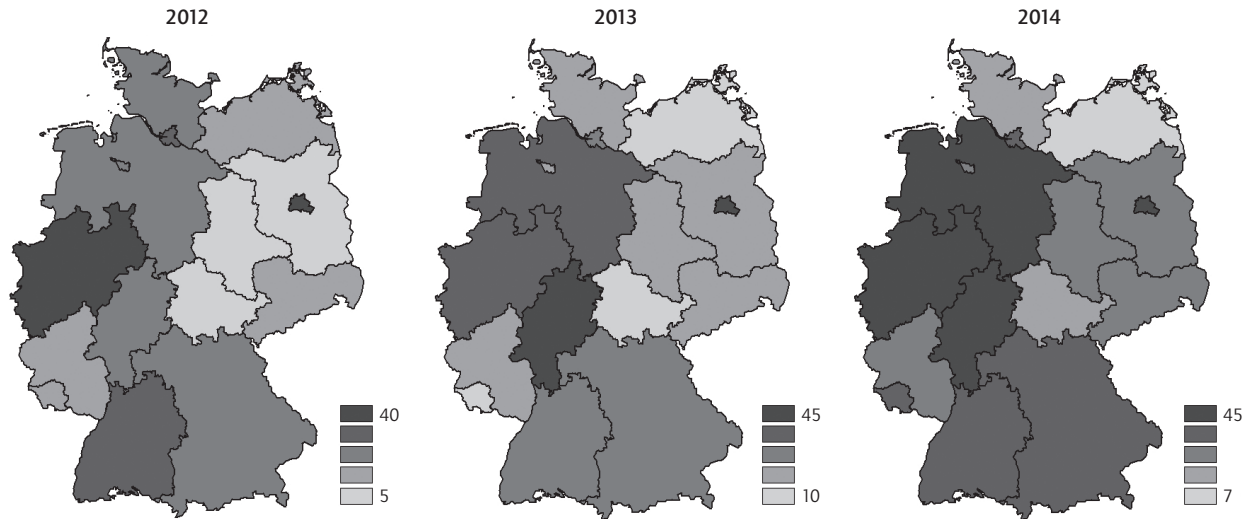
For a further systematic comparison of the two data sources, three classes of crime were formed: the first category contains various types of internet crime, the second consists of terms related to property crime, and the third comprises crimes involving death and physical injury. While Google Trends does provide data for regions below the level of federal states, these data are not complete and not available for every administrative district. For comparison purposes, therefore, the localized data from social networks were aggregated to represent federal states, and Google Trends data were obtained for the period from June to Octo-

ber 2014 (Figure 2 shows a comparison of geographic distributions).¹⁷

The contrast between social media and Google Trends regarding the class of internet crime is striking: Google data contains more searches from northern and eastern German states; the search volume for southern and

¹⁷ A direct comparison of the colors used in the maps (Google Trends and social media) is not possible because the colors represent percentages of different maximum values. Meaningful inferences about the relative frequency of terms can therefore only be made for each of the maps separately.

Figure 3

Google Trends search queries regarding property offences¹

¹ 2012 – 2014, equal intervals

Source: Bug, M.; Kroh, M.; Meier, K.; Rieckmann, J.; van Um, E.; Wald, N. (2015): WISIND-datasets: social media/Google Trends. Calculations by DIW Berlin.

© DIW Berlin 2015

The number of search queries concerning property crimes increased all over the country between 2012 and 2014 (from left to right), in most federal states.

western states is smaller in comparison. These results are similar to the fear of crime measurements made by the WISIND project, which found that northern Germany tended to have higher fear rates.¹⁸ The opposite is true of expressive Internet use, i.e., the use of social media. While the data from social media must be treated with caution—for instance, it was possible to localize only five percent of the total posts on internet crime—this contrast presents an interesting topic for further research.

In the class of property crimes, the German states of Berlin, Hesse, and North Rhine-Westphalia along with Lower Saxony are at the top of the Google Trends ranking with high search volumes for terms like “burglary,” “theft,” and “alarm system.” In social media there is a slightly more distinct north-south divide, with at the same time less obvious differences between West and East Germany.

Crimes involving death and physical violence show both similarities and differences between data sources and regions. In the north-western German federal states (with the exception of Bremen) as well as in Berlin posts about this class of crimes are frequent. Search queries

concerning the respective crime field are less frequent in Baden-Wuerttemberg, Bremen and Saarland than in Bavaria, the exact opposite holds true for social media.

Comparing Data with Actual Crime Rates and Fear of Crime

Google Trends makes it possible to select a time period and compare data spanning several years. Annual averages were calculated on the basis of weekly or monthly data for each of the German states.¹⁹ Between 2012 and 2014, searches related to property crime tended to rise in most states (see Figure 3). Every year, Google users—with the exception of those in Mecklenburg-Western Pomerania and Schleswig-Holstein—searched for terms related to the topic of property crime more frequently than in the previous year, although the increase has tailed off across Germany since 2014.

¹⁸ See also the article by M. Bug, M. Kroh, and K. Meier in this issue of DIW Economic Bulletin.

¹⁹ Seasonal variation in the occurrence of the search term “burglary” would be another interesting topic of study, but one which cannot be examined more closely here. Limiting the data sample from social media to the period between June and October 2014 does not allow any reliable conclusions to be drawn about seasonal changes in the attitudes and interests of Internet users.

The very limited suitability of search engine data as an indicator of objective crime rates—not fear of crime—becomes evident when they are compared to PKS data²⁰. PKS reports for the years 2012 and 2013 show only a slight rise of about 0.1 percent in the number of burglaries and thefts in Germany.²¹ Based on these figures, it would seem that Google Trends cannot be used to make any direct inferences about actual crime rates in Germany. At the federal state level, however, Google Trends correctly reflected developments in crime rates 11 times in the period from 2012 to 2013. It is highly doubtful, however, that these figures can be used to form a reliable overall picture, much less to make predictions.

A corresponding comparison of trends in the data from social media and the PKS cannot be made because of the data collection period, but here too differences can be seen, specifically differences between the number of posts and the number of cases associated with particular locations.

To assess whether the findings presented here from social media and search engines are suitable for use in mapping fear of crime (as opposed to objective crime rates), a comparison with representative survey data is required. A survey was carried out by DIW Berlin's WISIND project; it polled 12,000 people in Germany regarding their concerns about becoming the victim of various crimes.

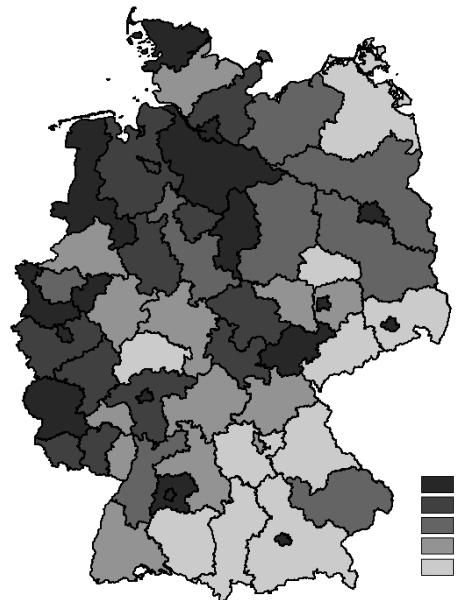
A comparison with the data measuring fear of crime²² in 2014 (presented in another article in this issue of DIW Economic Bulletin) shows that some of the geographic patterns associated with the ten crime categories are reflected in the regional distribution of the crime indicator “subjective fear” (see Table 4).²³ This is particularly true of Schleswig-Holstein, Berlin, the northern part of Lower Saxony, and large parts of North Rhine-Westphalia. The districts around Stuttgart²⁴ and the administrative regions of Karlsruhe and Freiburg can be seen in both maps, both of which show the same levels of Internet activity and fear of crime.

At the same time, the comparison reveals several differences at the regional level. At first glance, these might be attributable to the differing data collection periods; a fur-

Figure 4

Post frequencies in social media for all ten offence groups¹

per 100,000 inhabitants in quintiles on regional level



1 June through October 2014

Source: Bug, M.; Kroh, M.; Meier, K.; Rieckmann, J.; van Um, E.; Wald, N. (2015): WISIND-datasets: social media/Google Trends. Calculations by DIW Berlin.

© DIW Berlin 2015

The spatial distribution of posts within all ten offence groups displays similarities with the mapping of the WISIND crime fear indicator on regional level.

ther distinction can be seen between the simple mention of issues related to crime, on the one hand, and specific questioning about concerns regarding various forms of crime on the other.²⁵ In some cases, this results in regions being depicted on the maps in contrasting colors. The region of Trier shows a relatively low level of subjective fear of crime, i.e., people there tend to worry less that they will become victims of crime. On social media, however, this region is among those with the highest numbers of posts relating to crime. The same contradiction can be observed for Middle Franconia, Upper Palatinate and Lower Bavaria. Overall, the data collected from the Internet clearly reflect fear of crime much more accurately than actual crime rates, and they could be a good indication of differences in attitudes between regional populations.

²⁰ This statement assumes there is no significant change in the number of unreported crimes.

²¹ Federal Criminal Police Office, Jahrbuch Polizeiliche Kriminalstatistik (2012–2014).

²² For a graphic showing crime indicators of subjective fear, see also the above-mentioned article by Bug, Kroh, and Meier.

²³ The graphic is based on a rendering of regional social media activity in quintiles.

²⁴ It consists of the Rems-Murr, Böblingen, Esslingen, and Ludwigsburg districts.

²⁵ Media-use behavior, influenced by demographic factors, can also play a key role in regional differences.

The ability to regionalize the social media data sample widens its potential application beyond simple comparisons with other figures. The prominent position in the data of the Braunschweig-Salzgitter-Wolfsburg region is striking: it has the maximum value by a large margin for frequency of terms related to religious fundamentalism. This was the case even months before media coverage of the arrests of young Islamic extremists in Wolfsburg and other coverage of the carnival parade in Braunschweig that was recently canceled due to direct threats of a terrorist attack by Islamic extremists.²⁶

Conclusion

The findings of the analysis of data from social media and search engines indicate that, because of their intrinsic qualities and also for systemic reasons, they are not suitable for creating a reliable picture of real regional crime rates. However, the data certainly can be used to form a picture of subjective perceptions of

crime in regional populations and as such can function as a cost-effective data source supplementing traditional surveys on fear of crime. The analysis of social media content shows three limitations to interpreting the data as an indicator of actual crime risk. First, social media are heavily influenced by media effects and episodes of heightened interest in certain topics; this typically takes the form of retweeting, reposting, and sharing excerpts from other media. Second, it is not easy to localize the data. Data can be localized only when users disclose this information.²⁷ Their willingness to do this, however, depends on the context of the posts, with the result that some terms related to specific crimes can be localized more easily than other terms. Third, the selection of Internet platforms is very important, as there are considerable variations in the types of media used to discuss different categories of crimes. Comparisons with actual crime rates are further limited by a sample bias in favor of younger users, who are much more active on social media.

26 The map depicting the frequency of terms related to the "religious fundamentalism" crime category is not included in this DIW Economic Bulletin article.

Johannes Rieckmann is Research Associate within the Department of Development and Security at DIW Berlin

JEL: H79, K42, R12, Z18

Keywords: Crime, crime statistics, indicator, perception, security, social media, social network, search engine

27 A further challenge is the use of place name spellings that deviate from official orthography, as the matching process cannot be fully automated and is time-consuming.

Jan-Lucas Schanze is Student Assistant within the Department of Development and Security at DIW Berlin