

Recounting Crime

Project overview and synthetic crime data

Jose Pina-Sánchez

School of Law, University of Leeds
j.pinasanchez@leeds.ac.uk

David Buil-Gil

Department of Criminology, University of Manchester
david.builgil@manchester.ac.uk



Ian Brunton-Smith



Jose Pina-Sánchez



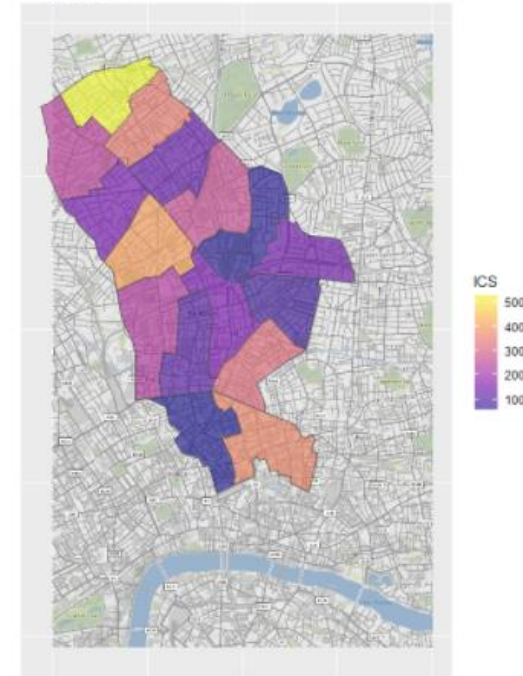
Alexandru Cernat



David Buil-Gil

Recounting crime is a research project funded by the ESRC Secondary Data Analysis Initiative. Our aim is to explore new statistical methods to help us improve the accuracy and precision of crime estimates.

Violent crime (estimates from local survey)
Islington, London



Crime statistics and measurement error

- Police-recorded crimes are used by:

Police forces → Design and evaluate policing strategies

Policy makers → Design and evaluate crime prevention policies

Academics → Explore the causes and consequences of crime and deviance

- However... police statistics are affected by:

Willingness to report crimes to police (varies by sex, age, ethnic group...)

Police control over areas (likelihood to witness crimes)

Counting rules

Police recording practices

} Dark
figure of
crime

Case Study:

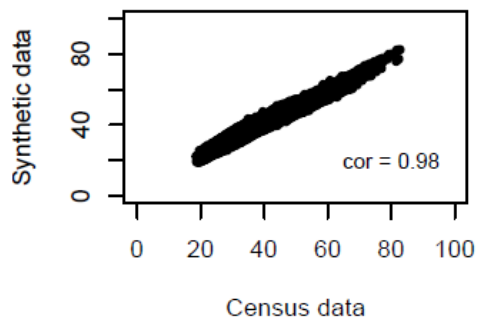
Can we generate reliable synthetic crime data for research and practice?

Method: Generating a synthetic population

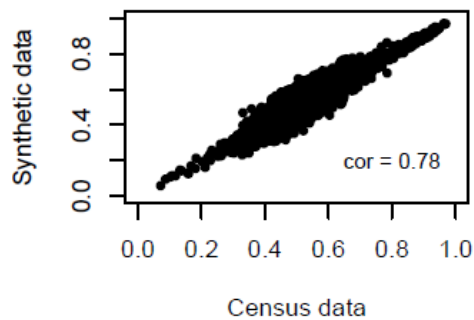
Simulation steps:

1. Simulating a synthetic population from Census 2011
 - Download census data aggregated in Output Areas
 - Obtain empirical parameters of age, sex, income, education, marital status and ethnicity
 - Generate synthetic population following a multivariate truncated normal and binary from empirical parameters in each area
2. Simulating crime victimisation from Crime Survey for England and Wales 2011/12
 - Estimate Negative Binomial regression models at individual level of (i) violent crime, (ii) property crime, and (iii) damage in CSEW
 - Same independent variables as in Step 1
 - Obtain regression parameter estimates
 - Simulate crime victimisation in synthetic population following Negative Binomial regression models

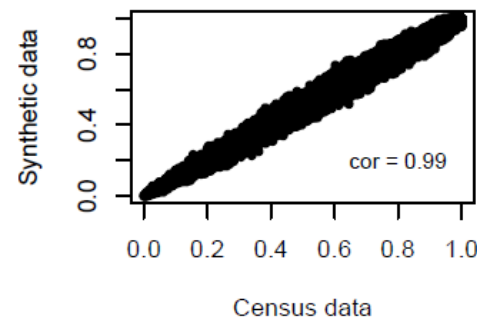
Mean age in OAs



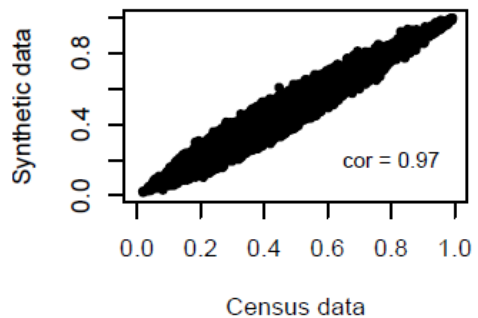
Proportion males in OAs



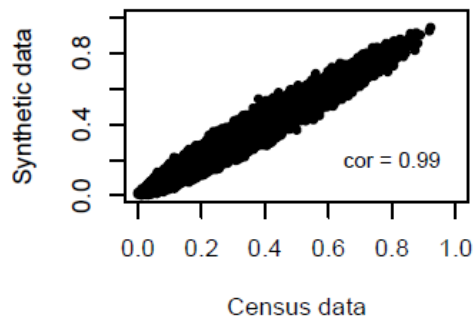
Proportion whites in OAs



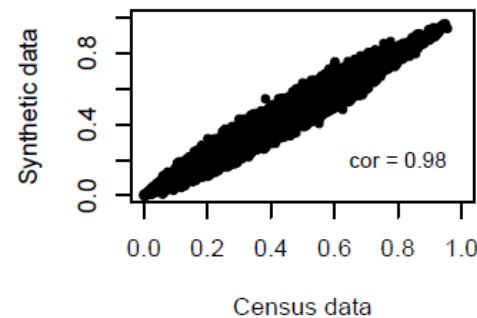
Proportion without income in OAs



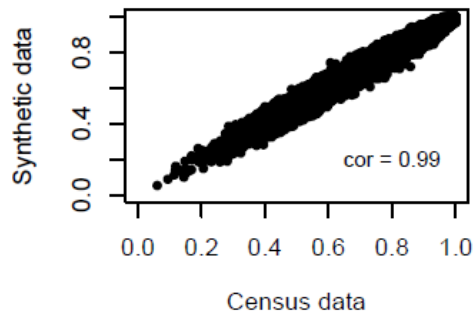
Proportion high education in OAs



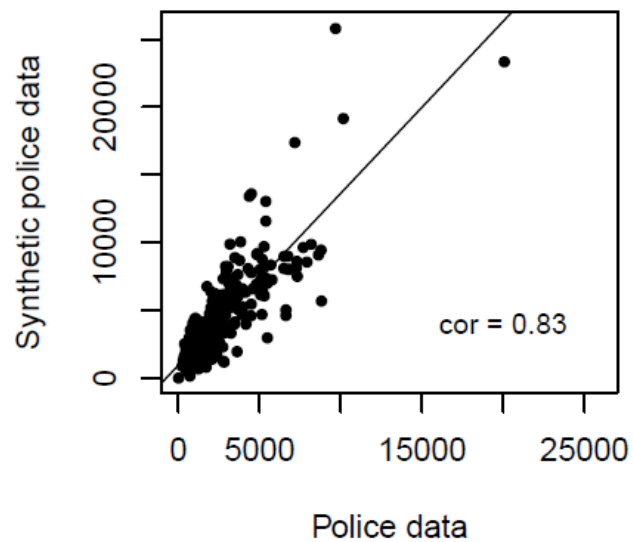
Proportion married in OAs



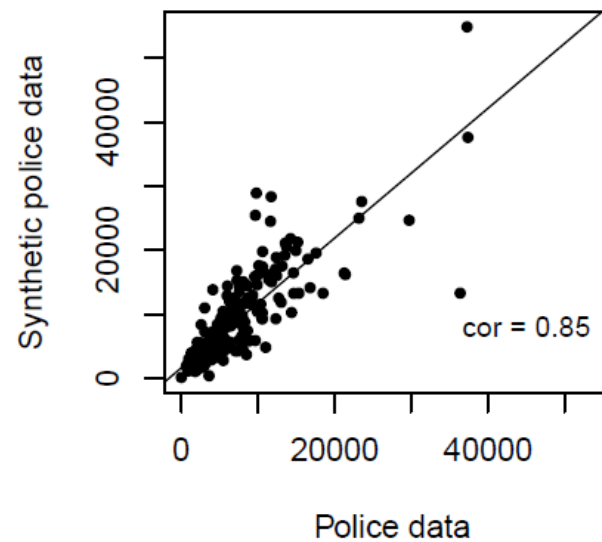
Proportion born in UK in OAs



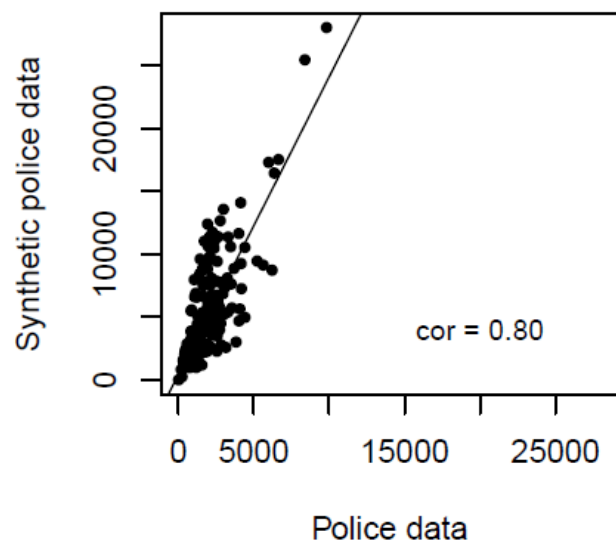
Violent crime in CSPs



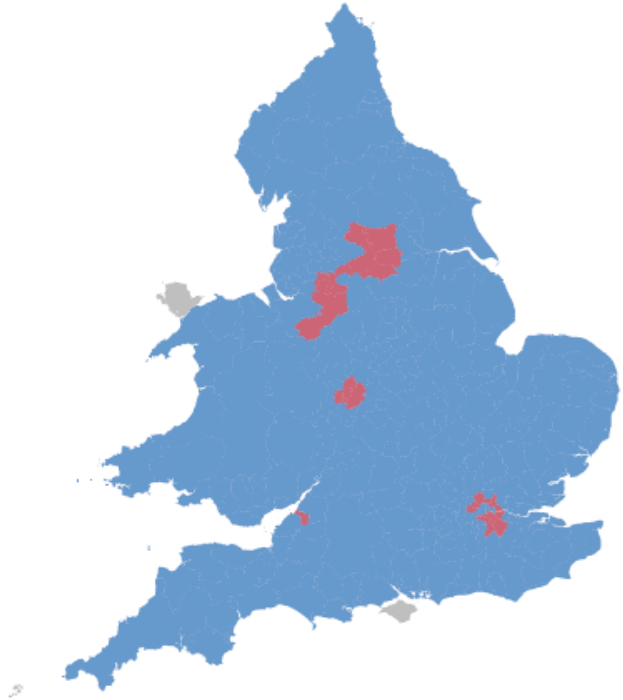
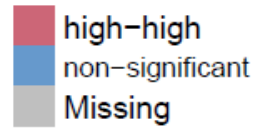
Property crime in CSPs



Damage crime in CSPs

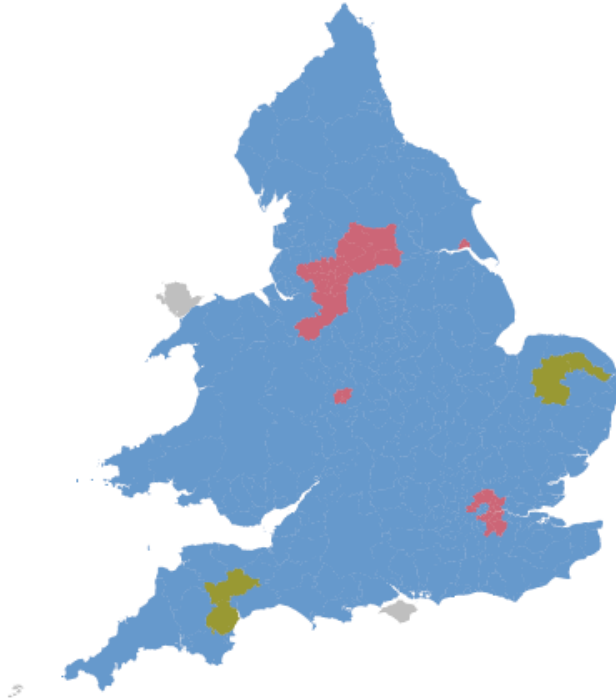
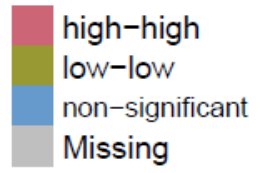


LISA All crime



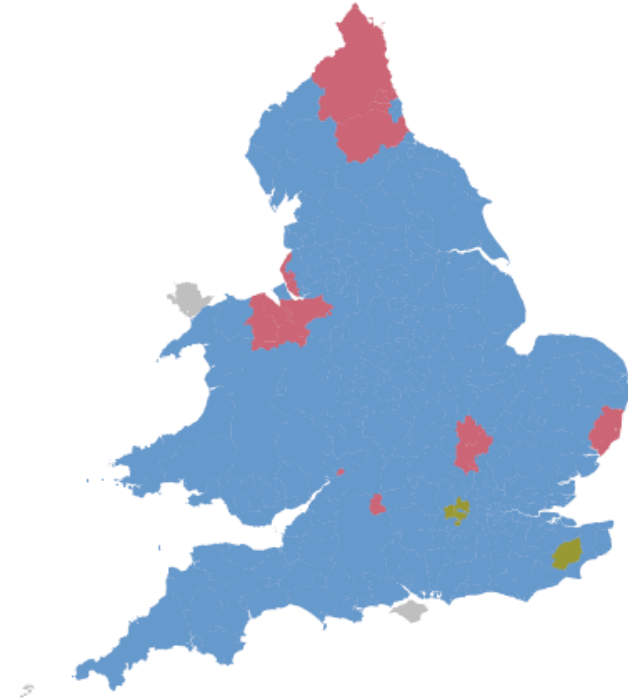
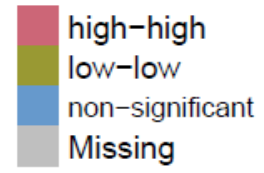
Moran's I = 0.24 (p=0.001)

LISA Police data



Moran's I = 0.36 (p=0.001)

LISA Survey data



Moran's I = 0.17 (p=0.001)

Conclusions

- Need to develop methods that account for ME in crime data
 - To build and test theories of crime and place
 - To design effective data-driven crime prevention approaches
- Synthetic crime data as an alternative approach to police records
- Synthetic data as a confidentiality-friendly alternative to crime statistics
- But more work is needed to ensure synthetic crime data are reliable and can be used in practice

Thank you for your attention!

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

David Buil-Gil
david.builgil@manchester.ac.uk