

Accessing and analysing CSEW data for low-level geographies

A case study about the dark figure of crime

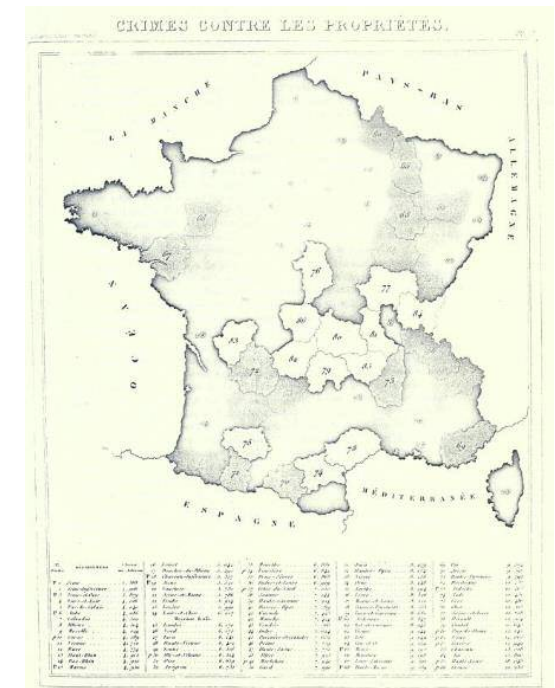
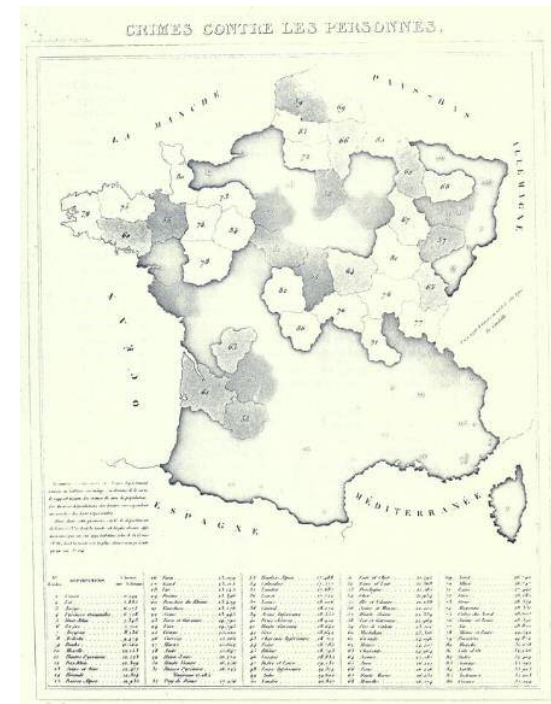
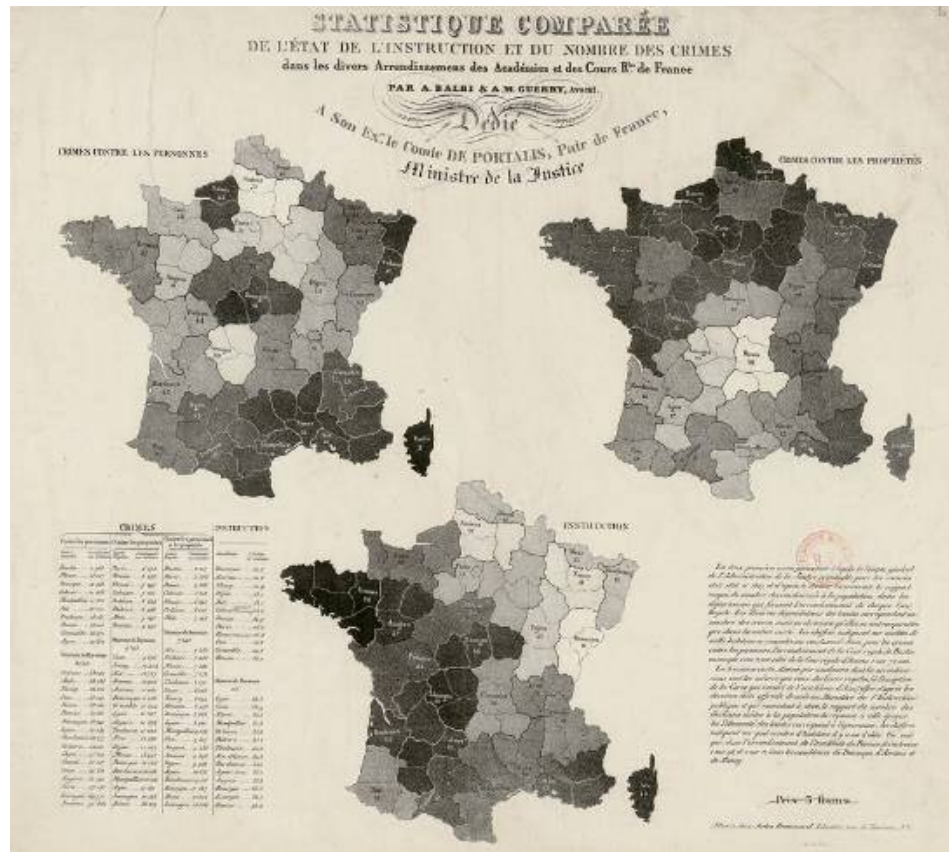
David Buil-Gil

 @davidbuil

david.builgil@manchester.ac.uk

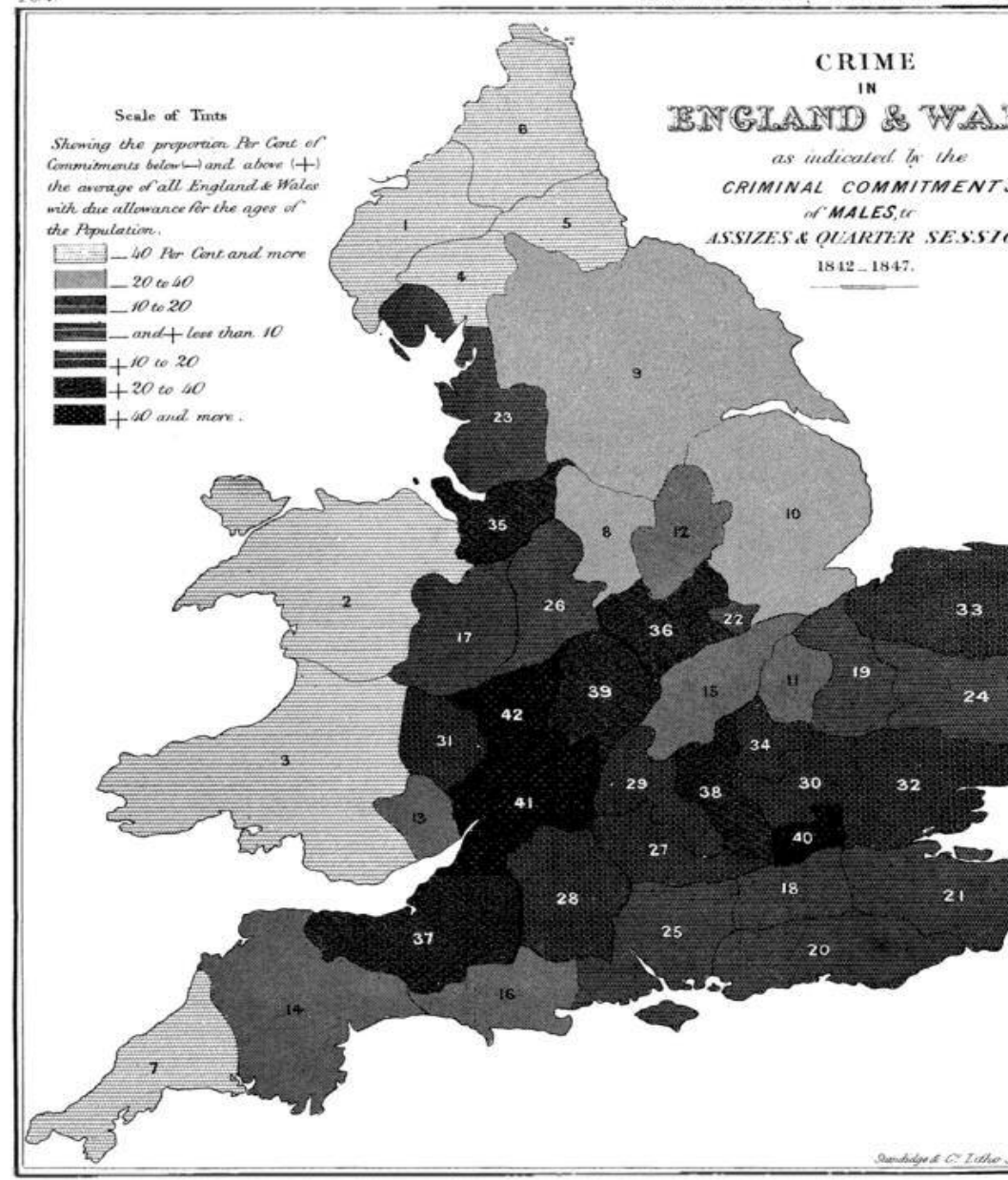
Overview

- 1.- Crime and space
- 2.- Police data is not enough
- 3.- The limitations of survey data for crime mapping
- 4.- Producing small area estimates of the dark figure of crime
- 5.- Conclusions



Guerry (1833)

*de Candolle (1837a, 1837b)



Fletcher (1849)

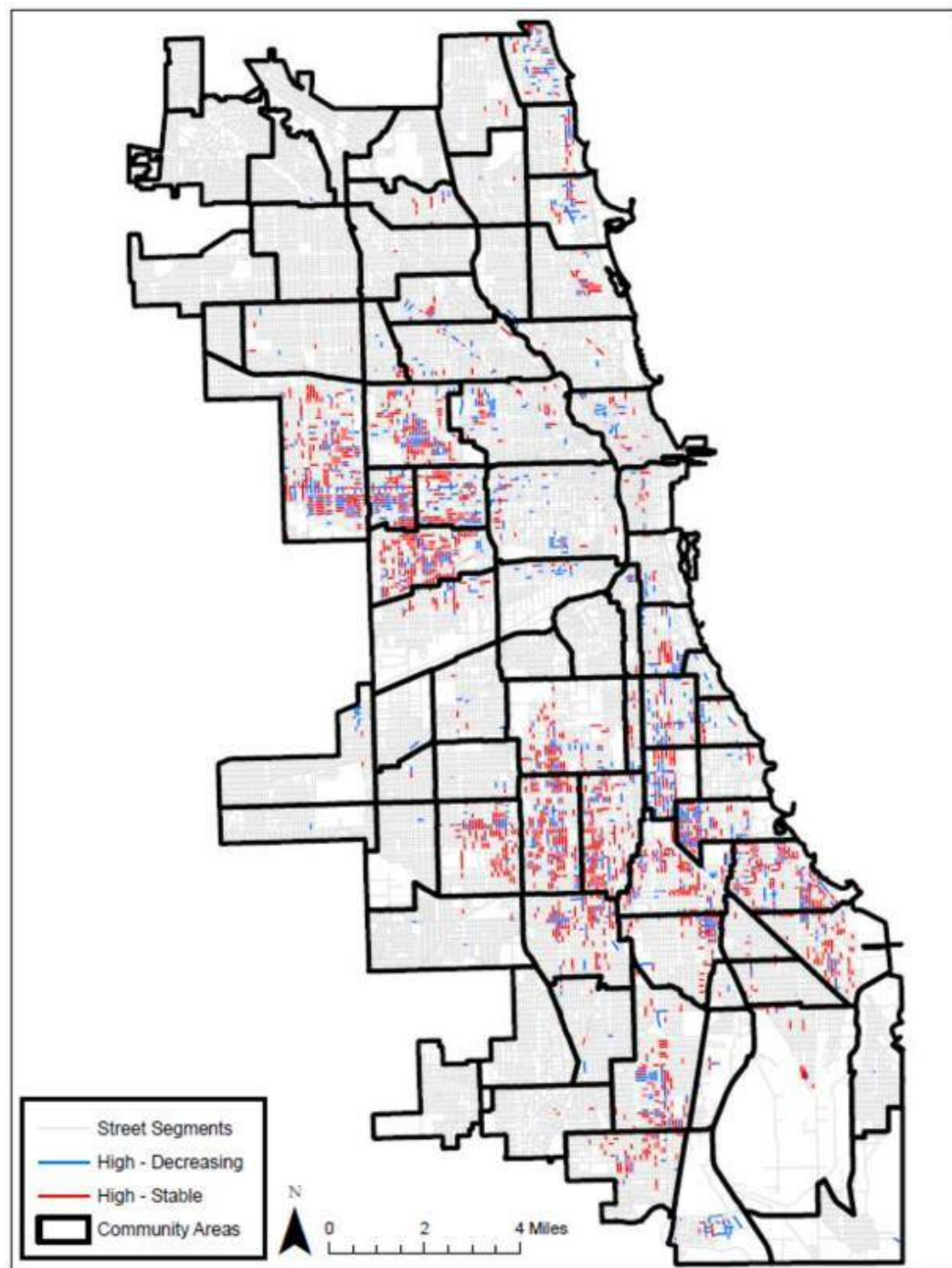
Crime concentration in micro places

- **Pierce et al. (1988)**: 2.6% of addresses in the city of Boston produced 50% of all crime calls to police services
- **Sherman et al. (1989)**: 3.5% of all addresses produced 50% of the annual crime calls to the police in Minneapolis
- **Weisburd et al. (2004)**: looked into the geographical distribution of crime statistics in Seattle over short periods of time between 1989 and 2002, and concluded that micro places with high concentrations of crime are stable over time

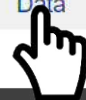
Crime concentration in micro places



There is a law of crime concentration in places!
“for a defined measure of crime at a specific **microgeographic** unit, the concentration of crime will fall within a narrow bandwidth of percentages for a defined cumulative proportion of crime” (Weisburd, 2015:138)



Crime concentration in
Chicago
Schnell (2017)



Welcome to data.police.uk

This is the site for open data about crime and policing in England, Wales and Northern Ireland.

You can download [street-level crime](#), [outcome](#), and [stop and search data](#) in clear and simple CSV format and explore the [API](#) containing detailed crime data and information about individual police forces and neighbourhood teams.

You can also download data on police activity, and a range of data collected under the [police annual data requirement \(ADR\)](#) including arrests and 101 call handling.

All the data on this site is made available under the [Open Government Licence v3.0](#).



DOWNLOADS

[Download Police.uk data in batches](#)



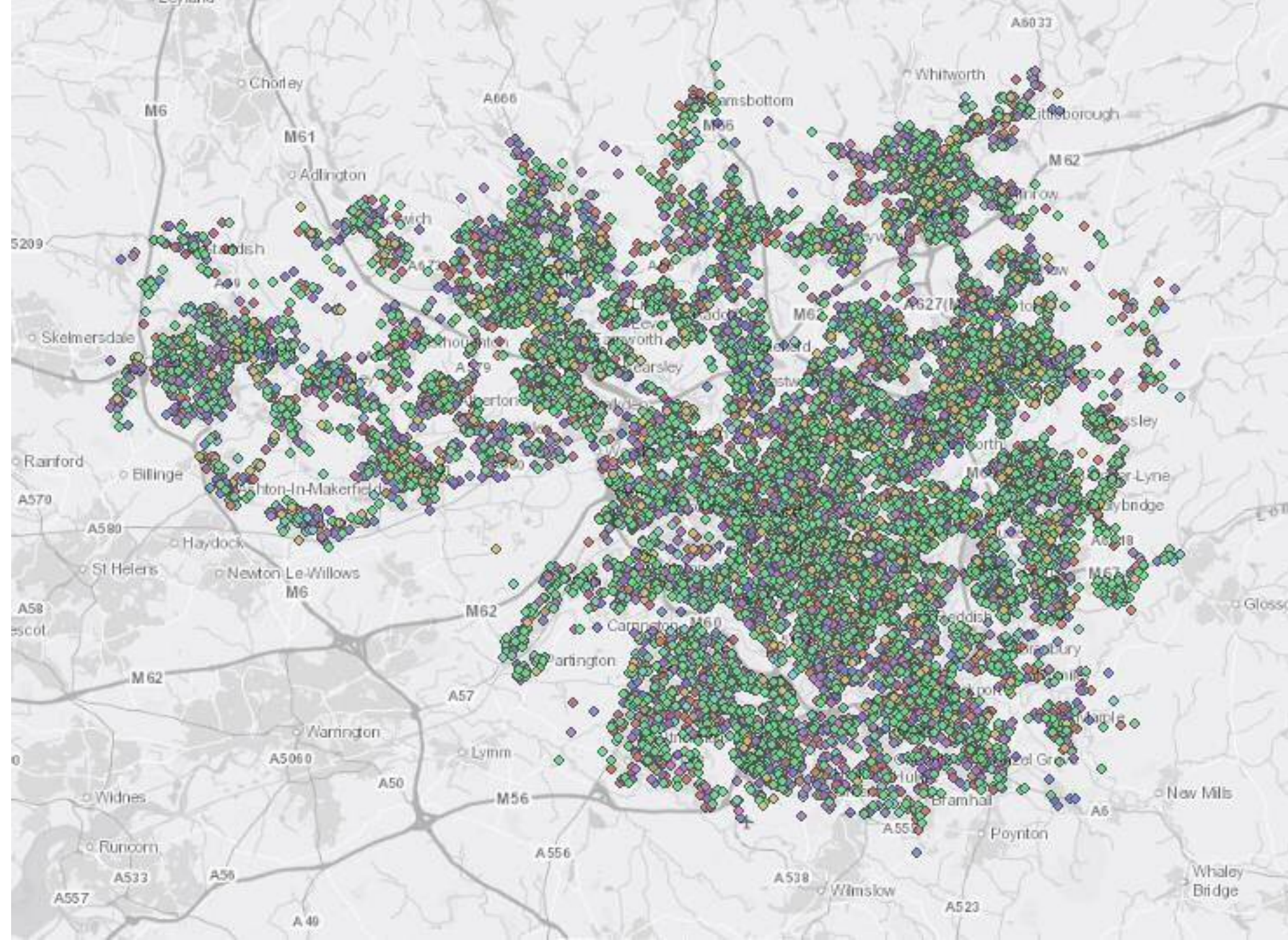
API DOCS

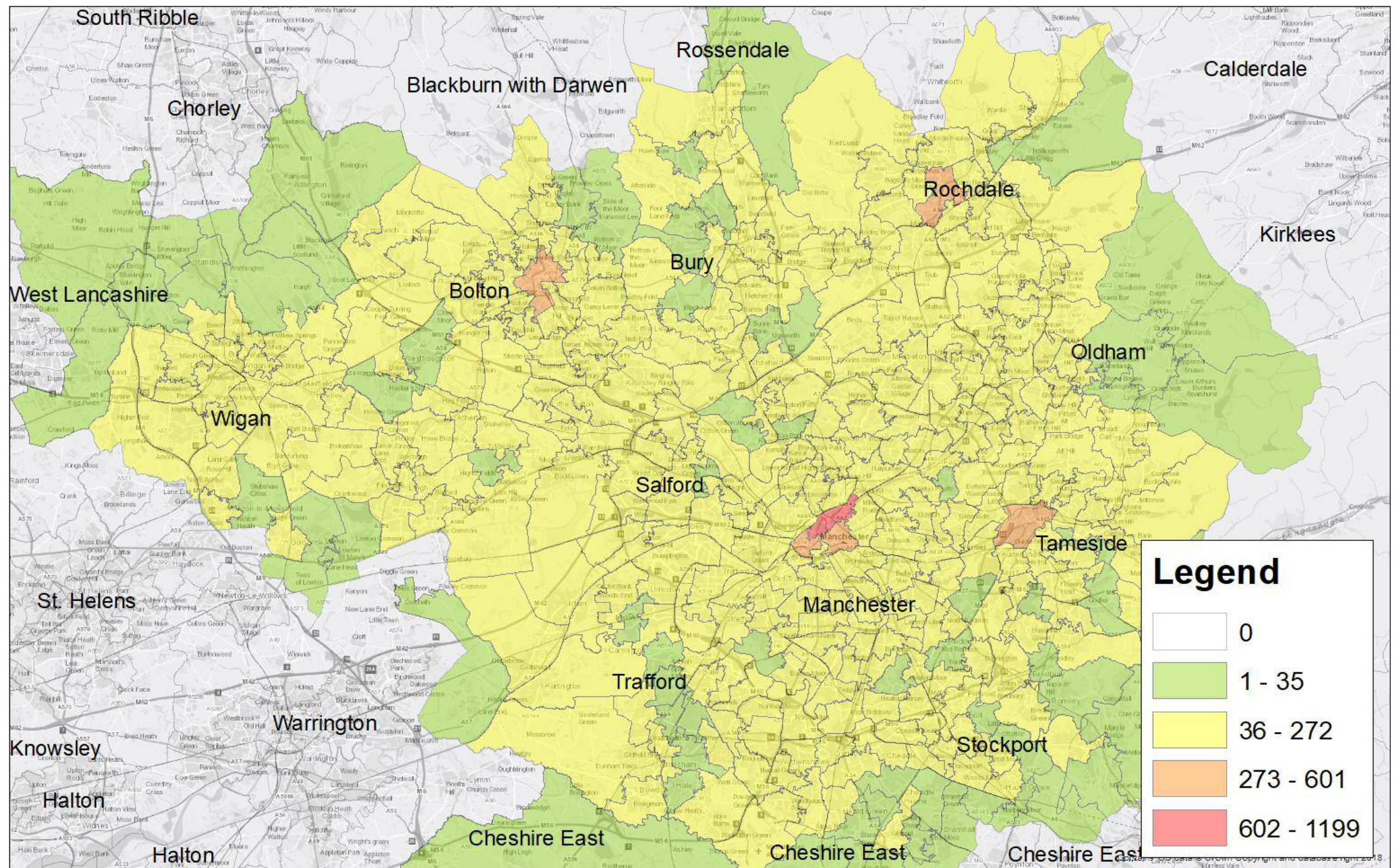
[Access Police.uk data via an API](#)



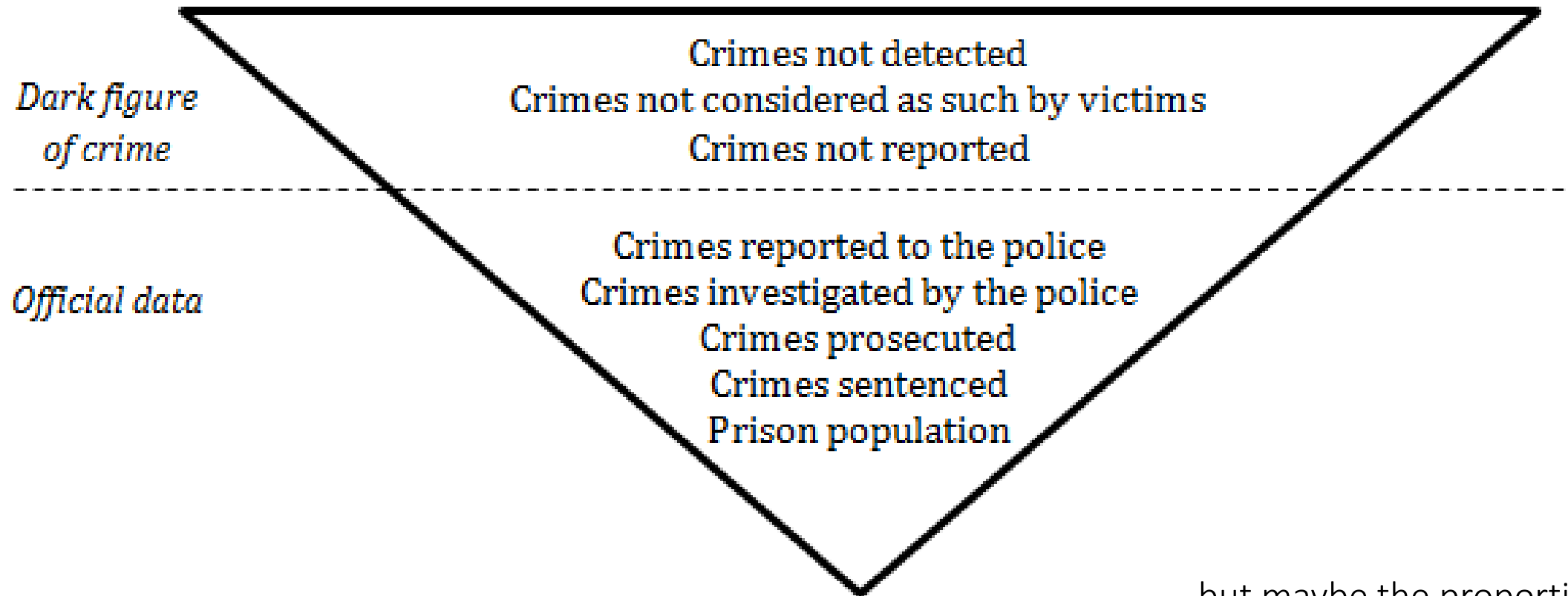
CHANGELOG

[See what's new](#), [see what's coming soon](#)





The 'funnel' of crime data



... but maybe the proportion of crimes unknown to police is the same across areas (MCAR)

The dark figure of crime

- Mawby (1979) compared police records, surveys and self-reported data in Sheffield - no important alterations across data sources
 - Bottoms and Wiles (1997): such results must not be overstated and cannot be overgeneralised
- Bottoms et al. (1987): similar study in seven areas in Sheffield - police data provide valid rates for comparisons across areas with similar housing types and within same police jurisdiction
- Goudriaan et al. (2004): cross-national survey data - victims are more likely to report crimes in countries where police are perceived to be competent
- Victims from suburban areas are less likely to report crimes to police than urban and rural citizens (Hart & Rennison, 2003; Langton et al., 2012)
- Residents from deprived neighbourhoods are less likely to report (Baumer, 2002; Berg et al., 2013; Goudriaan et al., 2006; Slocum et al., 2010)

Studies

Series

<http://doi.org/10.5255/UKDA-SN-8464-1>

Copy study DOI

Crime Survey for England and Wales 2017-2018

Page guide ?

Details

Documentation

Resources

Access data



Details

Title:	Crime Survey for England and Wales 2017-2018
Alternative title:	CSEW
Study number (SN):	8464
Access:	These data are safeguarded
Persistent identifier (DOI):	10.5255/UKDA-SN-8464-1
Series:	Crime Survey for England and Wales
Principal investigator(s):	Office for National Statistics

Sponsors and contributors



Citation and copyright



The citation for this study is:

Office for National Statistics. (2019). *Crime Survey for England and Wales 2017-2018*. [data collection]. UK Data Service. SN: 8464, <http://doi.org/10.5255/UKDA-SN-8464-1>

Select citation format:

APA



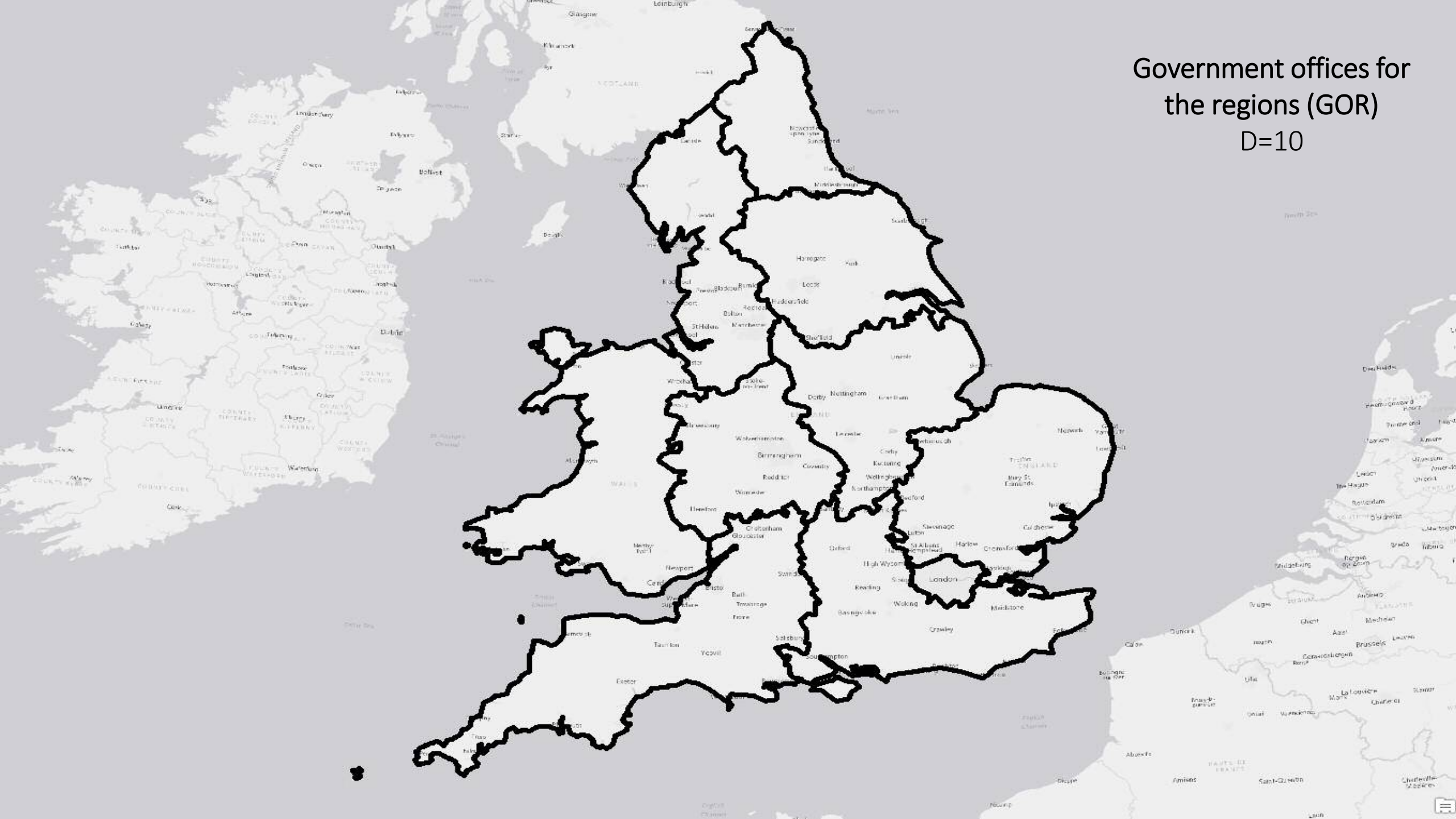
XML citation formats:

[CSL](#)

[EndNote](#)

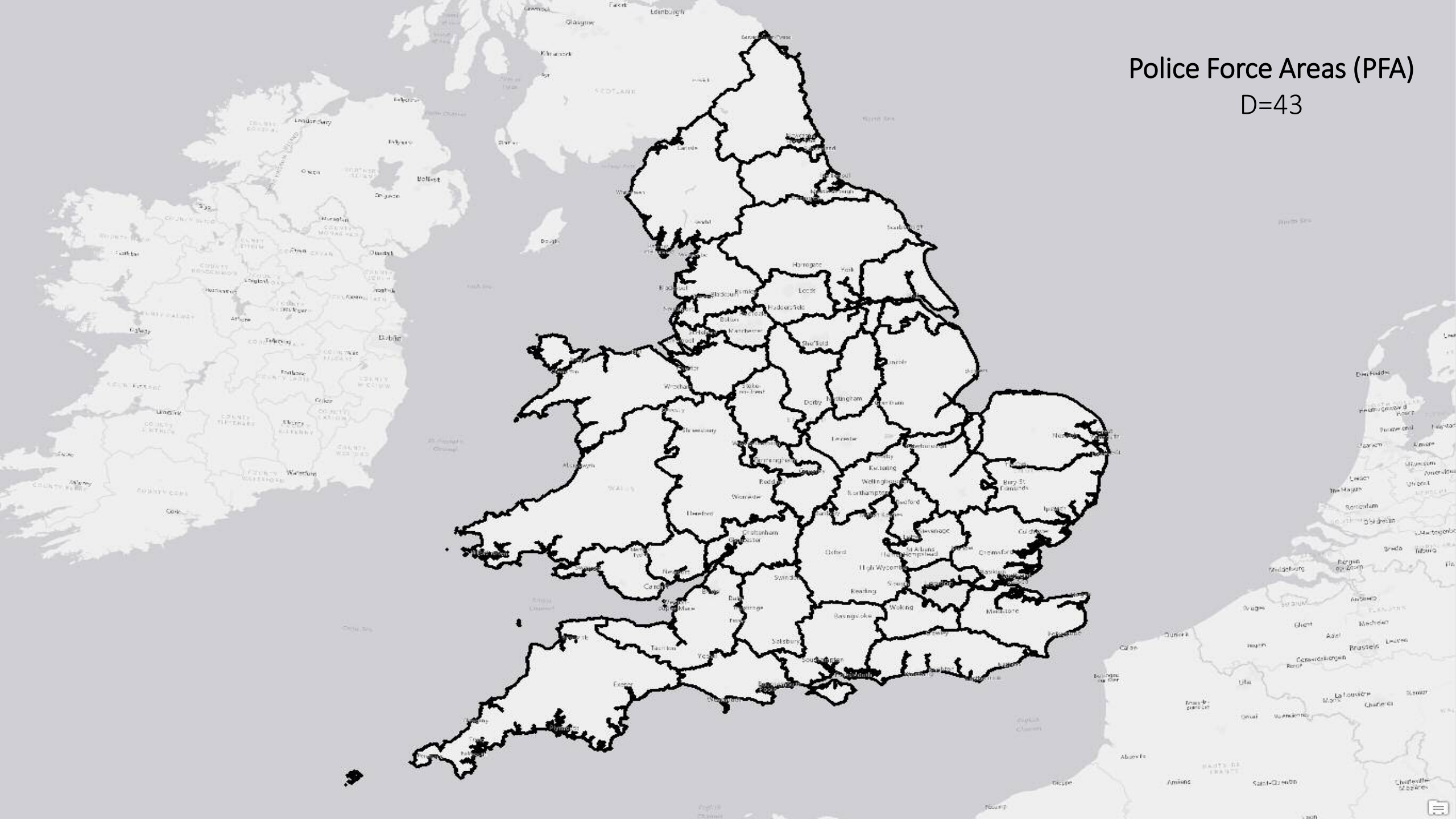
Government offices for
the regions (GOR)

D=10



Police Force Areas (PFA)

D=43



ukdataservice.ac.uk/use-data/secure-lab.aspx

Site search Accessibility FAQ Help Contact New users Sitemap

UK Data Service

About us Get data **Use data** Manage data Deposit data News and events

Home > Use data > Secure Lab

Secure Lab

"Enabling research through responsible use of detailed and sensitive data"

The UK Data Service is one of the few organisations who have the experience and expertise to build a world-leading secure access system.

What is the Secure Lab? Our collection Five Safes framework How it works

How do I apply?

SHARE

The UK Data Service Secure Lab provides secure access to data that are too detailed, sensitive or confidential to be made available under the standard [End User Licence](#) or [Special Licence](#). Our specialised staff apply statistical control techniques to ensure the delivery of safe statistical results.

Data accessed in this way cannot be downloaded. Once researchers and their projects

LOGIN / REGISTER

DATA CATALOGUE

GO

How to log in to the Secure Lab

HOW TO LOG IN TO THE SECURE LAB

View a video tutorial to show researchers who have been granted approval to use secure access level data how to log into their Secure Lab account.

QUICK ACCESS TO

About our data

Business microdata

Handbook Statistical Disclosure

[443 days]

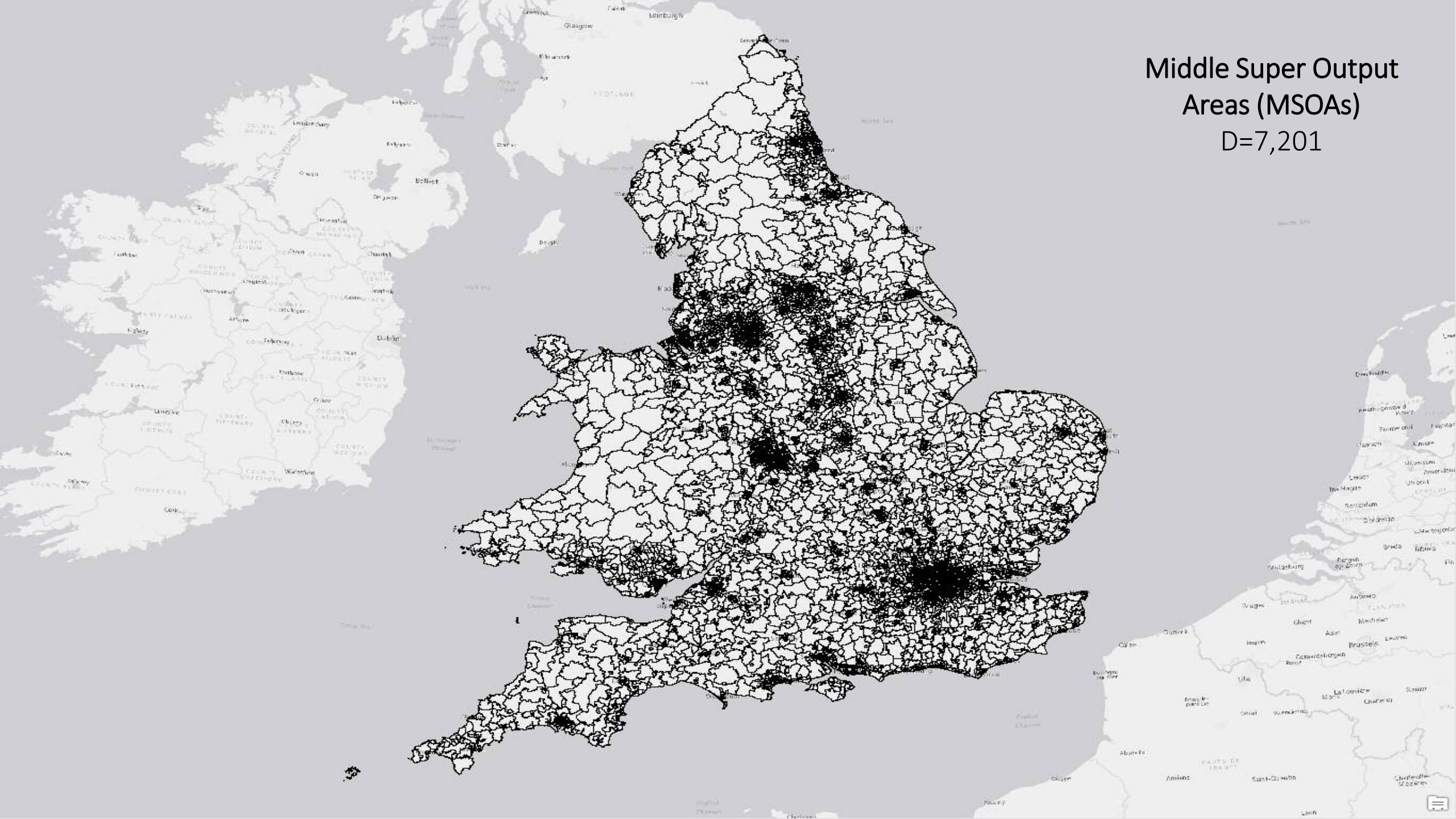
05/03/2017 -
22/05/2018

- (a) Apply for Accredited Researcher status
- (b) Go to London to take Safe User of Research Environments training
- (c) Submit a project proposal

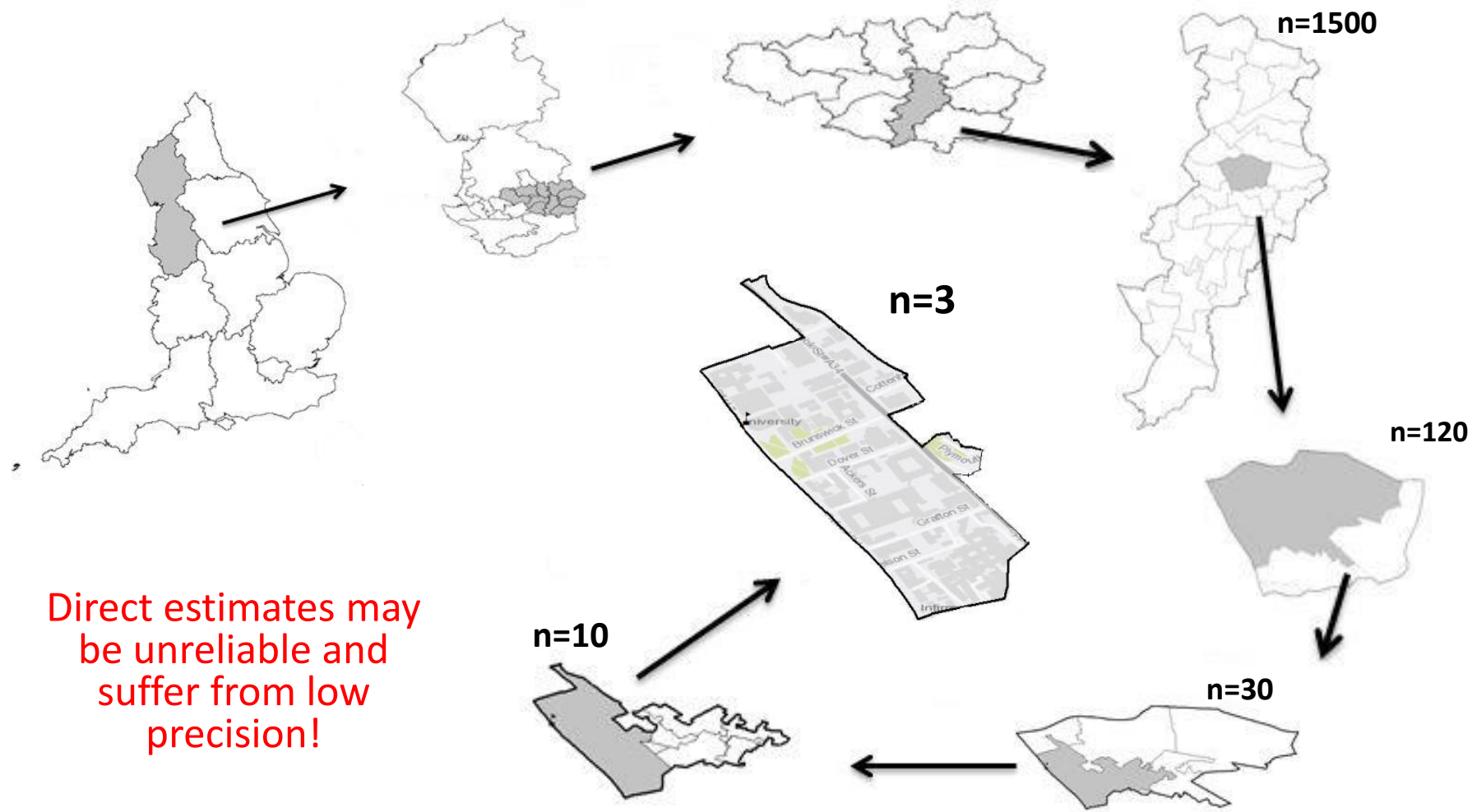
- (d) Complete a Secure User Agreement
- (e) Have your own office in your institution
- (f) Install Citrix Receiver
- (g) Send the R packages as zip files by email

For all
researchers
involved in the
project!

Middle Super Output Areas (MSOAs) D=7,201



Victimisation surveys: **Limitations**



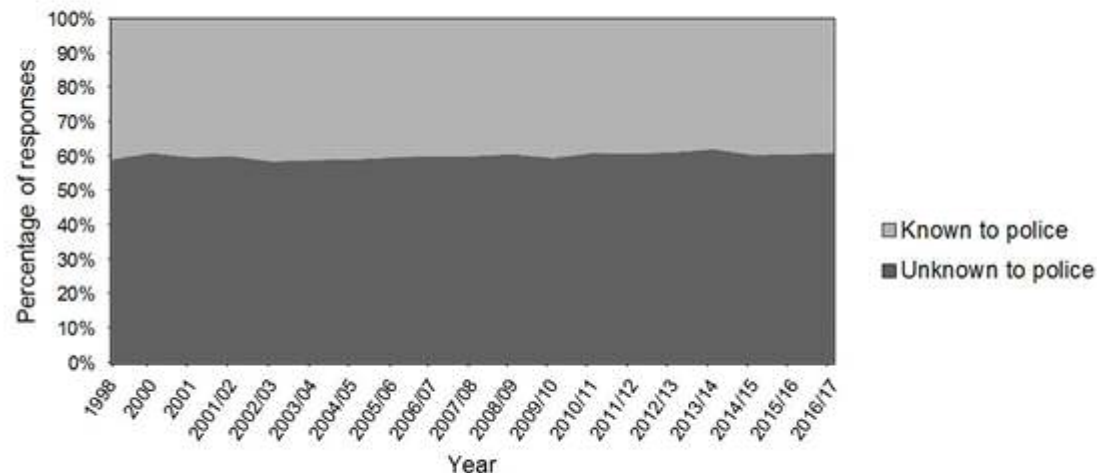
Direct estimates may be unreliable and suffer from low precision!

...we have a small area estimation problem

Data

- CSEW 2011/12 to 2015/16 – Dataset of victims of crime
- ‘Did the police come to know about the matter?’
- Temporal and spatial models for SAE

Figure 1. Percentage of crimes known and unknown to police (unweighted valid cases).



Source: Open access to the CSEW.

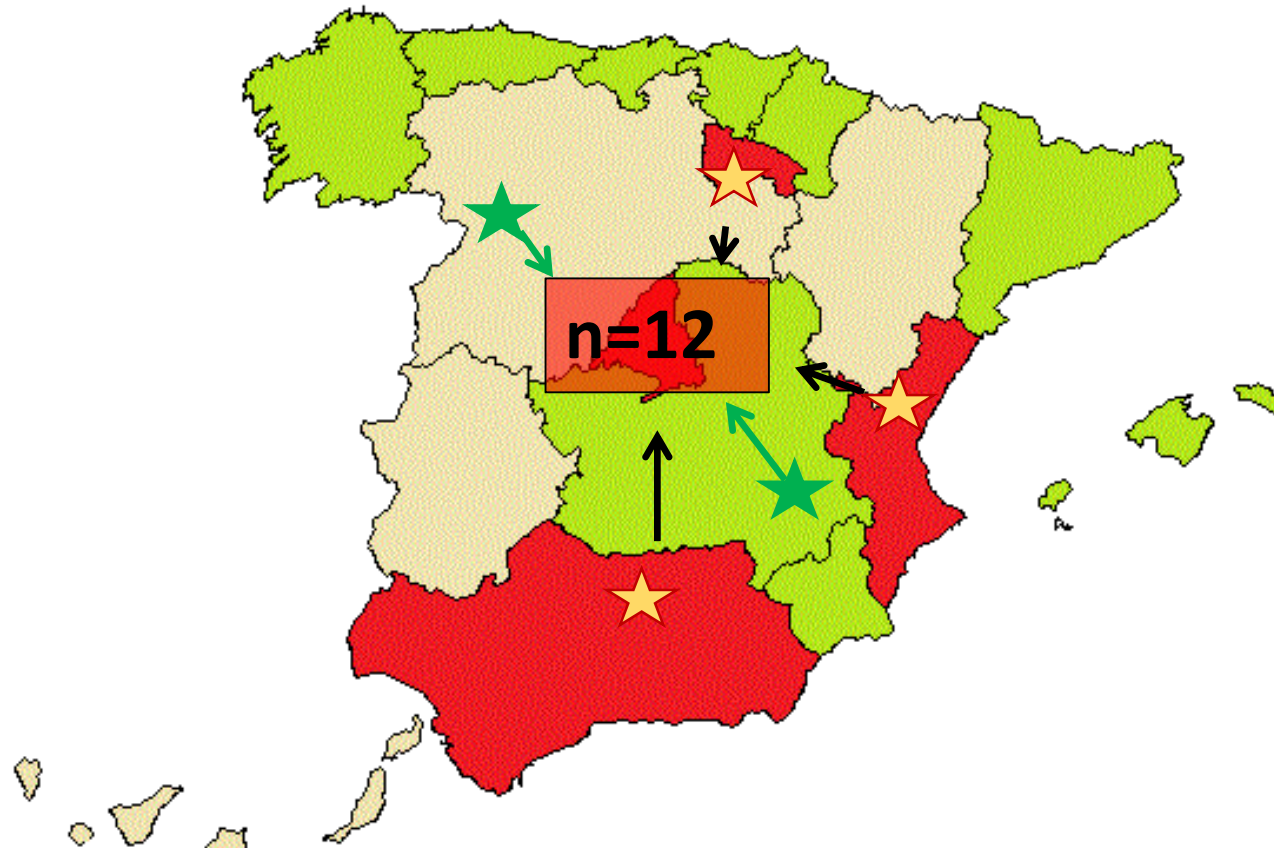
Table 1. Descriptive statistics about how the police come to know about crimes (unweighted valid cases).

		CSEW 2011/12	CSEW 2012/13	CSEW 2013/14	CSEW 2014/15	CSEW 2015/16	CSEW 2016/17
Police told by respondent	<i>f</i>	3445	2352	2147	1935	2004	1902
	%	63.2	63.2	65.0	63.4	64.0	65.5
Police told by another person	<i>f</i>	1759	1214	1032	974	978	872
	%	32.3	32.6	31.2	31.9	31.3	30.0
Police were there	<i>f</i>	127	79	65	65	84	58
	%	2.3	2.1	2.0	2.1	2.7	2.0
Police found out by another way	<i>f</i>	121	76	60	77	65	70
	%	2.2	2.0	1.8	2.5	2.1	2.4
Total	<i>f</i>	5452	3721	3304	3051	3131	2902

Source: Open access to the CSEW.

Small area estimation

- a. Direct estimation – Horvitz-Thompson (1952)
- b. Basic area-level EBLUP based on Fay-Herriot (1979)
- c. Spatial EBLUP (Petrucci and Salvati, 2006, Pratesi and Salvati, 2008)



Explaining the dark figure of crime

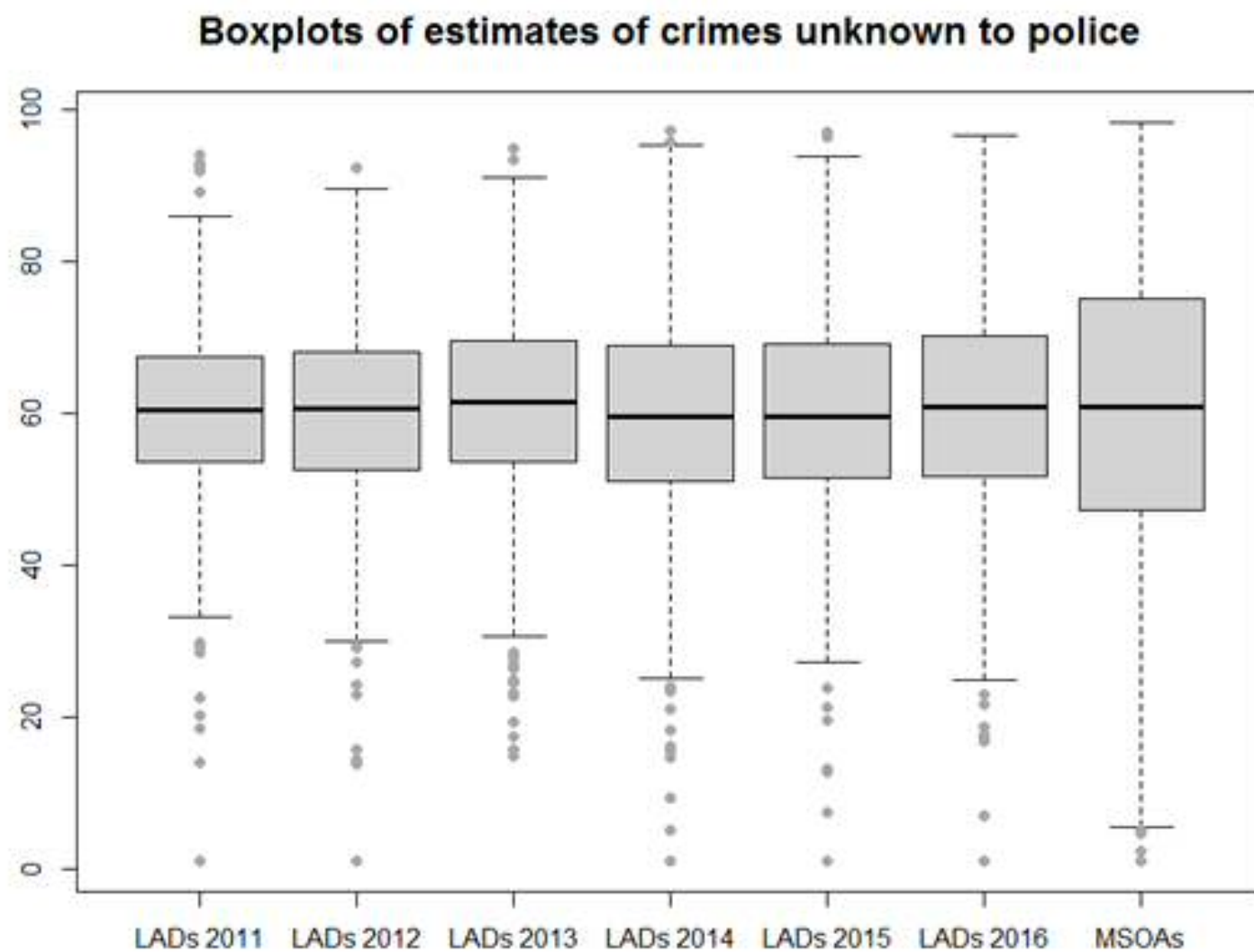
Table 5. Rao-Yu and STEBLUP models of crimes unknown to police at LAD level (standardised coefficients).

	Rao-Yu				STEBLUP			
	Beta	SE	t-value	p-value	Beta	SE	t-value	p-value
(Intercept)	59.920	0.3	181.83	0.000	60.96	2.1	28.37	0.000
ASS income	1.447	0.8	1.90	0.042	1.558	0.8	2.04	0.031
Conurbation	0.361	1.1	0.32	0.555	0.496	1.1	0.44	0.461
Small urban	2.455	1.0	2.46	0.013	2.246	1.0	2.24	0.025
Employed (%)	-0.317	0.8	-0.38	0.503	-0.493	0.8	-0.58	0.459
Whites (%)	-1.927	3.2	-0.59	0.349	-0.374	3.3	-0.11	0.511
Asians (%)	-1.040	2.9	-0.36	0.521	-0.025	2.9	-0.01	0.593
Mean house price	-0.572	1.0	-0.55	0.119	-1.071	1.1	-0.99	0.041
Crime rate	-0.809	0.9	-0.95	0.341	-0.543	0.9	-0.63	0.128
Males (%)	0.291	0.8	0.37	0.510	0.514	0.8	0.64	0.517
Spatial autocorrelation							0.21	
Temporal autocorrelation			0.11				0.09	
Log-likelihood			-8161.9				-8171.1	
Number of areas			331				331	

Table 6. EBLUP and SEBLUP models of crimes unknown to police at MSOA level (standardised coefficients).

	EBLUP				SEBLUP			
	Beta	SE	t-value	p-value	Beta	SE	t-value	p-value
(Intercept)	59.961	0.3	205.10	0.000	59.965	0.3	198.05	0.000
ASS income	1.227	0.7	1.74	0.083	1.257	0.7	1.71	0.048
Conurbation	-0.830	0.9	-0.84	0.403	-0.814	1.0	-0.81	0.419
Small urban	2.156	0.9	2.39	0.017	2.127	0.9	2.33	0.019
Employed (%)	-0.651	0.8	-0.83	0.408	-0.639	0.8	-0.81	0.420
Whites (%)	-2.621	4.3	-0.61	0.544	-2.747	4.4	-0.62	0.532
Asians (%)	-5.533	2.9	-1.86	0.063	-5.614	3.0	-1.85	0.044
Mean house price	-4.312	1.2	-3.59	0.000	-4.322	1.2	-3.56	0.000
Crime rate	-1.141	0.8	-1.33	0.183	-1.091	0.9	-1.27	0.204
Males (%)	0.633	0.7	0.87	0.387	0.616	0.7	0.84	0.401
No qualification (%)	4.530	1.7	2.70	0.006	4.453	1.7	2.63	0.008
High int. occupations (%)	-6.050	1.9	-3.24	0.001	-5.989	1.9	-3.18	0.001
Born UK (%)	-5.726	2.3	-2.46	0.013	-5.666	2.4	-2.41	0.016
Workday population density	1.292	0.9	1.47	0.141	1.237	0.9	1.39	0.094
Spatial autocorrelation							0.13	
Log-likelihood			-30554.82				-30556.88	
Number of areas			6657				6657	

Figure 2. Boxplots of model-based estimates of crimes unknown to police at LAD and MSOA levels.



The dark figure of crime

Figure 3. Model-based estimates of crimes unknown to police at the LAD level.

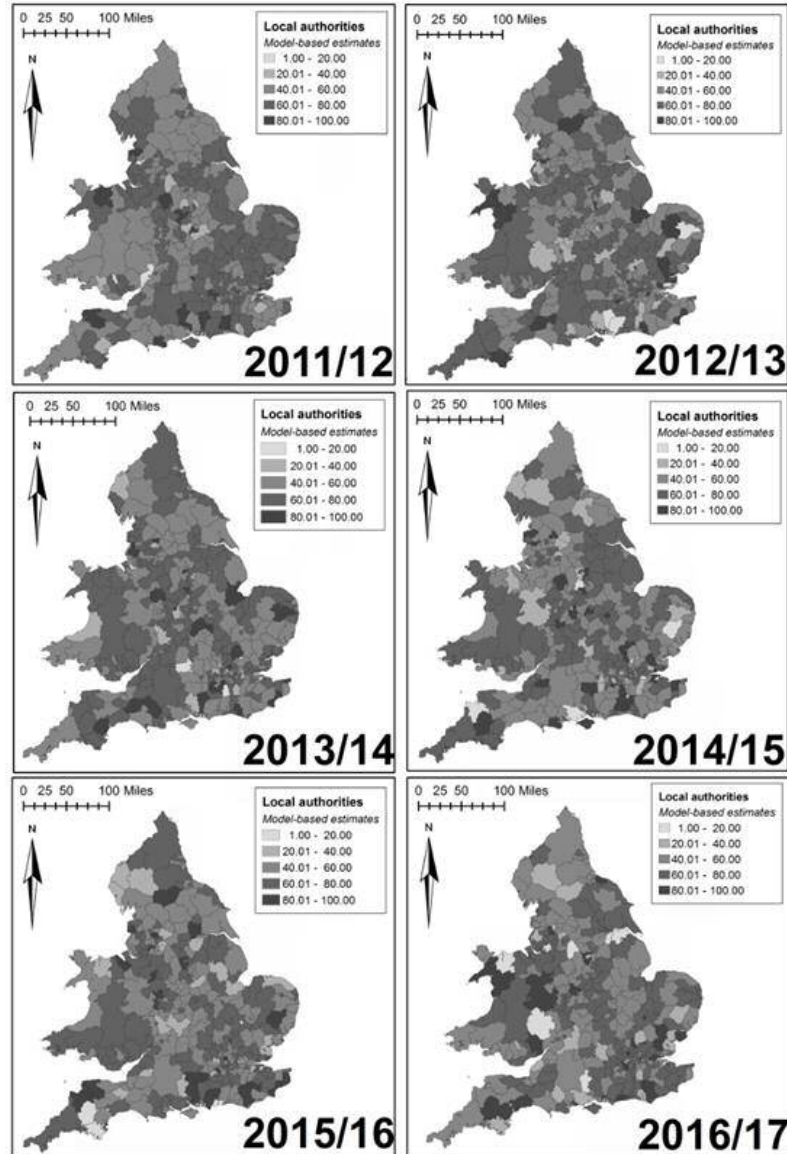
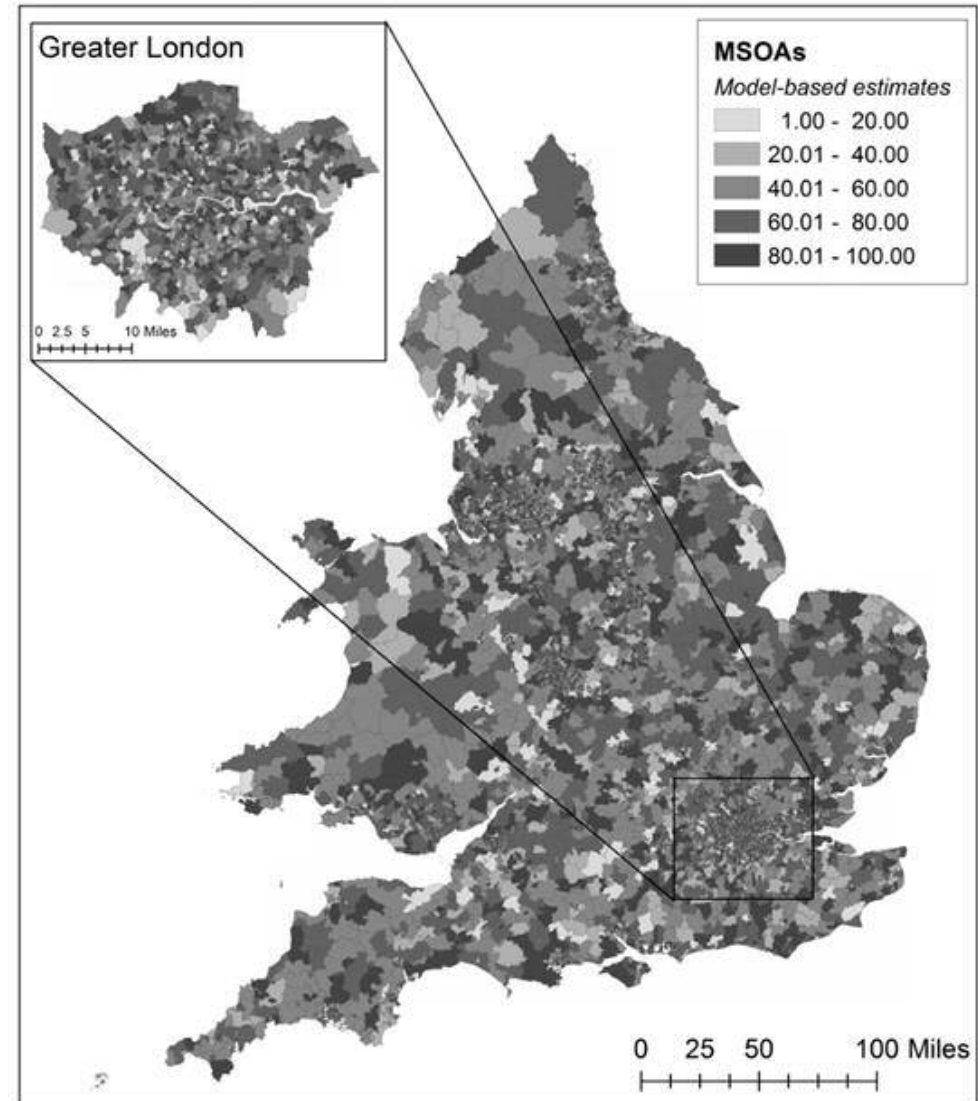
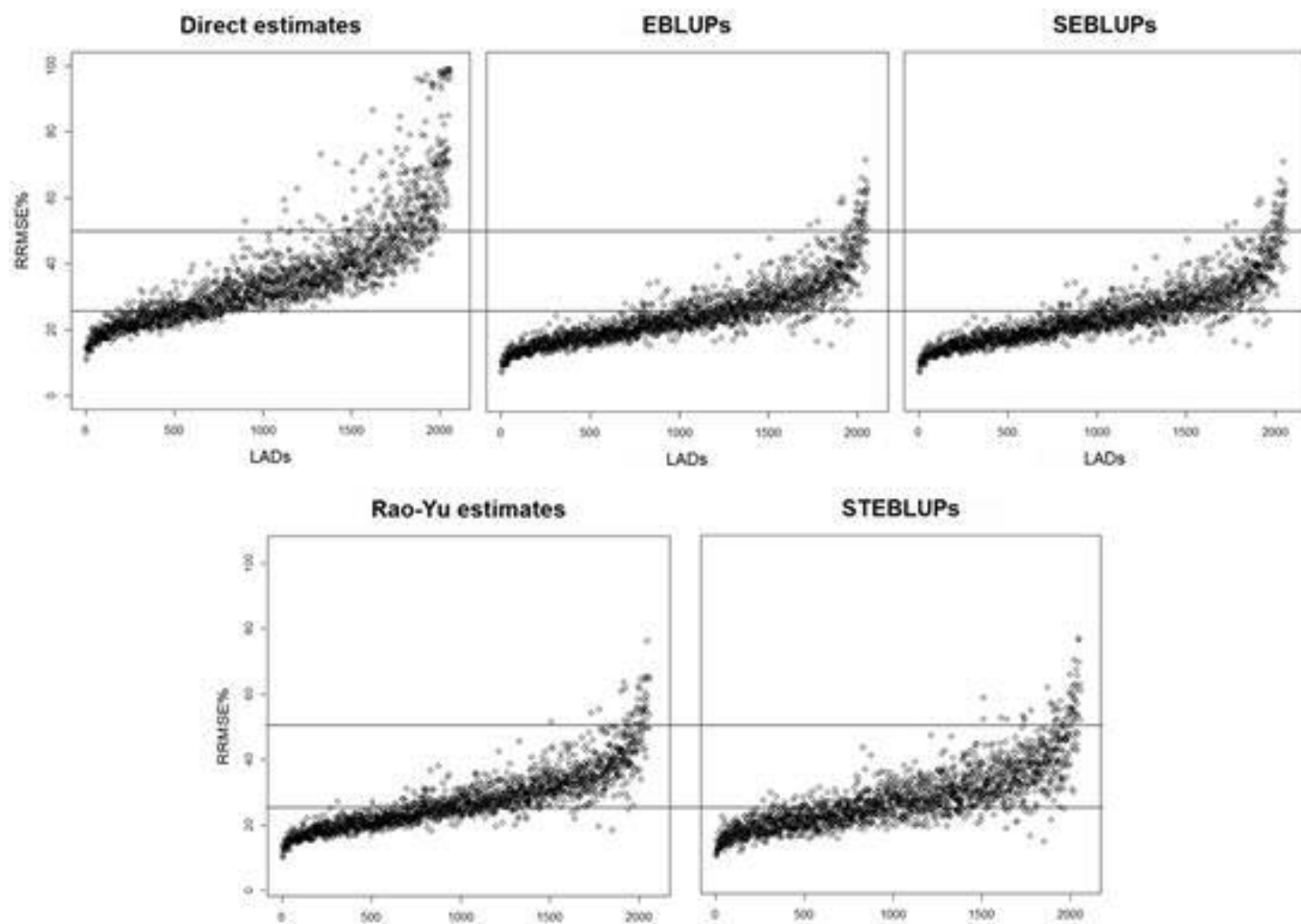


Figure 4. Model-based estimates of crimes unknown to police at the MSOA level (2011-2017).



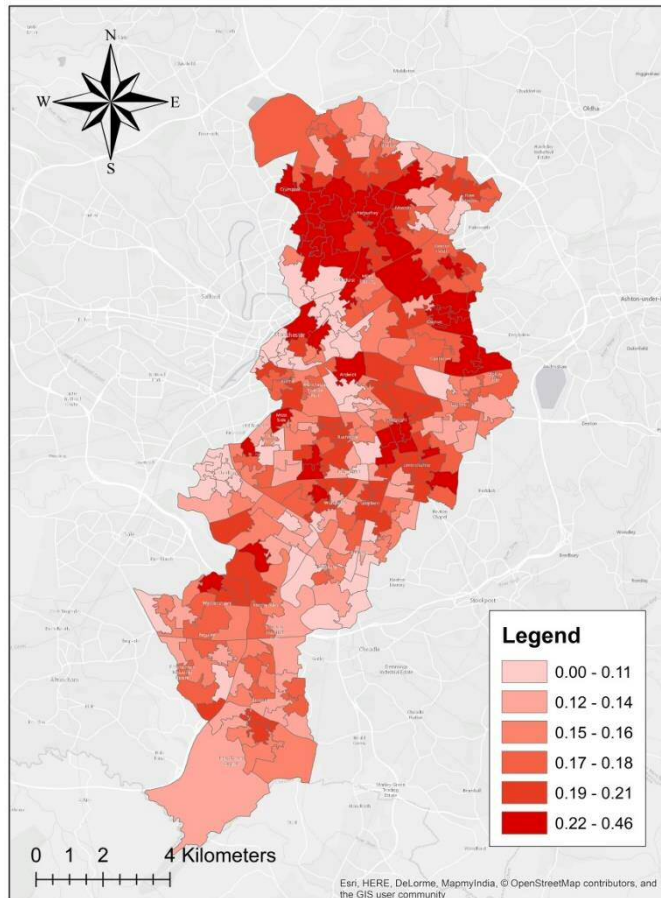
$\rho_2 = .09$

Figure 5. $RRMSE\%$ of small area estimates produced at LAD level (ordered by area sample size).



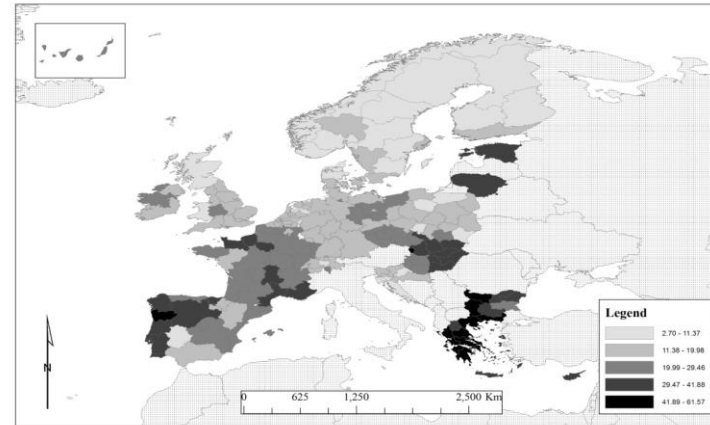
OTHER APPLICATIONS

Perceptions of disorder in Manchester



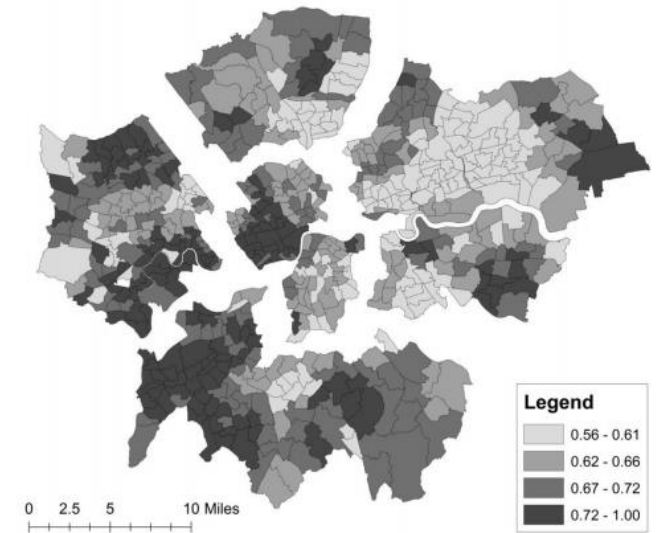
Buil-Gil, Medina and Shlomo (2019)

Dysfunctional worry about burglary at home

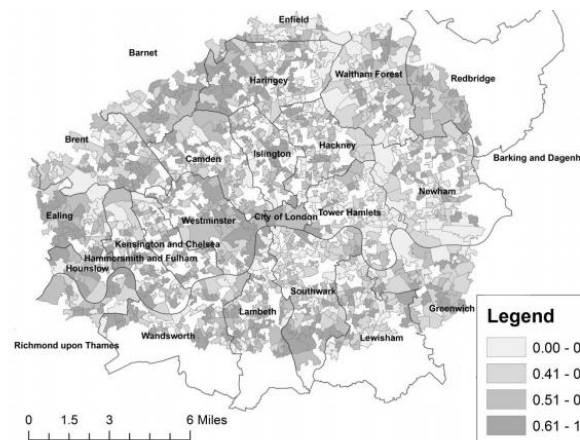


Buil-Gil, Moretti, Shlomo and Medina (2019)

Confidence in the police



Buil-Gil, Moretti, Shlomo and Medina (in press)



Perceived safety

Buil-Gil, Solymosi and Moretti (in press)

Conclusions (...for myself)

- If you produce maps from police data, think about potential measures of error and sources of bias.
- If you use a theoretical approach based on uniquely on police records, think if it is explaining reporting trends rather than crime.
- If you need access to the low-level geographical data of the CSEW, apply now!

Conclusions (...for criminologists)

- The dark figure of crime varies across areas (cities and neighbourhoods).
- The dark figure of crime is not stable over time.
- The dark figure of crime is associated with neighbourhood conditions (sometimes the same conditions that have been used to explain crime).

Thank you for your attention!

David Buil-Gil

 @davidbuil

david.builgil@manchester.ac.uk