Chapter 14
Are Dark Number Estimates of Crime Feasible and Useful?

Joras Ferwerda

Introduction

Criminals try to minimize the detection of their crimes. This makes crime hard to study, especially when attempting an overview, such as estimating the volume. For certain types of crime, few offenses are observed because the effects are not directly visible and detection rates are low—like for money laundering, terrorist financing, tax evasion, and certain types of fraud—or because both parties might prefer the transaction like for drug sales and illegal prostitution. But how much are we missing when focusing only on the known cases? While policymakers seek dark number estimates of the total volume, there is skepticism among academics about whether such estimates are valid, reliable, feasible, or even useful. This chapter explores whether such dark number estimates specifically for money laundering, which is rarely detected, are feasible or useful. To make the discussion more concrete, this chapter starts with outlining the critique on existing estimates of money laundering as an example. A normative discussion sets out the relevant evaluation principles. This chapter concludes dark number estimates should be attempted.

I thank Peter Reuter and Brigitte Unger for helpful comments.

J. Ferwerda (✉)
Utrecht University School of Economics, Utrecht, The Netherlands
e-mail: j.ferwerda@uu.nl

H. Nelen, D. Siegel (eds.), *Organized Crime in the 21st Century*,
https://doi.org/10.1007/978-3-031-21576-6_14

Critique on Dark Number Estimates of Crime

In the early 2000s, Naylor (2004) voiced his concern about politicians, press, and police telling the public about an "epidemic of crime sweeping through the increasingly globalized economy" and how such messages "help to justify a string of remarkable legislative and political initiatives" (Naylor, 2004, p. 1). It can indeed be troubling how dark number estimates of crime are sometimes used. However, this issue is broader than dark number estimates of crime; general crime narratives and examples of crime events and the occurrence of criminal gangs can be used in the same worrying manner. This chapter will focus on dark number crime estimates and the academic discussion and leave out how such estimates are or can be used in policy debates.

Taking the discussion around estimates of money laundering as an example, the most influential and in-depth critique on dark number estimates comes from Reuter (2013, p. 224). He states "[m]y view is that knowing how much money is laundered serves no important policy purpose." Continuing with: "It is simply one of those adornments for conversations about the phenomenon. That is very fortunate, since we have no methodology that plausibly would produce credible numbers. [..] [T]here is no prospect, either in surveys of experts or in studies of crimes themselves as reflected in criminal justice statistics, for developing persuasive estimates." A strong take on the matter from one of the most influential scholars in the field of money laundering.

Reuter (2013) explains why the amount of money laundering serves no important policy purpose. The eventual goal of anti-money laundering controls is not reduction in the volume of money laundering; it is about reducing predicate crimes and "global bads" such as terrorism and kleptocracy. Money laundering volume reduction should not be the primary measure of performance for agencies. Such a performance measure does not align with the eventual goal (reducing predicate crimes and global bads) and leads to unproductive strategic behavior, like picking low-hanging fruit.

But besides questioning the "why-question," both Reuter (2013) and Kruisbergen (2021) question the methods now in use for estimating money laundering. The existing estimates need some heroic assumptions and have such a range, that it is hard to use them for any policy or performance measure. When focusing on the leading model used for estimating money laundering, the Walker Model or the Walker-Unger Model, some assumptions and data sources are not well documented. Ferwerda et al. (2020) made the estimation procedure more transparent. Previous publications about the estimation model relied on talking the reader through the estimation procedure. While such a narrative needs to be precise, it also needs to remain readable. This can lead to small, seemingly uninteresting details being left out to keep the main text readable or to such details ending up in footnotes or appendices, which all makes it harder to replicate the whole estimation procedure for an outsider. Ferwerda et al. (2020) publishing in a journal of *Nature* were required to provide a detailed description of the estimation methodology. So, this estimation script is now downloadable with related data on the website of the publication.[1] Although this makes the estimation procedure more transparent and replication easier, Ferwerda et al. (2020) still relied on some criticized data sources, specifically expert judgment, from these former models.

Expert Judgment

Estimates of money laundering volume have relied heavily on expert judgment, since reliable data are missing. Reuter (2013) specifically criticizes expert judgment of fraud experts with only a 10% response rate, as used in the Walker model. Van Duyne and Soudijn (2010) doubt whether experts know much about relevant estimates for money laundering components. Such critique is in line with Kilmer et al. (2010) criticism of intelligence agency efforts to estimate US drug markets. Indeed, the current models have issues with using expert judgment. Comparing the expert elicitation used for estimating money laundering with the standards and best practices of Morgan (the leading scholar on this, see e.g. Morgan, 2014; Morgan et al., 1990; Morgan & Keith, 1995) shows how much improvements are still needed.

The dependence on expert elicitation reflects a lack of good data, even in official statistics, that are needed for a money laundering estimation. We can be optimistic and just claim that we have to wait for better data from national or international institutions, but we have to be realistic that such data needs focused academic research, since money laundering is by definition hidden from the eyes of official authorities.

The last issue worth discussing is research capacity. Kruisbergen (2021) argues that when research capacity is limited and when funded with public means, research budgets are better spent on other research questions than an estimation of money laundering that cannot be done well. Reuter (2013, p. 230) concludes in the same vein: "estimating the total volume [of money laundering] is a diversion of attention."

Before focusing on the feasibility, we first need to question whether we should want to have dark number crime estimates in the first place.

Are Dark Number Estimates of Crime Useful? Principles to Determine Whether Dark Number Estimates Are Worth Doing

For many types of crimes, we only see the tip of the iceberg: the cases that are known either because law enforcement agencies detected and investigated them or because criminologists were able to uncover them. These known cases provide the

[1] See: Ferwerda et al. (2020). Estimating money laundering flows with a gravity model-based simulation. Scientific Reports, 10(1), 1–11. https://doi.org/10.1038/s41598-020-75653-x

most concrete possibility to gain relevant insights into these crimes. Such observations are a logical starting point and can answer many relevant questions, but for some of the bigger research questions, it might not be the right focus. Criminologists can criticize law enforcement agencies for picking low-hanging fruit (see, e.g., Reuter, 2013). We, as a science, should then be open to the same critique. Without claiming that research on known cases is easy, we might be trying to pick the low-hanging fruit here. Good research starts by identifying what we want to know. For some research questions, we need the bigger picture. Dark number estimates of crime are hard, but the size of the challenge should not determine whether they need to be done. The challenge is a separate issue discussed later in this chapter, after we determine whether dark number estimates should be done at all.

We can use *relevance* as a guiding principle on whether research should be conducted. Relevance can be classified into two key types: scientific relevance, where a study increases our understanding, and societal relevance, where society directly benefits as a result of this increased understanding (Shaw & Elger, 2013). Not both types of relevance are needed to justify research. Fundamental research without a direct benefit to society is still relevant if it contributes to our understanding of human or social behavior. Dark number estimates of crime can be relevant for policy priority setting, policy evaluation, and further research. The first two allude to the societal relevance, while the latter indicates the scientific relevance.

Policy Priority Setting

Crimes and their related harms are important issues for societies. We need to understand crime well and formulate effective policies against it. But governments face many challenges with a limited budget, so setting priorities is key. Politicians and policymakers need to decide which problems will need to be tackled first and how much of the limited budget can be spent to deal with each issue. For an informed and justified priority decision, it is crucial to know the size of each problem and what its effects (both positive and negative) are. Dark number crime estimates—when credible, reliable, and sufficiently accurate—are relevant inputs for such policy decisions.

For money laundering estimates, the policy priority setting has become more explicit since 2012. The FATF, the international organization for the fight against money laundering, made it explicit that countries should understand money laundering risks for a proper risk-based fight against money laundering (FATF, 2012, 2013). Money laundering estimates can be helpful to understand (relative) money laundering risks within a country—especially when disaggregated for different sectors, products, and institutions—so that anti-money laundering policies can be truly risk-based (Ferwerda & Reuter, 2022, p. 24).

Policy Evaluation

For crime control policies, as for any other policies, the central question is whether such policies are effective. Since "water always finds its way" (Unger & den Hertog, 2012), criminals can react to new policies and try to circumvent them. Estimating the amount of crime properly is relevant to evaluating ex-post whether the policies have been effective and efficient. Did the amount of crime really decrease after the policy was introduced (effectiveness)? And was this worth the spending on the policy (efficiency)? The global anti-money laundering policy framework remains untested after 30 years, raising questions on whether the same results could be achieved more cheaply through other means (Ferwerda & Reuter, 2022). Dark number estimates of money laundering—when credible, reliable, and sufficiently accurate—are relevant inputs for such evaluations.

There has been a long discussion about how well policies can be evaluated and how to determine the causal relationship of policy and the effect. Policy evaluation as a field is still very much in development (see, e.g., Gertler et al., 2016; Nagel, 2002; Rossi et al., 2018). Determining the effectiveness of anti-money laundering policies is particularly challenging since much of the focus is on preventing money laundering. How to determine whether something did not happen due to the measures taken by authorities? (Nelen, 2008, p. 20) More and more focus on *evidence-based* policymaking and evaluating the effectiveness of policies can have disadvantages. Nelen (2008) warns of the danger of an *evidence maze,* where the social scientist is merely asked to provide the evidence for what works, while questions like why it works, for whom, and under which circumstances are more relevant to truly expand scientific insights. Indeed, measuring the effect of crime-prevention policies is hard, especially when there are poor measures of the volume of offenses (see, e.g., Van der Schoot, 2006).

However, there has been significant progress in the field called causal inference. The improvements in empirical research designs led Angrist and Pischke (2010) to even talk about a credibility revolution. This chapter cannot be the place to determine whether we should be optimistic or pessimistic about eventually being able to reliably estimate the prevention effect of policies generally. Consider the following thought experiment. If a policeman stands on a street corner for 40 years and never catches a criminal, this might seem a wasted resource. But it can also be argued that the police man prevented many crimes; criminals that therefore never had to be caught in the first place. How can we ever know the prevention effect? We could analyze what happens on days the police man did not attend the street corner (expectedly, like during weekends, and unexpectedly, due to, e.g., illness) and compare that with the other days. Theoretically at least, we could analyze many similar street corners (and account for the differences) with and without policemen and use econometrics or other data analytical tools to get an estimate of the prevention effect. We could go even further and randomize which street corners are watched by policemen, and when, to get more insights into the prevention effect (e.g., the work of Famega et al., 2017 is an inspiring example or randomization to test policing effects). Such a thought experiment shows how such an evaluation study is definitely not easy, but also how it is not impossible to get some relevant insights.

The answer to the *what-works*-question should not be the end product of scientific research. It should spur questions about why it works, how it works and in which circumstances (in line with Nelen, 2008, p. 23). Answering the *what-works*-question should eventually lead to new scientific insights.

Scientific Relevance

In theory, money laundering may have all kinds of (indirect) effects on societies. Money laundering can affect the real economy by distortion of consumption, savings, investment, inflation, competition, trade, and employment. Furthermore, money laundering can affect the financial sector with an increased risk to the solvability, liquidity, reputation, and integrity of the sector. But money laundering, on the other hand, could also be good for an economy, for example, because it increases the profits for the financial sector and leads to greater availability of credit (see Unger et al., 2006 for an overview of the effects of money laundering). Many of these effects can easily be identified in theoretical models, but empirical support for these effects is often lacking (Ferwerda, 2013). The empirical research on these effects, and especially the research on how big these effects are, is still in its infancy. Empirical research on the effects of money laundering is complicated because the most important variable (the cause) is so hard to estimate: the amount of money laundering. Dark number estimates of crime are relevant (or crucial, even) inputs to eventually test the suggested effects and unintended side effects (in line with the argumentation of Ferwerda et al., 2021 on illicit financial flows).

Research attempting to estimate the size forces researchers to think about underlying mechanisms and relevant determinants. This can generate related useful insights. For example, Ferwerda et al. (2020) simulated money laundering flows to estimate the amount of money laundering, but the real value might be the (limited) insight derived about which type of country money launderers prefer to send their money. A similar example is the study of Caulkins and Reuter (2022) that provides rough estimates for the volume of the heroin market in British Columbia. The main insight of that study is not the estimation itself but an understanding of the primary determinants for such an estimation and the role of anti-money laundering policy.

Even critics of money laundering estimates admit that wanting to know how big the problem is an understandable and rational wish (Kruisbergen, 2021, p.64). The main issue is then whether it is possible to have credible and reliable dark number estimates of crime.

Are Dark Number Estimates of Crime Feasible? Can We Have Reliable and Precise Estimates?

Will we ever be able to measure crime precisely? No. It is hard to imagine we would ever be able to have a measurement technique to know exactly how much crimes like money laundering are happening. Still, with credible estimates, we can get a better idea of the size, with possible flaws and uncertainty, but better. This inability to be precise is not unique to the estimation of the amount of crime.

Take the coastline paradox. It seems like we can see the length of a coastline on a map, but the coastline of a landmass does not have a well-defined length. The smaller the unit of measurement (200 km > 100 km > 50 km), the higher the estimated length. Figure 14.1 visualizes the paradox. Of course, on top of that, coastlines change, daily with the tides, but also over time. It is impossible to determine the exact length of a coastline. Still, imprecise coastline estimates are useful and certain statements based on such estimates can be credible. The estimates can indicate whether the coastline of an island is shrinking due to rising sea levels or that the coastline of the Netherlands is shorter than the coastline of the United Kingdom. Acceptance of a certain level of imprecision is needed to use such estimates.

Without going into the philosophical debate on whether we ever know anything (see the Gettier problem), we might want to focus on a more relevant question: can we produce sufficiently credible and reliable dark number crime estimates? Now or in the future.

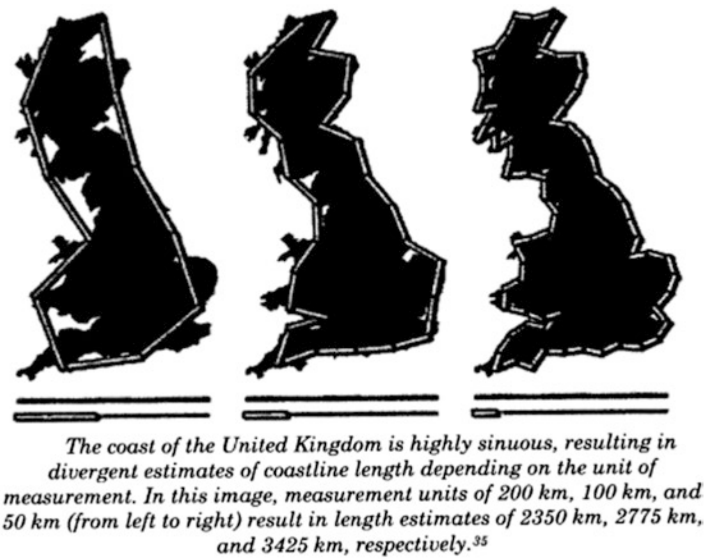

Fig. 14.1 The coastline paradox visualized. (Source: Stoa, 2019, p. 359)

At the moment, we lack the understanding and the data to have an estimate precise enough that it can be used for more than a rough idea about the amount of money laundering (see Reuter (2013) and other critics (e.g., Kruisbergen, 2021; Van Duyne, 2003; Van Duyne & Soudijn, 2010)). We might be able to answer questions about the order of magnitude (are we talking about millions or billions?), but current estimates cannot be used to test the effectiveness of anti-money laundering policies reliably. Dark number crime estimates still rely on strong assumptions and have large error margins. We might be stumbling at the gate, and that is fine. Dark number crime estimates are still in development.

There are many other examples where it initially seemed very hard to get reliable and credible estimates, like when estimating happiness (see, e.g., National Research Council, 2014; Voukelatou et al., 2021) and unemployment (see, e.g., Card, 2011). In these fields, there is still discussion about the precision of the estimates and adaptations are still being suggested (mostly on disentanglement). However, standardized estimation methods have been developed over time, with trial and error, to come to useable statistics useful for policymakers and related research. For unemployment estimates, the development started already in 1880. Largely accepted measurement techniques developed in the late 1930s (Card, 2011). Even today, most unemployment rates are derived from national statistics and are therefore not directly comparable. Only since 2001 has the OECD published the so-called harmonized employment rate,[2] a more internationally comparable unemployment measure. The development of feasible estimation methods takes time and trial and error. We have had insufficient of both for the development of money laundering estimates.

Estimates of unemployment and happiness became useful eventually. Critique about the estimates is needed and, when constructive, can help eventually, but there is little value in suggesting to stop trying. We are still at the initial phase where attempts and a diversity of methods should be welcomed. Failures should even be expected. Trial and error are needed to eventually come to a reliable and credible estimation model. In the same vein, we still need to determine which data are needed and how to reliably get to this data. We cannot expect to already have the necessary data available. The UNODC is working on making crime statistics more internationally comparable and recently published a conceptual framework for the statistical measurement of illicit financial flows (UNODC, 2020). Such endeavors might not directly provide all the relevant data that is needed but show progress in data collection and standardization.

[2] https://stats.oecd.org/glossary/detail.asp?ID=3094 (accessed May 28, 2022).

Conclusion

Some crimes and their effects cannot be observed directly. There is skepticism about dark number estimates of these crimes and how these estimates can be used by politicians, the press, and police. There are legitimate concerns about dark number estimates of crime, but being unable to precisely estimate crimes with current models and data should not be a reason to stop research. Whether to conduct dark number crime estimates should be determined by their relevance. Dark number estimates of crime can be relevant for three main reasons. First, having an estimate of the size of the problem can help politicians and policymakers determine whether scarce resources should be spent to fight the problem and how much. Second, estimates of the amount of crime are crucial to eventually test the effectiveness of crime-fighting policies. Third, dark number crime estimates are scientifically relevant to eventually better understand, among others, the underlying mechanisms, the determinants, and the effects. We might currently be unsure about which data to use, uncertain about the right research method, and in doubt about the appropriate model, but these doubts and uncertainties should not terminate the search. Actually, doubt and uncertainty should be a driving force for research.

References

Angrist, J. D., & Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives, 24*(2), 3–30.

Card, D. (2011). Origins of the unemployment rate: The lasting legacy of measurement without theory. *American Economic Review, 101*(3), 552–557.

Caulkins, J. P., & Reuter, P. (2022). How much demand for money laundering services does drug selling create? Identifying the key parameters. *International Journal of Drug Policy, 103*, 103652.

Famega, C., Hinkle, J. C., & Weisburd, D. (2017). Why getting inside the "black box" is important: Examining treatment implementation and outputs in policing experiments. *Police Quarterly, 20*(1), 106–132.

FATF. (2012). *International standards on combating money laundering and the financing of terrorism & proliferation: The FATF recommendations*. FATF/OECD. http://www.nrb.org.np/fiu/notices/Revised_FATF_Recommendations(2Feb2012)_FATF-WGEI(2012).pdf

FATF. (2013). *FATF guidance national money laundering and terrorist financing risk assessment*. http://www.fatf-gafi.org/media/fatf/content/images/National_ML_TF_Risk_Assessment.pdf

Ferwerda, J. (2013). 3. The effects of money laundering. In *Research handbook on money laundering* (p. 35). EconPapers

Ferwerda, J., & Reuter, P. (2022). *National assessments of money laundering risks* (Equitable growth, finance & institutions insight, p. 45) [World Bank report]. World Bank.

Ferwerda, J., van Saase, A., Unger, B., & Getzner, M. (2020). Estimating money laundering flows with a gravity model-based simulation. *Scientific Reports, 10*(1), 1–11.

Ferwerda, J., Unger, B., Rossel Flores, L., & Ferwerda, J. (2021). How big are illicit financial flows? The hot phase of IFF estimations. In *Combatting financial fraud and empowering regulators*. Oxford University Press.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2016). *Impact evaluation in practice*. World Bank Publications.
Kilmer, B., Caulkins, J. P., Bond, B. M., & Reuter, P. H. (2010). *Reducing drug trafficking revenues and violence in Mexico: Would legalizing marijuana in California help?* Rand.
Kruisbergen, E. (2021). Het witwassen van criminele winsten. *Justitiele Verkenningen, 47*(4), 47.
Morgan, M. G. (2014). Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National Academy of Sciences, 111*(20), 7176–7184.
Morgan, M. G., & Keith, D. W. (1995). Subjective judgments by climate experts. *Environmental Science & Technology, 29*(10), 468A–476A.
Morgan, M. G., Henrion, M., & Small, M. (1990). *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge University Press.
Nagel, S. S. (2002). *Handbook of public policy evaluation*. Sage.
National Research Council. (2014). *Subjective well-being: Measuring happiness, suffering, and other dimensions of experience*. National Research Council.
Naylor, R. T. (2004). *Wages of crime: Black markets, illegal finance, and the underworld economy*. Cornell University Press.
Nelen, H. (2008). *Evidence maze; het doolhof van het evaluatieonderzoek*. Maastricht University.
Reuter, P. (2013). 18. Are estimates of the volume of money laundering either feasible or useful? In *Research handbook on money laundering* (p. 224). EconPapers.
Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2018). *Evaluation: A systematic approach*. Sage publications.
Shaw, D. M., & Elger, B. S. (2013). The relevance of relevance in research. *Swiss Medical Weekly, 19*. https://doi.org/10.4414/smw.2013.13792
Stoa, R. B. (2019). The coastline paradox. *Rutgers University Law Review, 72*, 351.
Unger, B., & den Hertog, J. (2012). Water always finds its way – Identifying new forms of money laundering. *Crime Law and Social Change, Special Issue on Black and Grey Finance, 57*, 287.
Unger, B., Siegel, M., Ferwerda, J., de Kruijf, W., Busuioic, M., Wokke, K., & Rawlings, G. (2006). *The amounts and effects of money laundering*. Ministry of Finance. www.Minfin.Nl/Binaries/Minfin/Assets/Pdf/Actueel/Bijlage-Nieuwsberichten/2006/02/06-011a.Pdf; http://www.ftm.nl/wp-content/uploads/2014/02/witwassen-in-nederland-onderzoek-naar-criminele-geldstromen.pdf
UNODC. (2020). *Conceptual framework for the statistical measurement of illicit financial flows* (Produced by UNCTAD (Development Statistics and Information Branch in Consultation with Debt and Development Finance Branch) and UNODC (Research and Trends Analysis Branch)). https://www.unodc.org/documents/data-and-analysis/statistics/IFF/IFF_Conceptual_Framework_for_publication_FINAL_16Oct_print.pdf
Van der Schoot, C. R. A. (2006). *Organised crime prevention in the Netherlands: Exposing the effectiveness of preventive measures*. Eleven International Publishing.
Van Duyne, P. C. (2003). Money laundering policy, fears and facts. In P. C. van Duyne, K. von Lampe, & J. L. Newell (Eds.), *Criminal finances and organizing crime in Europe* (pp. 67–104). Wolf Legal Publishers. http://www.petrusvanduyne.nl/files/Witwassen%20Fears%20and%20Facts.pdf
Van Duyne, P. C., & Soudijn, M. R. J. (2010). Crime-money in the financial system: What we fear and what we know. In *Transnational criminology manual* (Vol. 2). Wolf Legal Publishers.
Voukelatou, V., Gabrielli, L., Miliou, I., Cresci, S., Sharma, R., Tesconi, M., & Pappalardo, L. (2021). Measuring objective and subjective well-being: Dimensions and data sources. *International Journal of Data Science and Analytics, 11*(4), 279–309.