The Darkfield of Cybercrime: Can Survey Data Reduce Administrative Data's Problem with Validity?

Julia Kleinewiese[1*]
University of Mannheim

Abstract

Cybercrime has become a major issue in digitalized societies. Addressing the rising amount of cybercrime necessitates high-quality research, beginning with examining its prevalence and trends. Examining the prevalence and trends of cybercrime requires a methodological approach that tackles the typical data quality issues of (cyber)crime data, such as validity. A primary problem is that different types of data (e.g., administrative process-generated data and survey data) do not show the actual number of crimes committed, leading to a large darkfield (dark figures). In order to tackle the methodological issue of the darkfield and concomitant validity problems, this article builds on prior research on administrative data and survey data as well as on a general background regarding underreporting issues in crime research. For instance, discussing the role of social desirability and trust in surveys. It then draws on the previous methodological research on crime data, generally, and on cybercrime data, specifically, to suggest an integrated "mixed-data" approach in which different data types (such as administrative data and survey data) are analyzed comparatively in order to gain more information on crime prevalence and trends. Embedded in the previous research field, it proposes this procedure in form of the "Data Combination Approach" (DCA). This approach is described and discussed, including potentials and methods of analyses as well as challenges, particularly regarding the differences between data types. In doing so, this article provides a solid foundation for future high-quality research on crime (particularly cybercrime) prevalence and trends.

Keywords: darkfield, dark figure, cybercrime, crime data, validity, process-generated data, administrative data, survey data, data combination

[1] Mannheim Centre for European Social Research, University of Mannheim
Email: kleinewiese@uni-mannheim.de
Julia Kleinewiese ID https://orcid.org/0000-0003-0053-219X
I have no known conflict of interest to disclose.
Correspondence concerning this article should be addressed to Julia Kleinewiese, Mannheim Centre for European Social Research, University of Mannheim, A5, 6, Building A, 68159 Mannheim, Germany.

The Darkfield of Cybercrime: Can Survey Data Reduce Administrative Data's Problem with Validity?

In the age of digitalization, all areas of social life increasingly involve technology. Every day, people are using their laptops, smartphones, tablets, smartwatches and even the so-called "internet of things". Digitalization makes crime using the internet (cybercrime) an increasingly attractive endeavor for criminals. Besides having much of their social life digitalized, it is now common for people to conduct activities that involve highly personal and financial information via the internet, such as online banking. People often receive phishing emails or their devices are infected with malware which steal or ransom information. Organizations, authorities and universities also rely heavily on technology and the internet, making them potential victims of phishing, distributed denial of service attacks and other cybercrimes. All of this shows why cybercrime has become a major issue in digitalized societies and needs to be investigated scientifically (Di Nicola, 2022).

While there is some research on cybercrime (Buil-Gil et al., 2021; Ghazi-Tehrani & Pontell, 2021; Ngo & Paternoster, 2011; Saridakis et al., 2016), very little of it (Lusthaus, Bruce, & Phair, 2020; Porcedda & Wall, 2021) is primarily methodological. Moreover, there is no methodological research that focusses on measuring the occurrence (also termed "prevalence" or "incidence") of cybercrime. It follows that there is a large research gap regarding the measurement of cybercrime. As a major first and foundational step, this article will build on previous research to suggest a new approach to measuring the occurrence of cybercrime, in order to better assess how much crime is actually taking place. It is a first step towards alleviating the central issue in measuring cybercrime occurrence: Its darkfield (crimes that take place but cannot be measured; also termed "dark figure") and the concomitant problem with validity.

Prior research theorizes, analyzes and discusses administrative data or survey data in regard to measuring crime (Ariel & Bland, 2019; Baur et al., 2020; Biderman & Reiss Jr, 1967; Comer, Jorgensen, & Carter, 2021; Decker, 1982; Kleinman & Lukoff, 1981; Skogan, 1974, 1977). Research comparing administrative data (also called process-produced data) and survey data – in general and in regard to criminal behavior, such as corruption – suggests that both types rely on the same five steps: Conceptualizing (paradigm, aim, operationalization), gathering data, archiving data, accessing and analyzing data and applying it to the relevant fields (e.g., scientific, policies). However, the characteristics of the data vary within most steps (all, except for archiving). For instance, in the gathering stage, surveys typically aim for random samples whereas administrative data relies on cases noted by an administrative process (such as policing and legal procedures that follow). Moreover, survey data are collected by recruiting respondents whereas administrative data come from registers (Baur, 2009; Baur et al., 2020; Bick & Müller, 1984). Both administrative data and survey data can suffer under issues located at any of the five steps. For example, when gathering survey data, errors can occur if the target population is not covered correctly. Moreover, problems with reliability or validity arise if mistakes are made when designing the survey instrument. Additionally, if analysis techniques are not selected or performed appropriately, errors or biases may arise

(Baur, 2009; Baur et al., 2020). While administrative process-generated data (as gathered during policing and legal processes) are often used as a point of departure for measuring crime, they also come with a number of drawbacks, for example, in the form of biases and errors (Bick & Müller, 1980, 1984). These can be – for instance – based on procedural rules and laws that dictate data collection, data handling and documentation (Baur et al., 2020). One example for this is that, in Germany, no victims-data are included in the registry-datasets for cybercrime, for administrative procedural reasons. Administrative data can also have issues in quality due to, for example, internal inconsistencies, bad data-formatting or outliers. The three dimensions that prior research suggests primarily contribute towards distorting administrative data are: (1) Administrative norms, procedural rules and other such contextual limitations; (2) the clerk and agency producing the data; and (3) the person (or people) affected by the process (Baur, 2009; Baur et al., 2020; Bick & Müller, 1984).

While all data quality issues should be resolved, the current article focusses on issues of validity that arise when measuring (cyber) crime by means of analyzing administrative data or survey data. It is particularly important to address the validity-issue because it is such a strong factor impeding the measurement of the occurrence of "deviant behavior", such as cybercrime. Since the darkfield is an issue encountered by research on crime in general (Skogan, 1977), it comes hand-in-hand with the implication that the validity of crime measurement is impeded. Both administrative and survey data suffer from this issue, although they are missing different cases (of criminal acts). Therefore, I suggest complementarily using both administrative and survey data to study cyber(crime) and uncover more of the actual crime taking place (see Figure 1). This would increase validity and reduce the darkfield.

To do so, this research article takes the previous standards and findings from crime research in general as a point of departure from which to approach the measurement of cybercrime. Therefore, it treats administrative data on cybercrime as a benchmark and looks at whether it is possible to reduce the darkfield of cybercrime, that exists when using this data, by additionally looking at survey data on cybercrime. I call this approach: Data Combination Approach (DCA).

The next sections will proceed as follows: The first section focusses on the conceptual background, including a review of previous research. It begins with a general overview of prior research on measuring "sensitive" topics (e.g., crime) in surveys and the resulting data-quality issues. Moreover, it discusses if administrative data provide solutions. Drawing on prior research, the section continues with data types used in crime research and the issue of the darkfield (including validity). The next subsection focusses on cybercrime, specifically, discussing data types and the darkfield (including validity). This leads directly to the research objectives, followed by the section on methodology, standards and analyses. It begins with a subsection on the conceptual proposition of using multiple data types to tackle the validity issue of cybercrime. The ensuing subsection proposes measuring (cyber)crime occurrence using the Data Combination Approach (DCA), describing steps, standards and how to proceed with analyses (including boundaries and opportunities). The final section draws some first conclusions by discussing the added value, challenges and an outlook towards future research.

Conceptual background and literature review

Prior research on measuring "sensitive" topics such as crime in surveys: Administrative data as a solution?

There are general societal developments which lead to lower trust in science and scientific surveys. This is a major reason why response rates are dropping (increasing the likelihood of nonresponse error) (Koen et al., 2018), and social desirability bias is increasing (also increasing measurement error) (Sakshaug, Yan, & Tourangeau, 2010). When researching sensitive topics (such as criminal behavior or discriminating attitudes) such a lack of respondent-trust is particularly likely to increase nonresponse and socially desirable responding behaviors (social desirability bias) (Stocké & Stark, 2006). In surveys, respondents typically self-report past actions or attitudes. Because responding honestly to a sensitive question can mean admitting that one has acted (or has an attitude) against a social or legal norm, respondents may fear consequences, leading to nonresponse or responses edited from the truth to a more socially desirable outcome (Tourangeau & Yan, 2007). Put plainly: Respondents are motivated to underreport socially undesirable behavior; they lie or provide biased responses, or do not respond at all. These editing-processes can be deliberate or automatic (Kammigan, Enzmann, & Pauwels, 2019; Krumpal, 2013; Wolter, 2012). This can lead to low data quality and biased results (Jann, Krumpal, & Wolter, 2019). Ong and Weiss (2000) show that – even under high anonymity and confidentiality – 25% of respondents still lie about having behaved deviantly. This demonstrates that when researching criminal behavior (including its prevalence and trends) survey data may not show "the whole picture", due to the sensitive nature of crime. Therefore, the question arises: Is administrative crime data a solution to survey datas' social desirability bias in measuring the prevalence of crime? Responding to this question is rather complex. On the one hand, administrative data does not rely on the responses of survey participants. On the other hand, administrative data relies on the reports of victims, the documentation of officials and other "data-recording" stages that can be affected by biases, including social desirability bias. This happens because the involved people are also aware of social and legal norms and may choose to lie, offer biased information or avoid giving specific information at all. Based on the findings of previous research, it seems that neither survey nor administrative data can objectively measure all criminal acts that are committed (Biderman & Reiss Jr, 1967; Çelik, 2021; Konstants, 2022).

Nonresponse or social desirability bias can lead to underestimating the prevalence of behaviors considered to be sensitive, such as different forms of crime, because fewer actions are recorded. This threatens the validity of (cyber)crime data and increases their darkfield – both in the case of survey data and administrative (process-generated) data (Skogan, 1977). Moreover, as presented and discussed in the introduction, both administrative and survey data come with a number of other issues pertaining data quality – both generally and in regard to criminal behavior, such as corruption (Baur et al., 2020; Bick & Müller, 1984). In regard to measuring the prevalence and trends of cybercrime, consideration of prior research, therefore, leads to the conclusion that both administrative and survey data come with a number of methodological drawbacks.

Measuring crime in general: Data types

Several types of data are used to quantitatively measure the occurrence of crime. Two general categories can be identified: (1) Process-generated data and (2) survey data.[2] Each of these categories encompasses a number of data types. For instance, process-generated data includes the data type "administrative data", which are gathered during the law enforcement process and then documented. One example of this is the German Federal Criminal Police Statistic (PKS German Federal Criminal Police Office, 1953-2021). Typically, these data are considered to be the result of several actions: (1) actual crimes that take place, (2) reporting behavior of victims and (3) the investigation, classification and recording by the police (Skogan, 1974). Additionally, differing legal and administrative frameworks of the police need to be considered as a factor of influence when examining administrative crime data (Baur et al., 2020). Administrative data can be information on victims of crime (e.g., number of victims, type of crime, gender, age), on (suspected) offenders of crime (perpetuators) (e.g., number of offenders, type of crime, gender, age) and on cases (e.g., number of cases, type of crime). The category process-generated data also includes the data type "digital behavioral data" (digital trace data), which consists of the digital traces of people's behavior (Veltri, 2020). Both process-generated data types – administrative data and digital behavioral data – are usually "big data" because they are very large datasets, i.e., have a large "volume" (Baur et al., 2020).

Survey data, also termed "victimization surveys", on crime are predominantly collected via population sampling, often focusing on one country or smaller geographical units within a country. The Eurobarometer 92.2 (European Commission and European Parliament, 2019), which inquires about experiencing cybercrime, however, is an example of a survey covering several countries – in this case, all 28 EU-member states (as of 2019, before Great Britain left the EU). Furthermore, factorial surveys on crime are sometimes conducted with samples of specific subpopulations (e.g., students, volunteer firefighters) regarding specific types of crime, such as illegally selling medication or not reporting colleagues' misdeeds (due to a "code of silence") (Kleinewiese & Graeff, 2021; Sattler et al., 2018). Other methods can sometimes be applied, such as quasi-randomized control experiments (Stickle & Felson, 2020). However, there are some limits to the methods of gathering data on crime due to ethical considerations. For instance, it would be highly unethical to conduct an experiment to see if a person physically attacks someone under a given treatment. Moreover, there are legal constraints limiting methods of data collection, for example, researchers need to make sure that they are not accidentally downloading illegal data when measuring crime with digital behavioral data.

By authorities and in crime research, administrative data has typically been used as a benchmark (Comer et al., 2021; Skogan, 1974). This perspective can be challenged (Decker, 1982), particularly because of the large darkfield of crime that

[2] Kleinman and Lukoff (1981), for example, posit a third category of crime data: self-reports by criminals. However, even they contend that such data are likely to be highly unreliable and methodologically weak, particularly, since the reporting offenders are probably motivated to distort or omit information. Based on these fundamental issues, self-reported crime data are not included in this discussion of crime data categories.

remains. However, even critical perspectives concede the utility of administrative crime data for measuring why some geographical places have more crime than others or how the occurrence of crime changes over time (i.e., crime trends) (Skogan, 1974). Moreover, survey data is also criticized for sampling and response biases (e.g., low response rates of people with high income) which reduce reliability and validity of the data (Skogan, 1974). Hence, the majority of crime-occurrence research still treats administrative police data as a benchmark. Building upon both the research in favor of administrative data and that in favor of survey data, the DCA conceptually treats administrative police data as a benchmark and then complements it with an additional data source, such as survey data.

Crime datas' darkfield: A problem with validity

As addressed above, measurements of the occurrence of crime struggle with a darkfield. This leads to issues of validity (Skogan, 1974). Since the darkfield consists of those crimes that are unaccounted for, the extent of the problem depends on the *number* of crimes that take place but are unaccounted for. The DCA takes the validity issue of administrative crime data into account (because of the large, presumed, darkfield) by applying the assumption that combining it with survey data can reduce the darkfield (see Figure 1 and Figure 2).

While administrative data is still prevalently considered to be the benchmark data, there is much disagreement on which data type (administrative or survey) has higher validity (Ariel & Bland, 2019). Ariel and Bland (2019); (Decker, 1982), for instance, criticize that in England and Wales, the crime survey has taken the former position of administrative police data as "national statistics". They posit that there is no clear argument for considering survey data to be more reliable or valid than administrative data. Rather than arguing for one position or the other, the DCA posits that the complementary information from both data types reduces the darkfield of (cyber)crime and, thereby, increases validity.

Measuring cybercrime: Data types

The categories and types of quantitative data on cybercrime are alike to those of other types of crime (see subsection "Measuring crime in general: Data types"). The two categories are: (1) Process-generated data and (2) survey data. Administrative (process-generated) data on cybercrime shares the advantages and drawbacks of administrative data on other types of crime. Furthermore, there is an additional drawback in comparison to administrative data on many other types of crime, at least in Germany: While for many forms of crime, there are administrative data on the number of victims, there are no such victims-data regarding cybercrime. Upon request, the German Federal Police Office stated that this is because in the German Federal Criminal Police Statistic (PKS), information on victims is only recorded if highly personal legal interests are affected. For other crimes (this includes cybercrimes), no victims-data are recorded in the PKS.

Cybercrime can also be investigated using other process-generated data such as digital behavioral data. Considering that these are online data regarding crimes committed via the internet, this data can be highly relevant for research on cybercrime. Such data are often longitudinal. For instance, the Cambridge

Cybercrime Center has collected a number of such datasets leading, for example, to a publication on behavior in underground hacking forums, over a period of time (Pastrana et al., 2018). While it is very useful for research on cybercrime in general, digital behavioral data is – so far and to the best of my knowledge – not used specifically for measuring cybercrime occurrence (in regard to the overall occurrence). One reason is that when using digital behavioral data, it is highly challenging to determine in which country the involved people are (particularly offenders, as they are digitally apt).

Although research on cybercrime is increasing, since it is a rather new phenomenon (compared to other forms of crime, such as house burglary), there are still fewer surveys that make such data available. As with surveys on other types of crime, surveys on cybercrime are "victimization surveys"; they inquire about people's experiences of being a victim of cybercrime. In survey research on cybercrime, there is a lack of cross-country longitudinal survey data (as can be seen in an article by Reep-van den Bergh and Junger (2018). There is a cross-country survey: The Eurobarometer 92.2 (European Commission and European Parliament, 2019), which provides data for examining cybercrime in Europe, according to country. However, it is a cross-sectional dataset. Hence, this is not sufficient for longitudinal analyses. In Germany, starting in 2019, there is a yearly "Digitalbarometer" (Federal Office for Information Security [BSI] and Police Crime Prevention of the Federal States and the Federation [ProPK], 2019-2021), which collects data about cybercrime experiences. This dataset is longitudinal but does not allow for comparisons between countries. Moreover, its estimates of cybercrime occurrence are much lower than that of the Eurobarometer, indicating that it covers less of the darkfield and is, concomitantly, less valid regarding the measurement of cybercrime occurrence.

Figure 1 conceptually shows the relationship of both data types (administrative and survey data) in regard to the darkfield of cybercrime. The front, light rectangle shows the amount of actual cybercrime measured by administrative data. The other light rectangle represents the amount of cybercrime measured by the survey data. There is an overlap of both, which represents the crime that they both measure. The space in the largest rectangle (encompassing the others) that lies outside of the two described rectangles shows that there is actual cybercrime that is, presumably, still not being measured. Hence, this conceptualization posits that the amount of cybercrime not measured (i.e., the darkfield) can be reduced substantially, but not eliminated entirely, by combining several data types for measuring occurrence (DCA).

As in regard to other types of crime, the DCA takes the administrative data on cybercrime as a benchmark and complements it with other data types (such as survey data). This is particularly challenging when the units of interest (i.e., victims, offenders, cases) are not the same across data types. Moreover, some challenges are aggravated in the case of cybercrime data, such as determining in which country to locate the crime or criminal. Also, is important to try and make sure that such aspects are consistent across the data types/sets included, when using the DCA to measure cybercrime occurrence.

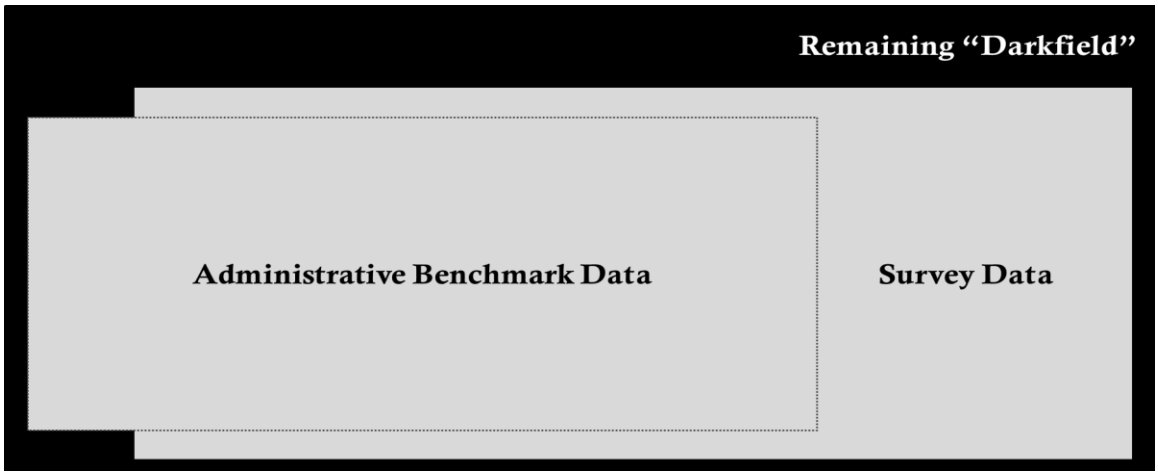

Figure 1. The relationship of administrative data, survey data and the darkfield of (cyber)crime in regard to actual (cyber)crimes committed

Cybercrime datas' darkfield: A problem with validity

Cybercrime data (both survey and administrative) have the same problems with reliability and validity as other crime data (Kshetri, 2013). As addressed above, we run into additional related problems due to the global nature of the internet, that transcends national borders. This makes the measurement of cybercrime even more challenging because not only do we need to identify the location of the crime, criminal or victim (and these may not all be the same) but different legal and social standards challenge a general definition. For example, an act may be considered a crime in the country of the victim but not in the country of the offender. Hence, it may be recorded in the administrative data in one country but not in another. In survey research, the crime may be reported by the victim in one country but not by a victim in another. There are further issues in measuring the occurrence (and trends) of cybercrime, such as that one crime (or case) may have many victims because technology and, particularly, the internet allow for rapid and often automatized distribution of criminal tools (e.g., phishing emails). All of the aforementioned point to particular difficulties in measuring the occurrence of cybercrime. This leads to the common conclusion that the darkfield of cybercrime is especially large. This would also mean that the validity of cybercrime data is particularly low. While this assumption has not been empirically validated, it should be taken into account when measuring the occurrence of cybercrime.

Figure 2 is a conceptualization of how much cybercrime presumably remains in the dark with different approaches to its measurement (not to scale). Figure 2 contains three rectangles, (1)-(3). Each rectangle represents the actual cybercrime occurrence. Each rectangle has a light section which shows the amount of cybercrime presumably measured by a given data type or approach. It is likely, that despite its advantage of measuring actual crimes, the administrative benchmark data leaves the largest darkfield. Other data types, such as survey data, show potential for measuring more of the darkfield but have their own drawbacks, such as no measurement of offenders or cases. By combining both data types, the DCA reduces the darkfield the most.

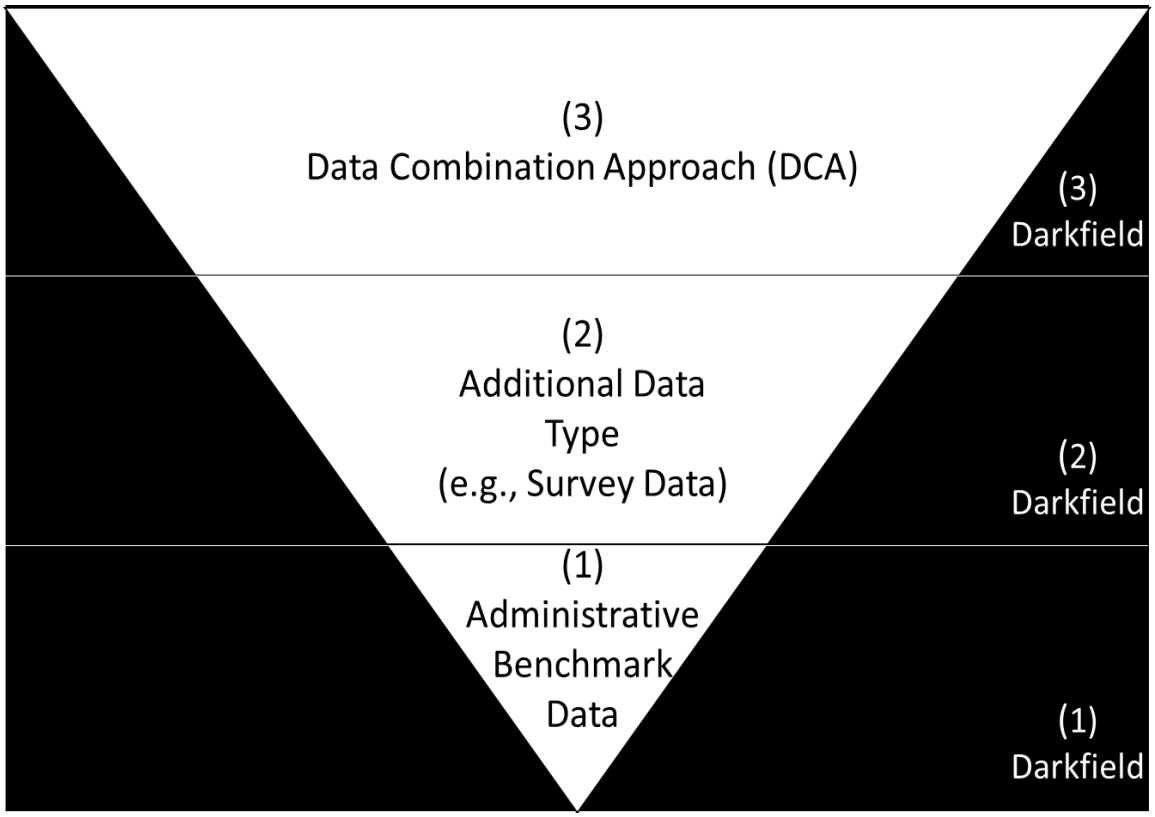

Figure 2. The measured (cyber)crimes and remaining darkfield for each data type/approach

Research Objectives

Based on the theoretical and empirical research in the relevant fields (for example, measuring sensitive topics, such as crime; administrative data – its advantages and issues and validity as well as the darkfield of (cyber)crime) and the considerations I have drawn from these previous studies, the main research objective of this article is to present a conceptual-methodological approach to more valid measurements of sensitive issues, such as cybercrime. Since both administrative and survey data have major advantages and drawbacks in their data quality, instead of using only one data type, it would be expedient to use a "mixed-data" approach which utilizes the respective strengths of each data type. Such an approach should increase the validity of (cyber)crime data. In order to do so, the following section builds on previous methodological literature to propose and discuss such a Data Combination Approach (DCA).

Methodology, standards and analyses

Using multiple data types

A comparison of administrative crime data and survey crime data by Skogan (1974) suggests that both data types are in accordance with underlying crime distributions, i.e., that absolute numbers of crime defer but the ranking of crime occurrence of the examined U.S.-cities is congruent. The DCA builds on this finding, positing that because there are similar crime patterns in different types of data, it is

useful to scrutinize them complementarily. This allows combining the strengths of both data types with each other. For instance, the administrative numbers are population data (not survey data) – recorded during processes that actually took place in a society. This can be complemented with survey data which I assume to capture more of the darkfield of (cyber)crime. All of these assumptions can be applied to examining the occurrence of different types of crime. Since cybercrime is assumed to be particularly underreported in the administrative data, complementing it with survey data is particularly useful for measuring more of the darkfield.

Measuring occurrence using the DCA: Steps and standards

This section focusses on how to measure the occurrence of cybercrime using the DCA. Figure 3 depicts the seven steps that are suggested for measuring (cyber)crime occurrence using the DCA. The steps also contain standards that should be upheld in order to ensure that the results of both datasets/-types can be studied comparatively or in combination, for example, that they are measuring the same geographical units. In steps six and seven, the DCA suggests first analyzing the datasets separately (to gain first insights and compare them) and then in combination (where methodologically possible).

The first step in a DCA is identifying the benchmark dataset. In order to do so, researchers should keep in mind for which units they are trying to measure (cyber)crime prevalence and/or trends. They should determine the geographical region and the unit of measurement. The units of measurement can be victims, offenders or cases. Geographical units in administrative data are often countries or smaller units within countries such as states or regions. When completing this first step, you need to keep in mind that to apply the DCA, you need a fitting additional dataset (such as survey data or digital behavioral data). For instance, if you select administrative data on cybercrime on several regions, you will need survey data measuring cybercrime for each of these regions. This is the intersection of the first and second step. The second step is searching for datasets that could be analyzed complementarily. The third step is identifying the dataset that is most compatible with the administrative benchmark data that you have selected. A typical pitfall to be avoided would be selecting a (survey) dataset on a different time period. The fourth step checks the compatibility of the two datasets regarding minimal requirements for examining the data complementarily, in order to cover more of the darkfield. The fifth step ensures transparency, by asking to clearly document and report the limitations and the similarities between datasets of two datatypes and how these aspects could affect the results of analyses. One such aspect could be how cybercrime is defined, as this affects its measurement. While administrative data adheres to legal and procedural definitions, survey data can use a number of items to measure cybercrime. This can lead to a more broad or narrow understanding of what cybercrime is. Once this step is completed, researchers should analyze each dataset independently and document the results. Analysis techniques may differ between datasets at this stage because survey data often allow for more complex multivariate analyses, whereas administrative cybercrime data are primarily analyzed descriptively. The final step of the DCA is combining the datasets via analyses. Such analyses would be descriptive. They are particularly expedient in regard to measuring trends of cybercrime or prevalence-rankings (e.g., between regions or personal characteristics).

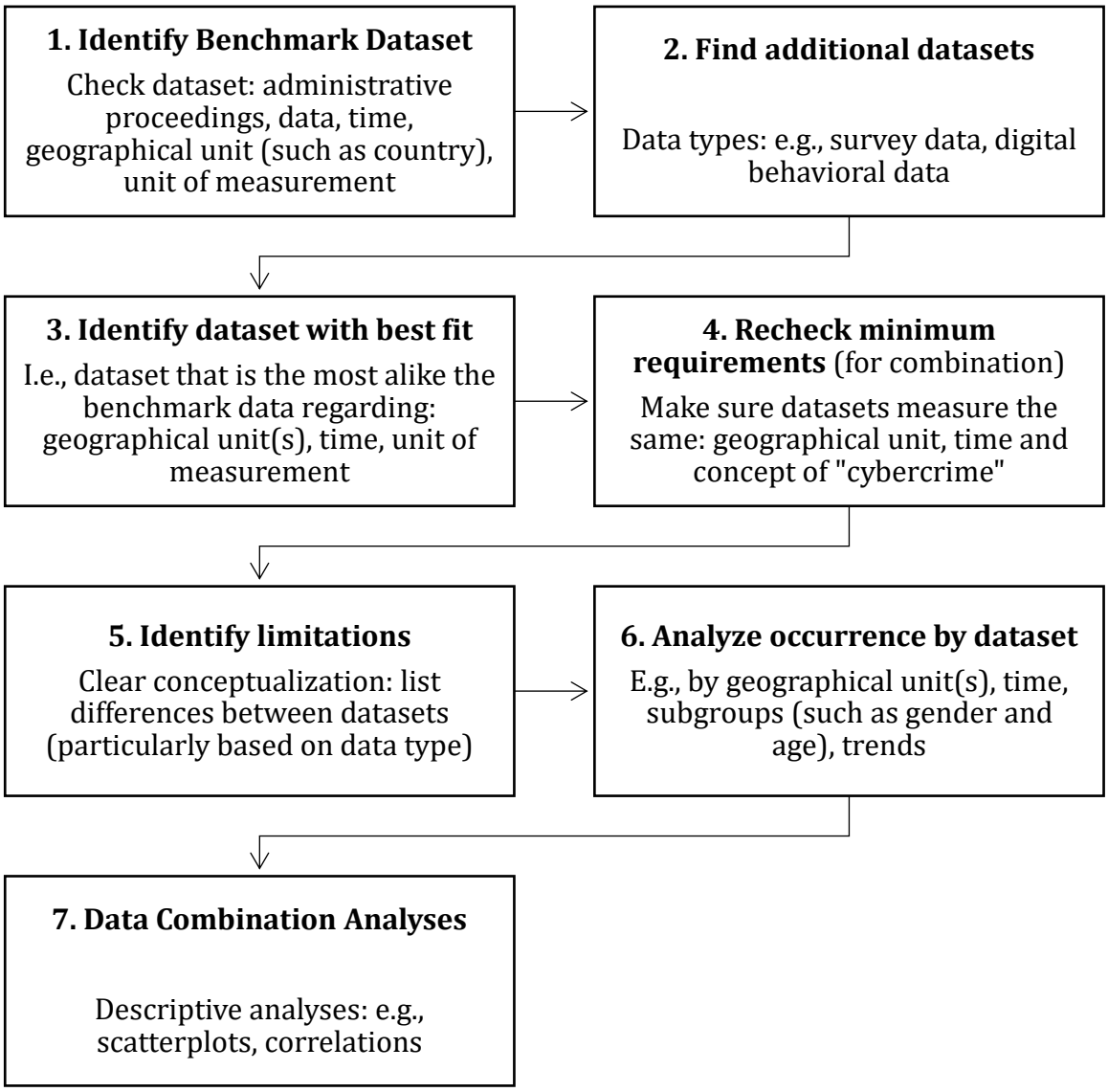

Figure 3. The seven steps and standards of the Data Combination Approach (DCA) to (cyber)crime

Analyses: Boundaries and opportunities

Skogan (1974) suggests that, because administrative crime data do not contain the exact occurrence (incidence) of crime (number of crimes actually committed), multivariate statistics will be affected by measurement error and will not be accurate. However, the administrative data accurately reflect which geographical locations (e.g., country, city) have more crime than others and how the occurrence of crime changes over time (trends). Therefore, Skogan (1974) posits that descriptive analyses such as scatterplots and correlations are the best methods of analysis for administrative crime data, as these would be useful in analyzing variations in crime occurrence and distribution. Taking this argument further, descriptive analyses of administrative crime data can also be used to examine which subgroups within populations commit more crime than others (e.g., according to gender) and how this,

potentially, changes over time (trends). When analyzing survey data on cybercrime, multivariate analyses are possible. However, for the measurement of occurrence, descriptive analyses are usually sufficient. Skogan (1974) empirically shows that rankings of geographical units in regard to crime occurrence are alike across administrative and survey data. This suggests that when analyzing the datasets together (as suggested in Figure 3, Step 7), descriptive analyses must be the selected tools. Hence, a boundary of the DCA is that multivariate analyses are not recommended. The opportunities are that the DCA is expedient for measuring occurrence (including according to smaller subgroups or geographical units), creating rankings and measuring trends.

Conclusions: Added value, challenges and outlook

This research article proposes a Data Combination Approach (DCA) to measuring the occurrence of cybercrime. It argues that the DCA allows for measuring more of the actual cybercrime occurrence than each data type would allow for individually. This effectively reduces the darkfield of cybercrime. In doing so, it also contributes towards tackling the problem with validity in the measurement of cybercrime occurrence. In the DCA, administrative data are used as a benchmark and then complemented with another data type (e.g., survey data).

The current research article contains a conceptual, methodological approach that provides the basis for empirical studies on cybercrime occurrence. Even though real data sets available may often not be perfectly matching to one another (e.g., one longitudinal and one cross-sectional data set), the DCA still allows for a reduction of the darkfield by providing a more informed estimation of cybercrime occurrence. Moreover, in accordance with Skogan (1974) results that rankings of geographical units in regard to crime prevalence as well as trends are alike across administrative and survey data, it appears that DCA applications should allow for descriptive analyses regarding the aforementioned. This article also suggests that these analyses should be refined by examining occurrence and trends according to subgroups.

While, in theory, the DCA is very fruitful because it allows for measuring more of the actual cybercrimes taking place, in empirical applications, researchers may run into a number of issues. The many differences between survey data and administrative data limit their comparability. Moreover, in many cases minimal requirements are hard to fulfill. For instance, one may have cross-sectional survey data for one year and administrative data for a number of years in research on measuring trends of cybercrime. Future methodological research should aim to tackle these limitations, for instance, via computerized methods such as simulations of trends. Moreover, projects collecting survey data on cybercrime should consult administrative data when designing their survey, in order to ensure meeting the requirements. Such developments would enable researchers to apply the DCA with more accuracy and in many contexts.

This research article shows how useful the DCA is in regard to measuring cybercrime occurrence. It is an approach that can and should be developed further in future research. For example, considering a DCA using three or more data types. Finally, while the current focus lies on researching cybercrime, future studies should also consider using the DCA to measure occurrence and trends of other types of crime.

Acknowledgements

The publication of this article was funded by the Mannheim Centre for European Social Research (MZES).

References

Ariel, B., & Bland, M. (2019). Is crime rising or falling? A comparison of police-recorded crime and victimization surveys. In *Methods of criminology and criminal justice research* (Vol. 24, pp. 7-31). Emerald Publishing Limited. https://doi.org/10.1108/S1521-613620190000024004

Baur, N. (2009). Measurement and selection bias in longitudinal data. A framework for re-opening the discussion on data quality and generalizability of social bookkeeping data. *Historical Social Research/Historische Sozialforschung, 34*(3), 9-50. https://doi.org/10.12759/hsr.34.2009.3.9-50

Baur, N., Graeff, P., Braunisch, L., & Schweia, M. (2020). The quality of big data. Development, problems, and possibilities of use of process-generated data in the digital age. *Historical Social Research/Historische Sozialforschung, 45*(3), 209-243. https://doi.org/10.12759/hsr.45.2020.3.209-243

Bick, W., & Müller, P. J. (1980). The nature of process-produced data: towards a social-scientific source criticism. In *Historical social research : the use of historical and process-produced data* (Vol. 6, pp. 369-413). Klett-Cotta. https://www.ssoar.info/ssoar/handle/document/32656

Bick, W., & Müller, P. J. (1984). *Sozialwissenschaftliche Datenkunde für prozeßproduzierte Daten: Entstehungsbedingungen und Indikatorenqualität*. Klett-Cotta. https://www.ssoar.info/ssoar/handle/document/33074

Biderman, A. D., & Reiss Jr, A. J. (1967). On exploring the" dark figure" of crime. *The Annals of the American Academy of Political and Social Science, 374*(1), 1-15. https://doi.org/10.1177/000271626737400102

Buil-Gil, D., Miró-Llinares, F., Moneva, A., Kemp, S., & Díaz-Castaño, N. (2021). Cybercrime and shifts in opportunities during COVID-19: a preliminary analysis in the UK. *European Societies, 23*(sup1), S47-S59. https://doi.org/10.1080/14616696.2020.1804973

Çelik, A. (2021). 'Keep your mouth shut in the day and your door shut at night.'Intra-Kurdish Violence in the Shadow of the State: The case of Hizbullah in Kurdistan of Turkey. *Kurdish Studies, 9*(1), 37-57. https://doi.org/10.33182/ks.v9i1.563

Comer, B. P., Jorgensen, C., & Carter, D. (2021). Reported Crime Frequencies: A Statistical Comparison of State Crime Reports and the UCR. *American Journal of Criminal Justice*, 1-25. https://doi.org/10.1007/s12103-021-09623-y

Decker, S. H. (1982). Comparing victimization and official estimates of crime: a re-examination of the validity of police statistics. *American Journal of Police, 2*(2), 193-202. https://www.ojp.gov/ncjrs/virtual-library/abstracts/comparing-victimization-and-official-estimates-crime-re-examination

Di Nicola, A. (2022). Towards digital organized crime and digital sociology of organized crime. *Trends in Organized Crime*, 1-20. https://doi.org/10.1007/s12117-022-09457-y

European Commission and European Parliament, B. (2019). *Eurobarometer 92.2 (2019)*. Cologne: GESIS Data Archive. https://doi.org/10.4232/1.13657

Federal Office for Information Security [BSI] and Police Crime Prevention of the Federal States and the Federation [ProPK]. (2019-2021). *Digital Barometer*. Berlin: Ipsos Public Affairs. https://www.ipsos.com/en-us

Ghazi-Tehrani, A. K., & Pontell, H. N. (2021). Phishing evolves: Analyzing the enduring cybercrime. *Victims & Offenders, 16*(3), 316-342. https://doi.org/10.1080/15564886.2020.1829224

Jann, B., Krumpal, I., & Wolter, F. (2019). Social Desirability Bias in Surveys–Collecting and Analyzing Sensitive Data. Special Issue. *MDA Methods, Data & Analyses, 13*(1), 3-6. https://mda.gesis.org/index.php/mda/article/view/247

Kammigan, I., Enzmann, D., & Pauwels, L. J. (2019). Over-and underreporting of drug use: A cross-national inquiry of social desirability through the lens of situational action theory. *European Journal on Criminal Policy and Research, 25*(3), 273-296. https://doi.org/10.1007/s10610-018-9397-y

Kleinewiese, J., & Graeff, P. (2021). Ethical decisions between the conflicting priorities of legality and group loyalty: scrutinizing the "code of silence" among volunteer firefighters with a vignette-based factorial survey. *Deviant Behavior, 42*(10), 1228-1241. https://doi.org/10.1080/01639625.2020.1738640

Kleinman, P. H., & Lukoff, I. F. (1981). Official crime data: Lag in recording time as a threat to validity. *Criminology, 19*(3), 449-454. https://doi.org/10.1111/j.1745-9125.1981.tb00429.x

Koen, B., Loosveldt, G., Vandenplas, C., & Stoop, I. (2018). Response rates in the European Social Survey: Increasing, decreasing, or a matter of fieldwork efforts? *Survey methods: Insights from the field*, 1-12. https://doi.org/10.13094/SMIF-2018-00003

Konstants, L. (2022). The Impact of Economic Sanctions (on Russia) on Labor Migrants from Kyrgyzstan and on their Remittances. *Remittances Review, 7*(2), 129-151. http://doi.org/10.47059/rr.v7i2.2416

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & quantity, 47*(4), 2025-2047. https://doi.org/10.1007/s11135-011-9640-9

Kshetri, N. (2013). Reliability, validity, comparability and practical utility of cybercrime-related data, metrics, and information. *Information, 4*(1), 117-123. https://doi.org/10.3390/info4010117

Lusthaus, J., Bruce, M., & Phair, N. (2020). Mapping the geography of cybercrime: A review of indices of digital offending by country. In *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)* (pp. 448-453). IEEE. https://doi.org/10.1109/EuroSPW51379.2020.00066

Ngo, F. T., & Paternoster, R. (2011). Cybercrime victimization: An examination of individual and situational level factors. *International Journal of Cyber Criminology, 5*(1), 773-793. https://digitalcommons.usf.edu/cjp_facpub_sm/17

Ong, A. D., & Weiss, D. J. (2000). The impact of anonymity on responses to sensitive questions 1. *Journal of Applied Social Psychology, 30*(8), 1691-1708. https://doi.org/10.1111/j.1559-1816.2000.tb02462.x

Pastrana, S., Thomas, D. R., Hutchings, A., & Clayton, R. (2018). Crimebb: Enabling cybercrime research on underground forums at scale. In *Proceedings of the 2018 World Wide Web Conference* (pp. 1845-1854). ACM Digital Library. https://doi.org/10.1145/3178876.3186178

PKS German Federal Criminal Police Office. (1953-2021). *German Federal Criminal Police Statistic*. Wiesbaden. https://www.bka.de/DE/Home/home_node.html

Porcedda, M. G., & Wall, D. S. (2021). Modelling the Cybercrime Cascade Effect in Data Crime. In *2021 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)* (pp. 161-177). IEEE. https://doi.org/10.1109/EuroSPW54576.2021.00025

Reep-van den Bergh, C. M., & Junger, M. (2018). Victims of cybercrime in Europe: a review of victim surveys. *Crime science, 7*(1), 1-15. https://doi.org/10.1186/s40163-018-0079-3

Sakshaug, J. W., Yan, T., & Tourangeau, R. (2010). Nonresponse error, measurement error, and mode of data collection: Tradeoffs in a multi-mode survey of sensitive and non-sensitive items. *Public Opinion Quarterly, 74*(5), 907-933. https://doi.org/10.1093/poq/nfq057

Saridakis, G., Benson, V., Ezingeard, J.-N., & Tennakoon, H. (2016). Individual information security, user behaviour and cyber victimisation: An empirical study of social networking users. *Technological Forecasting and Social Change, 102*, 320-330. https://doi.org/10.1016/j.techfore.2015.08.012

Sattler, S., Graeff, P., Sauer, C., & Mehlkop, G. (2018). Der illegale Verkauf verschreibungspflichtiger Medikamente zur kognitiven Leistungssteigerung–Eine Vignetten-basierte Studie rationaler und normativer Erklärungsgründe. *Monatsschrift für Kriminologie und Strafrechtsreform, 101*(3-4), 352-379. https://doi.org/10.1515/mks-2018-1013-408

Skogan, W. G. (1974). The validity of official crime statistics: An empirical investigation. *Social Science Quarterly, 55*(1), 25-38. https://www.jstor.org/stable/42859308

Skogan, W. G. (1977). Dimensions of the dark figure of unreported crime. *Crime & Delinquency, 23*(1), 41-50. https://doi.org/10.1177/001112877702300104

Stickle, B., & Felson, M. (2020). Crime rates in a pandemic: The largest criminological experiment in history. *American Journal of Criminal Justice, 45*(4), 525-536. https://doi.org/10.1007/s12103-020-09546-0

Stocké, V., & Stark, T. (2006). Trust in surveys and the respondents' susceptibility to item nonresponse. *Rationalitätskonzepte, Entscheidungsverhalten und ökonomische Modellierung, 6*. https://madoc.bib.uni-mannheim.de/2601/

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological bulletin, 133*(5), 859–883. https://doi.org/10.1037/0033-2909.133.5.859

Veltri, G. A. (2020). *Digital social research*. Polity Press. https://www.abebooks.it/9781509529315

Wolter, F. (2012). Heikle Fragen in Interviews. In *Heikle Fragen in Interviews* (pp. 27-118). Springer. https://doi.org/10.1007/978-3-531-19371-7_2