

RESEARCH

Open Access



Exploring the impact of measurement error in police recorded crime rates through sensitivity analysis

Jose Pina-Sánchez^{1*} , Ian Brunton-Smith², David Buil-Gil³ and Alexandru Cernat³

Abstract

It is well known that police recorded crime data is susceptible to substantial measurement error. However, despite its limitations, police data is widely used in regression models exploring the causes and effects of crime, which can lead to different types of bias. Here, we introduce a new R package ('rcme': Recounting Crime with Measurement Error) that can be used to facilitate sensitivity assessments of the impact of measurement error in analyses using police recorded crime rates across a wide range of settings. To demonstrate the potential of such sensitivity analysis, we explore the robustness of the effect of collective efficacy on criminal damage across Greater London's neighbourhoods. We show how the crime reduction effect attributed to collective efficacy appears robust, even when most criminal damage incidents are not recorded by the police, and if we accept that under-recording rates are moderately affected by collective efficacy.

Keywords Police data, Crime rates, Under-reporting, Bias

Introduction

Few social phenomena are as consequential as crime. For example, crime is explored as a precursor of key constructs across all Social Science disciplines, such as economic growth (Detotto & Otranto, 2010), residential segregation (Keels et al., 2005), academic performance (Boxer et al., 2020), or public health (Janke et al., 2016). However, one major methodological problem has, and continues to impact the robustness of studies exploring the causes and effects of crime. Many studies are still largely dependent on police recorded crime data, which is affected by different forms of measurement error (Coleman & Moynihan, 1996; Lohr, 2019). Namely,

a large share of crime fails to be counted as it remains unreported or undetected (Hart & Rennison, 2003; Tarling & Morris, 2010), while standards of crime recording (rules and everyday practices) within and between police forces have been shown to be inconsistent (Boivin & Cordeau, 2011; Eterno et al., 2016; Her Majesty Inspectorate of Constabulary, 2014).

The measurement error mechanisms impacting police recorded crime are complex, with Gibson and Kim (2008), Pudney et al. (2000) and Pina-Sánchez et al. (2022) showing that they are, at a minimum, systematic (under-estimates because of under-reporting and under-recording that may vary across police forces/areas in unpredictable ways) and multiplicative (errors are proportional to the underlying level of criminality), but may also be differential (under-reporting may be related to other features of the area). These types of errors are likely to have a substantial biasing effect in causal estimates of statistical models relying on police data. Yet, despite the problem being well-known, relatively few studies have

*Correspondence:

Jose Pina-Sánchez
j.pinasanchez@leeds.ac.uk

¹ School of Law, University of Leeds, Leeds, UK

² Department of Sociology, University of Surrey, Guildford, UK

³ School of Social Sciences, University of Manchester, Manchester, UK



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

sought to adjust for its impact (see some notable exceptions in Levitt, 1998; Mohler et al., 2019; Pudney et al., 2000).

Such paucity of crime studies exploring the impact of measurement error could be attributed to the lack of simple adjustment methods tailored to the types of error mechanisms seen in police recorded crime data. For example, adjustment methods popularised in Epidemiology and Econometrics, such as SIMEX (Stefanski & Cook, 1995) or instrumental variables (Buzas & Stefanski, 1996), assume simplified forms of measurement error, the default being classical errors (i.e. completely random). The assumption of non-differentiability is also common in most measurement error adjustment methods, implying independence between the measurement error process and any other variables included in the substantive model of interest where the variable prone to measurement error is used. Whilst this assumption can be defended for the sake of parsimony and generalisability, there is good reason to anticipate that it will not hold when measurement error in police recorded crime is considered. There is abundant research pointing at victims' characteristics affecting their willingness to report an incident to the police (Baumer, 2002; Hart & Rennison, 2003), and it could also be expected that these associations are also present when shifting from the individual to the area level. For example, Xie and Baumer (2019) note differences in crime reporting based on the proportion of immigrants in the neighbourhood.

Other common adjustment methods, such as multiple imputation (Cole et al., 2006) or latent variable estimation (Wansbeek & Meijer, 2001), are not based on similarly restrictive assumptions of the error term, but instead require auxiliary data. This is typically in the form of a 'gold standard' (i.e. error free) measure for a subsample of cases, or 'repeated observations' (a series of additional imperfect measures tapping into the same construct). Victimization surveys have often been proposed as an alternative, more accurate, measure of crime. However, because of their small sample sizes relative to their wide coverage, estimates from crime surveys are generally only reliable at large regional levels (Rosenbaum & Lavrakas, 1995).

In this article, we present the package 'rcme' (Recounting Crime with Measurement Error) for R software (R Core Team, 2022); a user-friendly sensitivity analysis tool to anticipate and assess the impact of measurement error in any form of regression models relying on police recorded crime rates. This tool allows researchers to incorporate the knowledge that measurement error in crime rates can be expected to be systematic and

multiplicative. Beyond this, it also enables researchers to relax the assumption of non-differentiability (whereby errors are not allowed to be associated with any other variable included in the substantive model of interest), providing a more comprehensive picture of the range of ways that measurement error may impact the accuracy and precision of empirical analyses. To do this, 'rcme' adjusts police recorded crime rates directly via simulation of the (user-defined) expected measurement error mechanisms. These simulated crime rates are then fed into the substantive model of interest to produce a range of error-adjusted estimates. This approach is unique both in terms of its simplicity and generalisability, and can be deployed across different crime types, area levels, and empirical questions of interest.

Measurement error mechanisms present in police recorded crime rates

For an effective adjustment strategy, it is key to describe the measurement error mechanisms present in police recorded crime data accurately. However, a degree of parsimony is also required if the measurement error model is to be generalisable. To meet the trade-off between internal and external validity, we focus on measurement error mechanisms that can be expected to: (i) be present between and within police forces, across jurisdictions and time; and (ii) lead to severe biasing effects on key estimates derived from regression models. In consonance with these criteria, we have defined the error term affecting police recorded crime rates as: multiplicative, systematic, and differential.¹

Multiplicative errors

As demonstrated by Gibson and Kim (2008) and Pina-Sánchez et al. (2022), the measurement error affecting police recorded crime is likely to be multiplicative in

¹ Gibson and Kim (2008) and Pina-Sánchez et al. (2022) also identify a random component to the errors in recorded crime data across areas. These random errors might be reflective of, for example, inconsistent policing priorities and recording practices between police forces (Burrows et al., 2000; Maltz & Targonski, 2002; von Hofer, 2000) as well as technical failures (Whelan, 2021) and external political pressures (Boivin & Cordeau, 2011; Eterno et al., 2016; Maltz, 2006). However, random errors cannot be simply adjusted through simulations without additional assumptions about the relationship between the prevalence of the measurement error and its biasing impact in the estimate of interest (see Stefanski & Cook, 1995), and generating reliable estimates of the magnitude of these errors is far from straightforward. Therefore, to keep our approach as parsimonious as possible, we assume that the average recording rate could only vary when associated with the focal variable investigated as either a cause or consequence of crime (i.e. in the presence of differential errors). Recent work by Brunton-Smith et al. (2023b) finds that random errors in police data account for less than 10% of the total variation across areas, suggesting this simplification is unlikely to materially impact conclusions drawn using rcme in most applications.

form. Multiplicative errors are typically present in count data such as self-reported spells of unemployment (Pina-Sánchez et al., 2016), duration data such as recalled dates of developmental milestones (Pickles et al., 1996; Skinner & Humphreys, 1999), or for any other variables bound from zero to infinity that tend to be right-skewed such as personal income (Glewwe, 2007). In these instances, the measurement error term, U , is seen as proportional—rather than independent—to the unobserved true value, X . This is reflected by representing the observed variable, X^* , as,

$$X^* = X \cdot U \tag{1}$$

This multiplicative model provides a better fit to crime data than the standard additive model ($X^* = X + U$).² The intuition is that we should expect larger errors in areas where crime rates are higher, and smaller errors where crime rates are lower. We can see how that is the case if we consider: (i) each crime as a different event with a given probability of being counted accurately; (ii) as a result of under-reported and under-detected crimes, the probability of false negatives in recording is higher than that of false positives; and (iii) police recorded crime rates are nothing more than the count of crimes recorded by the police for a given period and location, divided by population.

Systematic errors

Perhaps the most noticeable and pervasive form of measurement error present in police statistics is the systematic under-estimation of the total number of crimes. This under-estimation occurs for three main reasons. Firstly, a decision by victims not to report some crimes to the police. The Crime Survey for England and Wales (CSEW) estimates that only around 41% of crimes are typically reported, with reasons for non-reporting including a perceived lack of offence severity, a belief that the police will be ineffective at responding, and fear of reprisals (Office for National Statistics, 2020). Of course, this overall figure masks considerable variability between offence types with as many as 93% of vehicle thefts reported, compared with around one third of criminal damage offences and 15% of frauds. Second, are those instances where the police fail to record a crime. The majority of these are instances where a reported incident is not recorded, with a recent audit of police recording practice estimating that as many as 800,000 crimes reported to the police go unrecorded each year in the UK (Her Majesty

Inspectorate of Constabulary, 2014). There is also a comparatively small number of police detected incidents that remain unrecorded and absent from official records, with the same HMIC audit suggesting around 3–7% of incidents are incorrectly ‘no-crimed’ each year. Finally, police statistics also fail to record those instances where the victim is unaware of its condition or there is no direct victim of crime (e.g. tax fraud, drug crime) and the crime fails to be detected by the police (Skogan, 1977).

Differential errors

Prior research has tended to assume that measurement errors affecting crime data are non-differential (Gibson & Kim, 2008). For the simple case of measurement error affecting a single explanatory variable in a model of with response variable, Y_i , this can be expressed as, $E(Y_i|X_i^*, X_i) = E(Y_i|X_i)$.³ However, depending on the research question of interest, this assumption may not hold. For example, using the International Crime Victimization Survey, Estienne and Morabito (2016) found that economic inequality is negatively associated with reporting crime to the police.

In Eq. 2, we show how differential errors can be reflected in our measurement error model - again, assuming that the error-prone variables is used as an explanatory variable. We do so by disentangling the measurement error term in two parts: (i) the systematic error R (i.e. the average recording rate); and (ii) the association between the measurement error term and the response variable of the substantive model of interest, Y_i , with D representing the strength of that association,

$$U_i = R + DY_i \tag{2}$$

Lastly, we constrain the possible values of U_i to a zero to one interval. This is equivalent to considering that none of the error mechanisms will lead to overestimates of crime in police recorded data, nor will they push the recording rate to non-sensical negative values. To ensure that such hard constraints can be held regardless of the dispersion of recording rates across areas and strength of differential errors, we estimate Eq. 2 with a binomial link function.

‘rcme’: recounting crime with measurement error

The ‘rcme’ sensitivity tool is designed to explore the robustness of estimates obtained from regression models where police recorded crime rates are used. This is achieved by the following four steps:

² This was tested for property crimes measured at police force area level in England and Wales and homicides measured at the State level in the US by Pina-Sánchez et al. (2022); and for car theft, theft from car, burglary, attempted burglary, robbery, and assault and threat cases, measured at the country level in the US by Gibson and Kim (2008).

³ More generally, for any kind of model including any number of explanatory variables, non-differentiality holds true when, $\sigma_{U,\varepsilon} = 0$, where ε represents the substantive model’s error term, and $\sigma_{U,\varepsilon}$ represents the covariance between the measurement and the substantive model’s error terms.

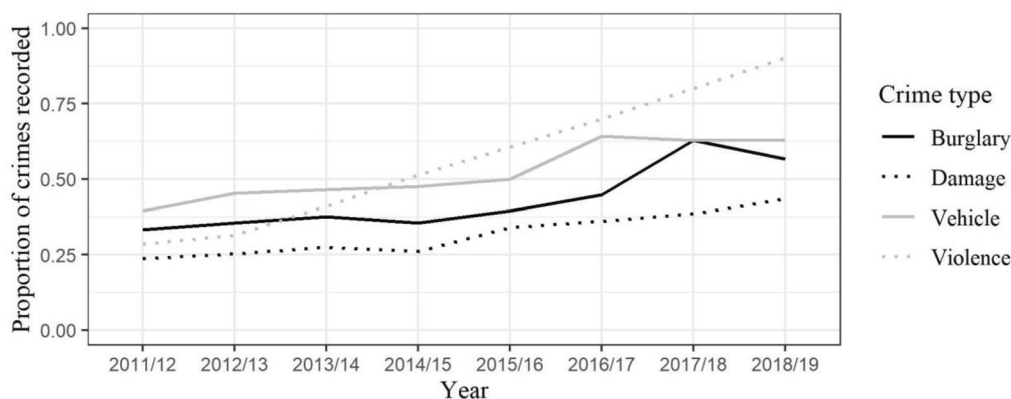


Fig. 1 Recording crime rates in England and Wales from 2011/12 to 2018/19

- i) Estimate the substantive model of interest using police recorded crime rates, and retrieve the key estimate of interest, which we call the ‘naïve estimate’.
- ii) Simulate the likely measurement error term, U , by identifying a range of plausible values representing the expected proportion of experienced crimes that were recorded by the police, R , and another range of plausible values representing the association between the recorded rate and the key variable of interest considered as either the cause or effect of crime rates, D .
- iii) Calculate a series of adjusted police recorded crime rates obtained after dividing the original police recorded crime rates by the range of values implied by the measurement error term as derived from step 2.
- iv) Re-estimate the substantive model using each of the adjusted crime rates derived from step 3, and report the range of adjusted values for the estimate of interest.

‘rcme’ has been implemented within the R programme language (R Core Team, 2022) and can be downloaded and installed directly from Github (`devtools::install_github("RecountingCrime/rcme")`). The package provides a series of functions to facilitate the calculation of adjusted crime rates based on the choice of a range of values for recording rates and differential errors,⁴ and to visualise the adjusted estimates of interest.

Selecting measurement error values

The selection of plausible values for the expected errors present in police recorded crime rates is essential for ensuring the sensitivity analysis is appropriately calibrated. Ideally, these should be based on previous studies in the literature, auxiliary data, or derived from expert opinions (Gosling, 2018). Most often, however, users will only be able to provide an educated guess, which is why we suggest a series of rules of thumb aiming to simplify this process.

First, when deciding on the average recording rate, R , it should be remembered that this can vary widely across crime types (Tarling & Morris, 2010), countries (Estienne & Morabito, 2016), and over time (Xie, 2014). Hence, it should be considered how well estimates from the literature, experts, or auxiliary datasets, can be ‘transported’ to the study in question where the sensitivity analysis is to be employed, and use wider or narrower ranges of recording rates accordingly.

Figure 1 provides an overview of the variability of recording rates across crime types and time for the case of England and Wales. These rates are derived as the ratio between the number of crimes recorded by the police and those estimated in the Crime Survey for England and Wales. We can see the expected differences across crime types, with criminal damage being less commonly recorded, but also some substantial changes across time.

The selection of plausible values for differential errors is significantly more challenging. As for the case of systematic errors, the presence of differential errors could be expected to vary by crime type, time, and jurisdiction; but more importantly, they will vary widely depending on the focal variable and the set of controls included in the substantive model of interest. To reflect this uncertainty, we suggest considering a wide range of differential errors covering both positive and negative associations between recording rates and the focal variable.

⁴ The type of sensitivity analysis through simulations advocated here could also be undertaken outside R by following the considerations regarding the reconstructing of the measurement error term described in the previous section. See also Gallop and Weschle (2019), who provide a blueprint for the simulation of a wider range of non-random types of measurement error.

‘rcme’ requires users to specify the magnitude of the differential errors as a risk ratio. This reflects the multiplicative form of the error mechanism and has the effect of increasing or decreasing the magnitude of the average recording rate proportionally to the identified predictor variable. For example, a risk ratio of 1.1 would increase the recording rate across areas by 10% for a one-unit increase in the focal variable.⁵ By default ‘rcme’ uses a range 0.9 to 1.1 risk ratios, but we encourage users to explore the literature in search for evidence regarding associations between the focal variable and recording or reporting rates. Even if the evidence available does not match the specific settings of their study (i.e. even if findings from the literature are not perfectly transportable as they are based on different crime types or area levels), this information could still be used to place more emphasis on the possible direction of any hypothetical differential error mechanism.

Application: exploring the robustness of the effect of collective efficacy on crime

We illustrate the value of sensitivity analysis using ‘rcme’ with an examination of the crime reducing effect of collective efficacy. Collective efficacy is a widely studied concept in Criminology, understood as the combination of ties between local residents with a willingness to intervene to prevent crimes (Mazerolle et al., 2010; Sampson & Wikström, 2008; Sampson et al., 1997). Here, we examine its effect on criminal damage rates (per 1000 residents) in 2011, across Middle layer Super Output Areas (MSOAs) in Greater London (N = 982). MSOAs contain

Table 1 Results for the naïve model (unadjusted) exploring the effect of collective efficacy on criminal damage rates

Predictor	Estimate	95% CI	p-value
(Intercept)	2.71	(2.62, 2.78)	< 0.001
Collective efficacy	- 0.26	(- 0.37, - 0.16)	< 0.001
% Unemployed	0.33	(0.20, 0.45)	< 0.001
% White British	0.29	(0.18, 0.41)	< 0.001
Median age	- 0.26	(- 0.37, - 0.15)	< 0.001
R ²	0.17		
N	982		

any young people here are causing trouble, people will tell them off’, ‘people call the police if someone is acting suspiciously’ and ‘if I sensed trouble, I could get help from people who live here’ (CFI=0.905, TLI=0.841, RMSEA=0.139). Our measure of collective efficacy in MSOAs is the area weighted average of individual latent scores. We also include three control variables that are commonly considered in the literature: the percentage of unemployed residents, percentage of white British residents, and the median age. These controls are also extracted from the 2011 Census.

In the first step in our assessment of the robustness of the effect of collective efficacy to measurement error, we estimate a linear model using ordinary least squares, with all the explanatory variables standardised. We call this the ‘naïve’ model:

```
lm(criminal_damage ~ collective_efficacy + unemployment +
    white_british + median_age,
    data = example)
```

between 5000 and 15,000 residents, or 2000–6000 households, and are often used as a measure of neighbourhoods in England and Wales.

Crime counts are accessed from data.police.uk and converted to rates using residential population counts taken from the 2011 Census. Collective efficacy is derived from the Metropolitan Police Service Public Attitudes Survey 2011–2013. Specifically, we use a confirmatory factor analysis model including the following items: ‘people in this area can be trusted’, ‘people act with courtesy to each other’, ‘people take pride in their environment’, ‘if

Table 1 reports the results from this model. We observe the expected negative association identified in the literature. Specifically, in MSOAs where collective efficacy is one standard deviation higher, criminal damage rates are 0.26 lower than in the average London MSOA. This is a statistically significant association, however the effect size is relatively modest if we take into consideration that the average criminal damage rate in our sample is 2.71 (per 1000 residents).

The second step of ‘rcme’ is crucial. It is here that the user defines the range of plausible values of systematic and differential errors. In Fig. 1, we saw how recording rates for criminal damage in England and Wales varied from roughly 25% in 2011/12 to 45% in 2018/19. We use the same range of values for our sensitivity analysis, noting that this is wide enough to be likely capturing

⁵ It is generally advisable to standardise continuous variables suspected to be prone to differential errors prior to their inclusion in the model to facilitate comparisons, so a one-unit increase in the focal variable always corresponds to a one standard deviation change, regardless of the unit of measurement used.

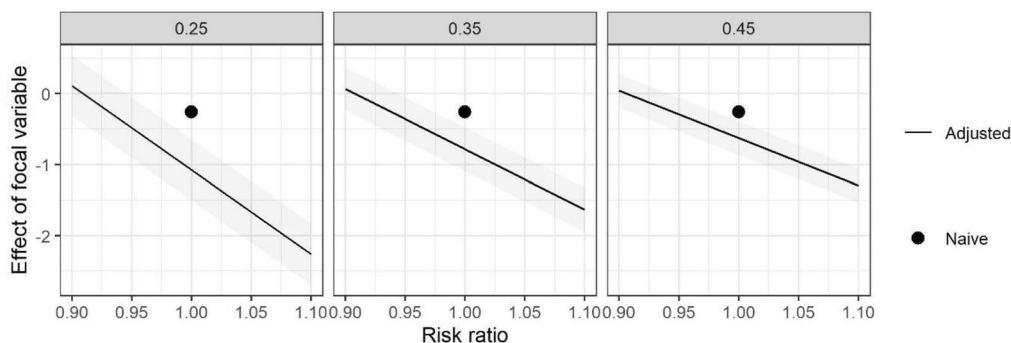


Fig. 2 Possible values for the ‘true’ effect of collective efficacy on criminal damage in the presence of different ranges of systematic (0.25 to 0.45 recording rates) and differential (0.9 to 1.1 risk ratios) errors affecting police recorded crime rates

the true but unobserved recording rate in our study. We also expect that with our study using data from 2011 it is more likely that the true recording rate is closer to the lower end of that interval.

We are less certain about the magnitude of the differential error. However, there are theoretical grounds to expect that crime reporting is positively associated with collective efficacy, since social trust and willingness to collaborate with the police are specific items used in

these default values in context, for the case of an average recording rate of 45%, a risk ratio of 1.1 means that recording rates in MSOAs with collective efficacy one standard deviation higher than the average are 49.5% (e.g. 10% higher than the average recording rate).

To pass on the range of values considered and calculate the adjusted crime rates (step 3) and adjusted estimates (step 4) we can use the command `rcme_out()`.

```
me <- rcme_out(formula = "criminal_damage ~ collective_efficacy +
  unemployment + white_british + median_age",
  data = example,
  focal_variable = "collective_efficacy",
  R = c(0.25, 0.35, 0.45),
  D = c(0.9, 0.95, 1, 1.05, 1.10))
```

the measurement of collective efficacy. This intuition has been empirically corroborated in the literature. For example, at the individual level, Hart and Colavito (2011) reported a positive association between collective efficacy and crime reporting amongst US college students, while at the Community Safety Partnership level, Brunton-Smith et al. (2023a) report a similar positive association between collective efficacy and crime recording. Yet, the precise magnitude of this positive association is far from certain, so we choose to retain the default ‘rcme’ values of 0.9–1.1 risk ratios, with the expectation that if present, differential errors in our study are more likely going to be positive (i.e. recording rates will be higher in areas with more collective efficacy) than negative. To put

To visualise the range of adjusted estimates obtained for the combination of systematic and differential errors considered, we can use the command `rcme_sim_plot()`.

```
rcme_sim_plot(me)
```

The resulting visualisation is shown in Fig. 2, where the three panels show each of the recording rates considered, the horizontal axis covers the range of differential errors, the vertical axis indicates the estimated effect of collective efficacy on criminal damage under such measurement error conditions, and the black dot indicates the naïve estimate for comparison.

We can see that the range of possible estimates widens when recording rates are lower. In those instances, results are more sensitive to the value attributed to differential errors. However, we can also see that under most scenarios, the effect of collective efficacy appears to have been underestimated. Specifically, such downward bias is stronger in the scenarios that we noted as most likely, low recording rates and positive differential errors, suggesting that the crime prevention effect has been underestimated. Even if we consider that recording rates and collective efficacy are independent, for an average recording rate of 25% we would have underestimated the effect of collective efficacy on crime reduction by a factor of roughly four.

Discussion

For decades (Biderman & Reiss, 1967; Skogan, 1977), if not centuries (de Candolle, 1830[1987]), it has been known that official crime records do not provide an accurate representation of the true extent of crime; with a substantial proportion of crimes failing to reach the attention of the police, and standards of crime recording varying substantially within and between police forces. However, whilst the limitations of police data are widely acknowledged, empirical assessments of the impact of such large measurement errors remain rare. As a result, the veracity of the existing evidence is unclear. In other words, studies that rely on police recorded statistics might be severely biased (Pepper et al., 2010; Pina-Sánchez et al., 2022).

The problem of measurement error is not unique to crime data, and a number of different adjustment methods are regularly used in other disciplines including Epidemiology, Biostatistics, Economics and Psychology. However, many of the methods that have been proposed in those disciplines cannot be readily adopted to adjust for the types of measurement errors present in police recorded crime rates, which are affected by multiple, and to some extent, non-conventional, measurement error mechanisms (Gibson & Kim, 2008; Pina-Sánchez et al., 2022).

'rcme' is a sensitivity analysis tool designed to enable researchers to explore the extent to which causal estimates derived from regression models where police recorded crime rates are used can be biased as a result of the measurement error mechanisms typically seen in that data. 'rcme' works by simulating plausible ranges for the effects of the types of error seen in police recorded crime rates, with the specific values defined by the user. Importantly, since the focus of the adjustment is on the police data itself, rather than on the estimation method,

it can be applied to any modelling setting. That is, across all types of regression models, regardless of the response and explanatory variables considered, or where in the model police recorded crime rates are to be introduced.

By applying this sensitivity tool to a typical research question commonly explored in the crime literature, we have illustrated the extent to which the evidence base might be biased, but also that the magnitude of the problem is context dependent. Specifically, our example showed that a typical study exploring the role of collective efficacy in reducing criminal damage using a linear model will likely underestimate the true effect of collective efficacy substantially. Further examples considering different types of models, crime types, and concepts widely explored as either precursors or consequences of crime can be found in our project website (www.recoupingcrime.com/workshops/).

'rcme' has the potential to make assessments of the sensitivity of results to the presence of measurement error routine in the study of crime. However, as with all such tools, the usefulness of the results stems from the precision with which the presence of different types of measurement error can be anticipated. The effectiveness of the sensitivity assessment will be proportional to our knowledge of the underlying measurement error mechanisms. It is therefore crucial that future research sheds light on the forms and prevalence of measurement error present in police data.

Author contributions

Jose Pina-Sánchez conceptualise the sensitivity analysis tool and wrote the 'Introduction', 'Measurement error mechanisms present in police recorded crime rates', 'rcme: recounting crime with measurement error' and 'Discussion'. Ian Brunton-Smith undertook the analysis and writing of the 'Application: exploring the robustness of collective efficacy on crime'; he also structured the paper in its current format. David Buil-Gil accessed and clean the data and contributed to the literature review. Alexandru Cernat wrote the R package. All authors read and approved the final manuscript.

Funding

This work was supported by the Secondary Data Analysis Initiative of the Economic and Social Research Council (Grant Ref: ES/T015667/1).

Availability of data and materials

The data used in this study can be retrieved from the following websites: police.data.uk, crimesurvey.co.uk, ukdataservice.ac.uk, and ons.gov.uk, the R package to undertake the sensitivity analysis can be found here: <https://github.com/RecountingCrime/rcme>.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 6 January 2023 Accepted: 12 July 2023

Published online: 25 July 2023

References

- Baumer, E. P. (2002). Neighborhood disadvantage and police notification by victims of violence. *Criminology*, 40(3), 579–616.
- Biderman, A. D., & Reiss, A. J. (1967). On exploring the 'dark figure' of crime. *The Annals of the American Academy of Political & Social Science*, 374(1), 1–15.
- Boivin, R., & Cordeau, G. (2011). Measuring the impact of police discretion on official crime statistics: A research note. *Police Quarterly*, 14(2), 186–203.
- Boxer, P., Drawwe, G., & Caplan, J. M. (2020). Neighborhood violent crime and academic performance: A geospatial analysis. *American Journal of Community Psychology*, 65(3–4), 343–352.
- Brunton-Smith, I., Buil-Gil, D., Pina-Sánchez, J., Cernat, I., & Moretti, A. (2023a). Using synthetic crime data to understand patterns of police undercounting at the local level. In L. Huey & D. Buil-Gil (Eds.), *The Crime Data Handbook*. Bristol University Press.
- Brunton-Smith, I., Cernat, A., Pina-Sánchez, J., & Buil-Gil, D. (2023b). Estimating the reliability of crime data in geographic areas. *CrimRxiv*, 1–16.
- Burrows, J., Tarling, R., & Mackie, A. (2000). *Review of police force's crime recording practices*. Home Office Research Study.
- Buzas, J. S., & Stefanski, L. A. (1996). Instrumental variable estimation in generalized linear measurement error models. *Journal of the American Statistical Association*, 91(435), 999–1006.
- Cole, S. R., Chu, H., & Greenland, S. (2006). Multiple-imputation for measurement-error correction. *International Journal of Epidemiology*, 35(4), 1074–1081.
- Coleman, C., & Moynihan, J. (1996). *Understanding Crime Data: Haunted by the Dark Figure*. Open University Press.
- de Candolle, A. (1830). Considérations sur la statistique des délits. *Déviance Et Société*, 11(4), 352–355.
- Detotto, C., & Otranto, E. (2010). Does crime affect economic growth? *Kyklos*, 63(3), 330–345.
- Estienne, E., & Morabito, M. (2016). Understanding differences in crime reporting practices: A comparative approach. *International Journal of Comparative and Applied Criminal Justice*, 40(2), 123–143.
- Eterno, J. A., Verma, A., & Silverman, E. B. (2016). Police manipulations of crime reporting: Insiders' revelations. *Justice Quarterly*, 33(5), 811–835.
- Gallop, M., & Weschle, A. (2019). Assessing the impact of non-random measurement error on inference: A sensitivity analysis approach. *Political Science Research Methods*, 7(2), 367–384.
- Gibson, J., & Kim, B. (2008). The effect of reporting errors on the cross-country relationship between inequality and crime. *Journal of Development Economics*, 87(2), 247–254.
- Glewwe, P. (2007). Measurement error bias in estimates of income and income growth among the poor: Analytical results and a correction formula. *Economic Development and Cultural Change*, 56, 163–189.
- Gosling, J. P. (2018). SHELF: The Sheffield Elicitation Framework. In L. Dias, A. Morton, & J. Quigley (Eds.), *Elicitation. International Series in Operations Research & Management Science* (pp. 61–93). Springer.
- Hart, T. C., & Colavito, V. (2011). College student victims and reporting crime to the police: The influence of collective efficacy. *Criminology, Criminal Justice, Law & Society*, 12(3), 1–19.
- Hart, T. C., & Rennison, C. (2003). *Reporting crime to the police, 1992–2000*. Bureau of Justice Statistics. <https://static.prisonpolicy.org/scans/bjs/rcp00.pdf>
- Her Majesty Inspectorate of Constabulary. (2014). *Crime-recording: making the victim count—The final report of an inspection of crime data integrity in police forces in England and Wales*. Justice Inspectorates. <https://www.justiceinspectorates.gov.uk/hmicfrs/wp-content/uploads/crime-recording-making-the-victim-count.pdf>
- Janke, K., Propper, C., & Shields, M. A. (2016). Assaults, murders and walkers: The impact of violent crime on physical activity. *Journal of Health Economics*, 47, 34–49.
- Keels, M., Duncan, G. J., DeLuca, S., Mendenhall, R., & Rosenbaum, J. (2005). Fifteen years later: Can residential mobility programs provide a long-term escape from neighborhood segregation, crime, and poverty. *Demography*, 42(1), 51–73.
- Levitt, S. D. (1998). Why do increased arrest rates appear to reduce crime: Deterrence, incapacitation, or measurement error? *Economic Inquiry*, 36(3), 353–372.
- Lohr, S. H. (2019). *Measuring crime: Behind the statistics*. CRC Press.
- Maltz, M. D. (2006). *Analysis of missingness in UCR crime data*. Criminal Justice Research Center, Ohio State University.
- Maltz, M. D., & Targonski, J. (2002). A note on the use of country-level UCR data. *Journal of Quantitative Criminology*, 18(3), 297–318.
- Mazerolle, L., Wickes, R., & McBroom, J. (2010). Community variations in violence: The role of social ties and collective efficacy in comparative context. *Journal of Research in Crime and Delinquency*, 47, 3–30.
- Mohler, G., Brantingham, P. J., Carter, J., & Short, M. B. (2019). Reducing bias in estimates for the law of crime concentration. *Journal of Quantitative Criminology*, 35, 747–765.
- Office for National Statistics. (2020). *Crime in England and Wales: Annual Trend and Demographic Tables*. <https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/datasets/crimeinenglandandwalesannualtrendanddemographictables/current>
- Pepper, J., Petrie, C., & Sullivan, S. (2010). Measurement error in criminal justice data. In A. R. Piquero & D. Weisburd (Eds.), *Handbook of Quantitative Criminology* (pp. 353–374). Springer.
- Pickles, A., Pickering, K., & Taylor, C. (1996). Reconciling recalled dates of developmental milestones, events and transitions: A mixed generalized linear model with random mean and variance functions. *Journal of the Royal Statistical Society: Series A (statistics in Society)*, 159(2), 225–234.
- Pina Sánchez, J. (2016). Adjustment of recall errors in duration data using SIMEX. *Metodološki Zvezki-Advances in Methodology & Statistics*, 13(1), 27–58.
- Pina-Sánchez, J., Buil-Gil, D., Brunton-Smith, I., and Cernat, A. (2022). The impact of measurement error in models using police recorded crime rates. *Journal of Quantitative Criminology*.
- Pudney, S., Deadman, D., & Pyle, D. (2000). The relationship between crime, punishment and economic conditions: Is reliable inference possible when crimes are under-recorded? *Journal of the Royal Statistical Society: Series A (statistics in Society)*, 163(1), 81–97.
- R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rosenbaum, D. P., & Lavrakas, P. J. (1995). Self-reports about place: The application of survey and interview methods to the study of small areas. In J. E. Eck & D. Weisburd (Eds.), *Crime and place. Crime prevention studies* (pp. 285–314). Criminal Justice Press.
- Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, 277(5328), 918–924.
- Sampson, R. J., & Wikström, P.-O.H. (2008). The social order of violence in Chicago and Stockholm neighbourhoods: A comparative inquiry. In S. N. Kalyva, I. Shapiro, & T. Masoud (Eds.), *Order, conflict and violence* (pp. 97–119). Cambridge University Press.
- Skinner, C., & Humphreys, K. (1999). Weibull regression for lifetimes measured with error. *Lifetime Data Analysis*, 5, 23–37.
- Skogan, W. G. (1977). Dimensions of the dark figure of unreported crime. *Crime & Delinquency*, 23(1), 41–50.
- Stefanski, L. A., & Cook, J. R. (1995). Simulation-extrapolation: The measurement error jackknife. *Journal of the American Statistical Association*, 90(432), 1247–1256.
- Tarling, R., & Morris, K. (2010). Reporting crime to the police. *The British Journal of Criminology*, 50(3), 474–490.
- von Hofer, H. (2000). Crime statistics as constructs: The case of Swedish rape statistics. *European Journal on Criminal Policy & Research*, 8, 77–89.
- Wansbeek, T., & Meijer, E. (2001). Measurement error and latent variables. *A Companion to Theoretical Econometrics*, 162–179.
- Whelan, Z. (2021). Technical blunder sees 150,000 arrest records 'accidentally wiped' from UK-wide police database. Manchester Evening News. <https://www.manchestereveningnews.co.uk/news/uk-news/technical-blunder-sees-150000-arrest-19630870>
- Xie, M. (2014). Area differences and time trends in crime reporting: Comparing New York with other metropolitan areas. *Justice Quarterly*, 31(1), 43–73.
- Xie, M., & Baumer, E. P. (2019). Neighborhood immigrant concentration and violent crime reporting to the police: A multilevel analysis of data from the National Crime Victimization Survey. *Criminology*, 57(2), 237–267.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.