


Please cite the Published Version

Buil-Gil, David, Moretti, Angelo , Shlomo, Natalie and Medina, Juanjo (2020) Applying the Spatial EBLUP to place-based policing. Simulation study and application to confidence in police work. *Applied Spatial Analysis and Policy*, 13 (4). pp. 901-924. ISSN 1874-4621

DOI: <https://doi.org/10.1007/s12061-020-09333-8>

Publisher: Springer (part of Springer Nature)

Version: Published Version

Downloaded from: <https://e-space.mmu.ac.uk/625195/>

Usage rights:  [Creative Commons: Attribution 4.0](https://creativecommons.org/licenses/by/4.0/)

Additional Information: This is an Open Access article published in *Applied Spatial Analysis and Policy*, published by Springer, copyright The Author(s).

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)



Applying the Spatial EBLUP to Place-Based Policing. Simulation Study and Application to Confidence in Police Work

David Buil-Gil¹  · Angelo Moretti²  · Natalie Shlomo³  · Juanjo Medina¹ 

Received: 29 May 2019 / Accepted: 27 January 2020 / Published online: 09 March 2020
© The Author(s) 2020

Abstract

There is growing need for reliable survey-based small area estimates of crime and confidence in police work to design and evaluate place-based policing strategies. Crime and confidence in policing are geographically aggregated and police resources can be targeted to areas with the most problems. High levels of spatial autocorrelation in these variables allow for using spatial random effects to improve small area estimation models and estimates' reliability. This article introduces the Spatial Empirical Best Linear Unbiased Predictor (SEBLUP), which borrows strength from neighboring areas, to place-based policing. It assesses the SEBLUP under different scenarios of number of areas and levels of spatial autocorrelation and provides an application to confidence in policing in London. The SEBLUP should be applied for place-based policing strategies when the variable's spatial autocorrelation is medium/high, and the number of areas is large. Confidence in policing is higher in Central and West London and lower in Eastern neighborhoods.

Keywords Spatial correlation · SAR · Contiguity matrix · Spatial model · Police legitimacy · London

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12061-020-09333-8>) contains supplementary material, which is available to authorized users.

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

¹ Department of Criminology, University of Manchester, G18 Humanities Bridgeford Street, M13 9PL, Manchester, UK

² Department of Computing and Mathematics, Manchester Metropolitan University, Manchester, UK

³ Social Statistics Department, University of Manchester, Manchester, UK

Introduction

Policing analyses and intelligence-led policing are moving towards the study of small geographic areas, or micro places, to develop place-based policing strategies to reduce crime and disorder (Hutt et al. 2018; Weisburd 2018). Place-based policing draws from the empirical observation that crime is concentrated at micro geographical units, which are sometimes referred to as ‘hot spots of crime’ (Weisburd 2015, 2018). Sherman et al. (1989) found that only 3.5% of addresses in the city of Minneapolis produce 50% of all annual crime calls to the police. Pierce et al. (1988) found similar results in Boston: 2.6% of addresses produce the 50% of police calls. Weisburd et al. (2004) examined the distribution of crime in Seattle from 1989 to 2002, and found that 50% of crimes were located at 4.5% of street segments, which showed that the concentration of crimes in small areas is stable across time. Therefore, Weisburd (2015) argues that there is a law of crime concentration, which states that “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). Place-based policing interventions target those areas with high levels of crime and are successful in reducing crime and disorder, as shown by Braga et al. (2014) in their meta-analysis of quasi-experimental evaluations of hot spots policing. Braga et al. (2014) also found that the crime control benefits of such strategies diffuse into areas surrounding targeted places. This shows the need for the study of small areas in policing research and practice. However, the police effectiveness in reducing crime in places highly depends on its relationship with the public (Bennett et al. 2014; Jackson et al. 2013; Tyler and Bies 1990; Weisburd 2018). Areas with higher confidence in police work tend to have larger citizens’ cooperation with the police, thus enhancing the police capacity to prevent crime and deviance. Moreover, government inspections into police forces assess not only their effectiveness in reducing crime, but also they expect the police to develop programs to enhance its legitimacy and public confidence in those geographical areas where public cooperation with police services is lower (HMICFRS 2017). The confidence in police work is also distributed at micro places (Williams et al. 2019), and thus should be taken into account to design place-based policing strategies.

Police-recorded offences and crime calls are relatively easy to geocode and map, and advanced geographical analyses can be drawn from crime maps with a high level of spatial accuracy (Hutt et al. 2018). However, the confidence in policing cannot be directly observed and is mainly recorded by crime surveys, such as the Crime Survey for England and Wales (CSEW) and the National Crime Victimization Survey (NCVS). Crime surveys are usually designed to record large samples and provide reliable direct estimates only for large geographies, such as regions or cities, and small areas within these are usually unplanned domains and have small or zero sample sizes. This is the reason why more advanced statistical methods are needed to map the confidence in police work. Groves and Cork (2008) argue that model-based small area estimation (SAE) techniques are a potential tool to overcome such limitations and produce reliable small area estimates from crime surveys. SAE seeks to produce reliable estimates for

unplanned areas where direct estimates are not precise enough (Rao and Molina 2015). Those estimates allow for advanced geographical analyses and precise maps of the confidence in policing and associated constructs.

In this paper we provide background information, a simulation study and an application to introduce model-based SAE techniques that account for spatially correlated random area effects to place-based policing. This is one of the first papers that evaluates and applies these methods in policing research and practice. Confidence in police work tends to show high levels of spatial clustering (Jackson et al. 2013; Williams et al. 2019), which can be taken into account in SAE models to increase the estimates' precision. In SAE, the use of spatially correlated random area effects is increasingly in use (Chandra et al. 2007; Petrucci and Salvati 2006; Pratesi and Salvati 2008; Salvati et al. 2014). Small area estimators that incorporate the spatial autocorrelation parameter have been shown to reduce the estimates' mean squared error when the level of spatial autocorrelation (henceforth ρ) is large. ρ measures the correlation of a variable with itself across neighboring areas. Thus, a large ρ means that geographically nearby areas tend to have similar values (i.e. high values of a variable in one area are surrounded by high values in neighboring areas and low values of a variable in one area are surrounded by low values in neighboring areas), while a ρ close to zero represents a geographically random phenomenon. Specifically, this paper introduces the Spatial Empirical Best Linear Unbiased Predictor (SEBLUP) to place-based policing. The SEBLUP is an extension of the Empirical Best Linear Unbiased Predictor (EBLUP), which is based on the Fay-Herriot (FH) model (Fay and Herriot 1979), considering correlated random area effects between neighboring areas through the simultaneous autoregressive (SAR) process (Cressie 1993; Salvati 2004).

The level of ρ of the variable of interest has shown to be relevant to improve SEBLUP estimates. Less attention has been paid to the effect of the number of areas under study, m , on SEBLUP's performance, and particularly how m interacts with ρ to explain the SEBLUP's increased precision. m measures the number of geographical areas for which we aim to produce estimates. For example, confidence in police work can be estimated in London at a metropolitan ($m = 1$), borough ($m = 32$) or ward level ($m = 610$), or even at lower geographical scales with larger number of areas. This is especially relevant for crime analysts and police departments aiming to select appropriate methods to estimate confidence in police work at different geographical scales with dissimilar number of areas. There are few studies examining the efficiency of the SEBLUP under different geographical conditions and these show contradicting results (Asfar and Sadik 2016; Petrucci and Salvati 2006; Pratesi and Salvati 2008; Salvati 2004). Thus, further examinations and applications of the method are needed.

This paper assesses the SEBLUP performance, in terms of bias and mean squared error, under different scenarios with unequal m and ρ , and provides an empirical evaluation and application to confidence in police work in London. The confidence in policing is measured here by the proportion of people who think that the police do a good job (Stanko and Bradford 2009). Thus, we gain evidence about the SEBLUP estimates' reliability under different conditions, to examine the cases in which this estimator provides better estimates than basic model-based estimators when applied to policing data. In the simulation study, quality measures

for SEBLUP estimates are compared to post-stratified and EBLUP estimates controlling for m and ρ . In the empirical evaluation, estimates of confidence in police work are produced at ward level in five London sub-regions with different number of wards. Furthermore, the application contributes to the increasing criminological research on understanding the geographical distribution of citizens' confidence in the police (Jackson and Bradford 2010; Jackson et al. 2013; Tankebe 2012).

Section 2 provides background information on the need for accounting for the confidence in police work in policing strategies, and section 3 bridges the gap between SAE and place-based policing. Section 4 describes the SEBLUP and results of previous studies. Section 5 presents the simulation study and its results. Section 6 applies SEBLUP to produce estimates of confidence in police work in London. Section 7 draws final conclusions.

Confidence in the Police and Policing Strategies

The police effectiveness in maintaining order and preventing crime depends on its relationship with the public (Jackson and Bradford 2010; Jackson et al. 2013). Citizens' willingness to cooperate and support police officers is essential for an effective policing service, and public cooperation with the police is shaped by the citizens' trust in police work (Bennett et al. 2014; Tyler 2004). The residents' confidence in police services, which shows heterogeneity between neighborhoods, affects the unequal police capacity to prevent crime in different areas. Thus, effective policing strategies need to develop measures to enhance the public confidence in police work, and inspections into police forces assess the efforts made by the police to increase their public confidence at different geographical areas (HMICFRS 2017). This is especially important in the case of place-based policing strategies, which have been criticised for having negative impacts on the perceptions about the police of targeted communities (Rosenbaum 2006).

Confidence in policing and police legitimacy are known to be driven by a series of demographic and social variables that operate at individual, micro and meso levels, and increasing research focuses on understanding their predictors at different scales. Several individual characteristics have been related with decreased confidence in police work and less willingness to cooperate with the police, such as being male and young, belonging to an ethnic minority, low education, poverty, negative perceptions of procedural justice and negative experiences with the police (Jackson et al. 2013; Sampson and Bartusch 1998; Tankebe 2012; Tyler 2004). Particular attention has been given to the study of the relationship between procedural justice and public confidence in police: citizens tend to be more confident in police services and legitimize police activities when police officers are perceived to treat people with respect and dignity (Tyler 2004; Tyler and Bies 1990).

Research has also found that confidence in policing is higher in certain neighborhoods than others, and the confidence and trust in the police are known to be influenced by neighborhood-level variables that operate at the scales of small communities (Jackson et al. 2013; Sampson and Bartusch 1998). Some of the variables used to explain the unequal distribution of the neighbors' confidence in police work and associated constructs are the average income, unemployment

rates, social cohesion, residential mobility, concentration of minorities and immigrants, and crime rates (Bradford et al. 2017; Dai and Johnson 2009; Jackson et al. 2013; Kwak and McNeeley 2017; Sampson and Bartusch 1998; Wu et al. 2009). Wu et al. (2009:150) argue that “racial composition, concentrated disadvantage, residential mobility, and violence crime rate are all good neighborhood-level predictors in determining public perception of police”. Sampson and Bartusch (1998) found that the combined effect of concentrated disadvantage, crime and ethnic concentration explains 82% of the variation between small areas in levels of satisfaction with police. Neighborhood poverty and unemployment, as forms of concentrated disadvantage, are known to shape neighbors’ social identities and decrease citizens’ attitudes and perceptions of policing services (Wu et al. 2009). Confidence in police work tends to be lower in deprived areas, while wealthy neighborhoods have more confidence in the police. While some argue that this is due to the larger police control and the more violent techniques used by the police in deprived areas (Dai and Johnson 2009), others argue that it is explained by differential social identities within cities: “residents of more socially integrated neighborhoods may feel they are connected to larger formal institutions such as the police” (Kwak and McNeeley 2017:10). People living in poor socioeconomic conditions are not only likely to be dissatisfied with the police, but with all government services (Dai and Johnson 2009).

The concentration of minorities and immigrants has also been used to explain neighborhood-level confidence in policing. Areas with larger concentrations of minorities and immigrants are likely to have lesser confidence in police work (Sampson and Bartusch 1998; Wu et al. 2009), although research conducted in the United Kingdom has found the opposite: “trust in the police was on average higher among immigrants to the United Kingdom than among the UK-born population” (Bradford et al. 2017:381). Dai and Johnson (2009) argue that the relationship between concentration of minorities and dissatisfaction with the police in the US is likely to be explained by the neighborhood concentrated disadvantage, as citizens from minority groups are disproportionately represented in deprived areas. In relation to crime rates, Kwak and McNeeley (2017) and Wu et al. (2009) found that, contrarily to what one might expect, these are not significant in predicting confidence in policing and dissatisfaction with the police. We will use this information to select covariates to fit our SAE models of confidence in policing.

Small Area Estimation in Place-Based Policing

Since 2008, when the US Panel to Review the Programs of the Bureau of Justice Statistics suggested the use of model-based SAE to produce estimates from the NCVS (Groves and Cork 2008), there have been several applications of SAE methods to policing data. Buelens and Benschop (2009) used the EBLUP based on the FH model to produce estimates of victimization rate per police zone in Netherlands. Fay and Diallo (2012) presented an extension of the temporal model developed by Rao and Yu (1994) and applied it to estimate crime by states in the US. Whitworth (2012) produced regression-based synthetic estimates of fear of crime in England and Wales. Taylor (2013) made use of multilevel models to produce synthetic estimates of perceived

antisocial behaviour in England and Wales. Williams et al. (2019) introduced the spatially correlated random area effects and produced neighborhood estimates of public confidence in policing from a spatiotemporal Bayesian approach. Wheeler et al. (2017) made use of spatial models to produce synthetic estimates of attitudes towards the police. Regression-based synthetic estimates, however, are known to suffer from a high risk of bias arising from possible misspecification of models (Rao and Molina 2015). Spatial microsimulation approaches have also been used to produce estimates of crime rates (Kongmuang 2006).

Several of these studies have shown the need for incorporating the spatial autocorrelation parameter to SAE when producing estimates for designing place-based policing strategies. The spatial autocorrelation accounts for the geographical concentration of attitudes towards policing and estimators that incorporate it tend to provide more precise estimates than basic model-based estimators. The SEBLUP has shown promising results not only in simulation studies (Asfar and Sadik 2016; Chandra et al. 2007; Pratesi and Salvati 2008; Salvati 2004), but also when it has been applied to social science research, such as the estimation of poverty (Salvati et al. 2014). Thus, the SEBLUP is expected to produce promising results in the field of place-based policing. Hence, we aim to bridge this gap by demonstrating its use for estimating confidence in police work at small area level. In order to gain evidence about cases in which the SEBLUP provides better estimates than basic model-based estimators when applied to policing data, we provide a simulation study and an application.

Model Description: SEBLUP

Let us consider a target population partitioned into m small areas. In our application, estimates of confidence in policing will be produced for London wards, thus, m equals 610. In the traditional EBLUP derived from the FH model (Fay and Herriot 1979), we assume that a linking model linearly relates the quantity of inferential interest (i.e. proportion of citizens who think that police do a good job), which is usually an area mean or total δ_i , to p area level auxiliary variables $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})'$ with a random effect v_i :

$$\delta_i = \mathbf{x}_i' \boldsymbol{\beta} + v_i, \quad i = 1, \dots, m, \quad (1)$$

where $\boldsymbol{\beta}$ is the $p \times 1$ vector of regression parameters and $v_i \sim iid(0, \sigma_v^2)$. In our case, δ_i represents the confidence in police work and \mathbf{x}_i denotes the covariates known to be associated to confidence in policing (e.g. unemployment, concentration of minorities, poverty). The model assumes that a design-unbiased direct estimate denoted y_i for δ_i , which is obtained from the observed sample, is available for each area $i = 1, \dots, m$:

$$y_i = \delta_i + e_i, \quad i = 1, \dots, m, \quad (2)$$

where $e_i \sim N(0, \psi_i)$ denotes the sampling errors, independent of v_i , and ψ_i refers to the sampling variance of the direct estimates (Rao and Molina 2015).

The SEBLUP borrows strength from neighboring areas by adding spatially correlated random area effects (Petrucchi and Salvati 2006; Salvati 2004). If we combine (1) with (2) we can write the following model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{v} + \mathbf{e}, \quad (3)$$

where $\mathbf{y} = (y_1, \dots, y_m)'$ is the vector of direct estimates of confidence in policing for m areas, $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_m)'$ denotes the covariates associated to the outcome measure for m areas, $\mathbf{v} = (v_1, \dots, v_m)'$ is a vector of area effects and $\mathbf{e} = (e_1, \dots, e_m)'$ is a vector of sampling errors independent of \mathbf{v} . We assume \mathbf{v} to follow a SAR process with unknown autoregression parameter $\rho \in (-1, 1)$ and a contiguity matrix \mathbf{W} (Cressie 1993):

$$\mathbf{v} = \rho \mathbf{W}\mathbf{v} + \mathbf{u}, \quad (4)$$

where ρ represents the spatial autocorrelation coefficient of our outcome measure (i.e. confidence in policing) and \mathbf{W} is a standardised matrix that relates each area with all neighboring areas.

We also assume $(\mathbf{I}_m - \rho \mathbf{W})$ to be non-singular, where \mathbf{I}_m is a the $m \times m$ identity matrix, so we can express (4) as follows:

$$\mathbf{v} = (\mathbf{I}_m - \rho \mathbf{W})^{-1} \mathbf{u}, \quad (5)$$

where $\mathbf{u} = (u_1, \dots, u_m)'$ satisfies $\mathbf{u} \sim N(\mathbf{0}_m, \sigma_u^2 \mathbf{I}_m)$. Thus,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I}_m - \rho \mathbf{W})^{-1} \mathbf{u} + \mathbf{e} \quad (6)$$

The vector of variance components are denoted as $\boldsymbol{\theta} = (\theta_1, \theta_2)' = (\sigma_u^2, \rho)'$. Then, the Spatial Best Linear Unbiased Predictor (SBLUP) of $\delta_i = \mathbf{x}_i' \boldsymbol{\beta} + v_i$ is given by

$$\tilde{\delta}_i^{SBLUP}(\boldsymbol{\theta}) = \mathbf{x}_i' \tilde{\boldsymbol{\beta}}(\boldsymbol{\theta}) + \mathbf{b}_i' \mathbf{G}(\boldsymbol{\theta}) \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) \{\mathbf{y} - \mathbf{X} \tilde{\boldsymbol{\beta}}(\boldsymbol{\theta})\} \quad (7)$$

where \mathbf{b}_i' is a $1 \times m$ vector $(0, \dots, 1, 0, \dots, 0)$ with 1 in position i . $\mathbf{G}(\boldsymbol{\theta})$, the covariance matrix of \mathbf{v} , is given by $\mathbf{G}(\boldsymbol{\theta}) = \sigma_u^2 \{(\mathbf{I}_m - \rho \mathbf{W})'(\mathbf{I}_m - \rho \mathbf{W})\}^{-1}$. $\boldsymbol{\Sigma}(\boldsymbol{\theta})$, which is the covariance matrix of \mathbf{y} , is defined as $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \mathbf{G}(\boldsymbol{\theta}) + \boldsymbol{\Psi}$, where $\boldsymbol{\Psi} = \text{diag}(\psi_1, \dots, \psi_m)$. And $\tilde{\boldsymbol{\beta}}(\boldsymbol{\theta})$, the weighted least squares estimator of $\boldsymbol{\beta}$, is obtained as $\tilde{\boldsymbol{\beta}}(\boldsymbol{\theta}) = \{\mathbf{X}' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) \mathbf{X}\}^{-1} \mathbf{X}' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) \mathbf{y}$.

The SEBLUP is obtained by replacing a consistent estimator of $\boldsymbol{\theta}$ by $\hat{\boldsymbol{\theta}} = (\hat{\sigma}_u^2, \hat{\rho})'$:

$$\hat{\delta}_i^{SEBLUP} = \tilde{\delta}_i^{SEBLUP}(\hat{\boldsymbol{\theta}}) = \mathbf{x}_i' \tilde{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}) + \mathbf{b}_i' \mathbf{G}(\hat{\boldsymbol{\theta}}) \boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\theta}}) \{\mathbf{y} - \mathbf{X} \tilde{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})\}. \quad (8)$$

If we assume the normality of the random effects, we can estimate σ_u^2 and ρ based on different procedures. In this research, we consider the Restricted Maximum Likelihood estimator, which takes into account for the loss in degrees of freedom

derived from estimating β , while other estimators, such as the Maximum Likelihood estimator, do not (Rao and Molina 2015). The assumption of normality of the random effects is reasonable in those cases in which area-level direct estimates are normally distributed, as tends to be the case in criminological studies looking into the confidence in police work (Williams et al. 2019), emotions about crime (Whitworth 2012) and rates of some crime types at large spatial scales (Fay and Diallo 2012). However, such assumption may be considered invalid in those cases in which the normality of direct estimates is not met. This may be the case of studies analysing specific crime types at detailed spatial scales, as these may show zero inflated skewed distributions and thus robust SAE techniques adjusted to non-normal distributions are needed (Dreassi et al. 2014).

Previous Studies Using the SEBLUP

The SEBLUP has not yet been used to estimate crime rates or confidence in the police. However, a series of simulation studies and applications analysing economic and agricultural outcomes have shown that the SEBLUP tends to outperform EBLUP estimators when ρ moves away from zero -especially when it is close to -1 or 1 (Chandra et al. 2007; Petrucci and Salvati 2006; Pratesi and Salvati 2008). There are very few simulation studies that investigate the impact of m , and the interaction between m and ρ , on the SEBLUP's performance, and these show contradicting results. Salvati (2004) examined the precision of SEBLUP estimates for m equal to 25 and 50, and $\rho = \{\pm 0.25, \pm 0.5, \pm 0.75\}$, and concluded that the improvement in the estimates' accuracy is higher when the spatial autoregressive coefficient increases, but also that "benefit is bigger as the number of small areas increase" (Salvati 2004:11). In policing research, the SEBLUP is thus expected to produce more reliable estimates than the EBLUP when the values of the variable of interest geographically cluster together, as observed in many studies on crime and crime perceptions (Baller et al. 2001; Williams et al. 2019), and when the number of areas for which we aim to produce estimates is large. Therefore, in cases like the one encountered by Gemmell et al. (2004), who produced estimates of drug use for ten local authorities in Greater Manchester, the EBLUP is expected to produce better estimates than the SEBLUP due to the small number of areas under study.

Asfar and Sadik (2016) analyzed the SEBLUP's relative mean squared errors under m equal to 16, 64 and 144, and they found large relative improvement of SEBLUP estimates even when ρ is very small ($\rho = 0.05$) and small ($\rho = 0.25$), also in cases of very few areas under study ($m = 16$). In addition, such improvement was sometimes larger when m was equal to 16 than in cases of m equal to 64 and 144. These results are not consistent with other simulation studies, which show that SEBLUP's relative performance improves as the number of areas increases (Salvati 2004), and the SEBLUP's precision is not improved if $\rho \approx 0$ in cases of m equal to 25 and 50 (Salvati 2004), 61 (Petrucci and Salvati 2006), 23 (Chandra et al. 2007) and 42 (Pratesi and Salvati 2008). Therefore, further research is needed to understand how both ρ and m affect the SEBLUP's relative precision, and we assess the performance of the SEBLUP in Section 5.

Simulation Study

In this section we describe the simulation study designed to assess the effect of m and ρ on the SEBLUP's performance in comparison to EBLUP and post-stratified estimators.

Generating the Population and Simulation Steps

The population is generated based on previous simulation studies such as Petrucci and Salvati (2006) and Pratesi and Salvati (2008). Similar approaches have also been used in Asfar and Sadik (2016), Molina et al. (2009) and Salvati (2004). Simulation parameters are based on previous simulation experiments to allow comparisons and reproducibility. The population is generated following a linear mixed-effect model with random area effects of neighboring areas correlated to the SAR dispersion matrix with fixed autoregressive coefficient:

$$y_{ij} = x_{ij}\beta + v_i + e_{ij}, \quad i = 1, \dots, m, \quad j = 1, \dots, N_i, \quad (9)$$

where x_{ij} is the value of the covariate x for unit j in area i , v_i denotes the area effect and e_{ij} is the individual error. The simulation parameters are given as follows: $\beta = 0.74$, $\sigma_u^2 = 90$, $\sigma^2 = 1.5$ (Petrucci and Salvati 2006). $\mathbf{v} = [v_1, \dots, v_m]'$ is generated from a $\text{MVN}\left(0, \sigma_u^2 [(\mathbf{I} - \rho \mathbf{W})(\mathbf{I} - \rho \mathbf{W}')^{-1}]\right)$, and $\mathbf{e} = [e_{11}, e_{12}, \dots, e_{ij}, \dots, e_{mN_m}]'$ from a $N(0, \sigma^2)$. x_{ij} values are generated from a uniform distribution between 0 and 1000 and $N_i = [N_1, \dots, N_m]$ is generated from uniform distribution between 100 and 300. The population size is $N = \sum_{i=1}^m N_i$. Thus, we simulate 42 different populations based on

different values of spatial autoregressive coefficient, $\rho = \{0, \pm 0.25, \pm 0.5, \pm 0.75\}$, and number of areas, $m = \{16, 25, 36, 64, 144, 225\}$. y_{ij} is then produced as a continuous and normally-distributed variable with random area effects of contiguous areas. As a result, area-level aggregates and estimates are continuous, normally distributed and geographically aggregated, as is usually the case of many criminological variables such as confidence in police services, fear of crime or general crime rates at large scales (Fay and Diallo 2012; Williams et al. 2019; Whitworth 2012). Future research should also examine different simulation parameters with smaller intra-class correlations.

All maps used are hypothetical maps based on perfect squares divided into m number of areas, where the maximum number of neighbors is 8 and the minimum is 3 at the corners (see Fig. 1). Future research should conduct similar studies using more realistic maps. Neighboring areas are defined based on a 'Queen Contiguity' matrix, typically the most common structure used in simulation studies, which defines as neighbors all areas that share borders or at least one vertex. The \mathbf{W} matrix is standardised by rows, so that every row adds up to 1.

The simulation consists in the following steps for each simulated population:

1. Selection of $t = 1, \dots, T$ ($T = 1000$) simple random samples without replacement. Sample sizes are drawn with the only constraint of a minimum of two units selected in each area (Salvati 2004). The average sample size per area is $\bar{n} = 48.8$.

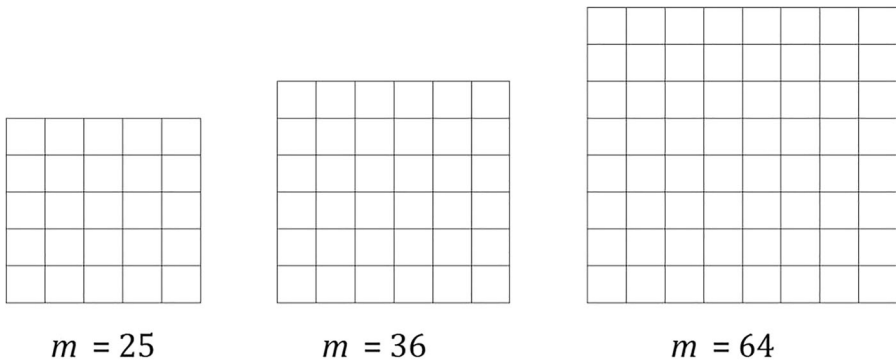


Fig. 1 Three examples of hypothetical maps used in simulation study

- In each sample, post-stratified, EBLUP and SEBLUP estimates are computed and compared based on Pratesi and Salvati (2008). The post-stratified estimator is given by the following:

$$\hat{Y}_i(pst) = \sum_{j \in s_i} \frac{y_{ij}}{n_i}, \quad (10)$$

where s_i is the set of n_i sample units falling in area i .

- The results are evaluated by the absolute relative bias, absolute relative error, relative root mean squared error, and mean squared error averaged through the samples and small areas (Petucci and Salvati 2006). These are denoted by \overline{ARB} , \overline{ARE} , \overline{RRMSE} , and \overline{MSE} , and given by the following formulas, respectively:

$$\overline{ARB} = \frac{1}{m} \sum_i \left| \frac{1}{T} \sum_{t=1}^T \left(\frac{\hat{Y}_{it}}{Y_i} - 1 \right) \right| \quad (11)$$

$$\overline{ARE} = \frac{1}{m} \sum_i \frac{1}{T} \sum_{t=1}^T \left(\left| \frac{Y_{it}}{Y_i} - 1 \right| \right) \quad (12)$$

$$\overline{RRMSE} = \frac{1}{m} \sum_i \frac{\left[\overline{MSE}(\hat{Y}_i)^{1/2} \right]}{Y_i} \quad (13)$$

with

$$\overline{MSE} = \frac{1}{m} \sum_i \frac{1}{T} \sum_{t=1}^T \left(\hat{Y}_{it} - Y_i \right)^2, \quad (14)$$

where \hat{Y}_{it} denotes the estimate (post-stratified, EBLUP or SEBLUP) for small area i in sample t and Y_i the true value observed in the population for area i .

The simulation study has been coded in R software (Molina and Marhuenda 2015) and results are detailed in Tables 1, 2, 3 and 4.

Results: Comparison of EBLUP and SEBLUP Estimates

Table 1 shows the \overline{RRMSE} , \overline{ARB} and \overline{ARE} of post-stratified, EBLUP and SEBLUP estimates from each simulated population. Both EBLUP and SEBLUP estimators outperform post-stratified estimators in all cases, in terms of \overline{RRMSE} and \overline{ARE} , regardless of the spatial correlation parameter and the number of areas under study. The post-stratified estimator performs better in terms of \overline{ARB} , as expected. ρ and m do not affect the EBLUP or SEBLUP's relative difference towards post-stratified estimates regardless of the quality measure selected. The relative difference between post-stratified and SEBLUP estimates' \overline{RRMSE} , which expresses the absolute percentage change of the estimate quality measure, has been calculated as follows:

$$RD\% = \frac{\overline{RRMSE} \left[\hat{\delta}^{SEBLUP} \right] - \overline{RRMSE} \left[\hat{Y}(pst) \right]}{\overline{RRMSE} \left[\hat{Y}(pst) \right]} \times 100 \quad (15)$$

Equation (15) gives the measure of efficiency of $\hat{\delta}^{SEBLUP}$ over $\hat{Y}(pst)$ estimates.

The relative difference between post-stratified and SEBLUP estimates' \overline{RRMSE} varies between a maximum of -5.83% in the case of $m = 64$ and $\rho = 0.75$ and a minimum of -14.29% in the case of $m = 16$ and $\rho = 0$, having also small values such as -13.99% in the case of $m = 25$ and $\rho = 0.25$, -13.40% in the case of $m = 144$ and $\rho = 0$, and -13.00% in the case of $m = 144$ and $\rho = -0.5$. In other words, neither ρ nor m can be used to interpret the increased precision, in terms of \overline{RRMSE} and \overline{ARE} , of EBLUP and SEBLUP estimates when compared to post-stratified estimates. However, both ρ and m have a large impact in the improvement of the SEBLUP estimates, which perform substantially better than EBLUP estimates for those cases with a medium and large spatial correlation parameter (especially $\rho = \{\pm 0.50, \pm 0.75\}$) and a large number of areas (notably $m = \{144, 255\}$) (see Tables 2, 3 and 4).

Table 2 shows the relative difference between EBLUP and SEBLUP estimates' \overline{RRMSE} , as shown in Eq. (15), formatting the cells based on a black-to-white colour scale. Darker scales represent positive values, meaning a better performance of EBLUP estimates with respect to their quality measure, and white scales refer to negative values, which show that SEBLUP estimates improve their quality measure when compared to EBLUP estimates. First, it is clear from Table 2 that

Table 1 Estimates' relative root mean squared error, absolute relative Bias and absolute relative error ($\times 100$)

			$m = 16$	$m = 25$	$m = 36$	$m = 64$	$m = 144$	$m = 225$
$\rho = -0.75$	$\hat{Y}(pst)$	$\overline{RRMSE}\%$	12.91	12.50	14.54	12.61	13.08	13.18
		$\overline{ARB}\%$	0.38	0.32	0.42	0.36	0.33	0.31
		$\overline{ARE}\%$	8.95	8.55	10.09	8.78	9.07	9.15
	$\hat{\delta}^{EBLUP}$	$\overline{RRMSE}\%$	11.99	11.53	14.29	11.30	11.80	11.89
		$\overline{ARB}\%$	2.95	2.50	3.85	2.58	2.80	2.57
		$\overline{ARE}\%$	8.56	8.16	10.08	8.13	8.46	8.51
	$\hat{\delta}^{SEBLUP}$	$\overline{RRMSE}\%$	12.23	11.57	14.25	11.21	11.42	11.51
		$\overline{ARB}\%$	2.99	2.53	3.87	2.57	2.78	2.55
		$\overline{ARE}\%$	8.69	8.19	10.05	8.09	8.25	8.34
$\rho = -0.5$	$\hat{Y}(pst)$	$\overline{RRMSE}\%$	12.32	13.09	12.40	12.99	12.92	13.15
		$\overline{ARB}\%$	0.29	0.31	0.33	0.33	0.36	0.31
		$\overline{ARE}\%$	8.57	9.07	8.57	9.04	8.94	9.12
	$\hat{\delta}^{EBLUP}$	$\overline{RRMSE}\%$	11.31	12.21	11.21	11.72	11.24	11.86
		$\overline{ARB}\%$	2.59	2.40	2.31	2.66	2.65	2.90
		$\overline{ARE}\%$	8.11	8.65	7.99	8.42	8.10	8.51
	$\hat{\delta}^{SEBLUP}$	$\overline{RRMSE}\%$	11.60	12.36	11.25	11.59	11.23	11.71
		$\overline{ARB}\%$	2.66	2.46	2.36	2.65	2.63	2.87
		$\overline{ARE}\%$	8.27	8.74	8.02	8.37	8.07	8.43
$\rho = -0.25$	$\hat{Y}(pst)$	$\overline{RRMSE}\%$	13.11	12.62	12.93	12.61	12.68	13.06
		$\overline{ARB}\%$	0.35	0.31	0.24	0.29	0.29	0.31
		$\overline{ARE}\%$	9.14	8.77	8.92	8.76	8.78	9.03
	$\hat{\delta}^{EBLUP}$	$\overline{RRMSE}\%$	12.35	11.40	11.71	11.49	11.18	11.34
		$\overline{ARB}\%$	2.80	2.35	2.57	2.51	2.56	2.79
		$\overline{ARE}\%$	8.79	8.16	8.35	8.23	8.04	8.18
	$\hat{\delta}^{SEBLUP}$	$\overline{RRMSE}\%$	12.50	11.52	11.71	11.41	11.09	11.25
		$\overline{ARB}\%$	2.86	2.39	2.58	2.51	2.54	2.77
		$\overline{ARE}\%$	8.88	8.22	8.34	8.20	8.02	8.15
$\rho = 0$	$\hat{Y}(pst)$	$\overline{RRMSE}\%$	11.97	12.47	12.77	12.65	12.69	12.99
		$\overline{ARB}\%$	0.36	0.28	0.33	0.36	0.33	0.35
		$\overline{ARE}\%$	8.34	8.65	8.86	8.75	8.79	8.97
	$\hat{\delta}^{EBLUP}$	$\overline{RRMSE}\%$	10.26	10.96	11.52	11.19	10.99	11.47
		$\overline{ARB}\%$	2.38	2.60	2.95	2.61	2.76	2.63
		$\overline{ARE}\%$	7.46	7.93	8.29	8.03	7.95	8.23
	$\hat{\delta}^{SEBLUP}$	$\overline{RRMSE}\%$	10.62	11.08	11.60	11.23	11.03	11.46
		$\overline{ARB}\%$	2.59	2.67	3.00	2.63	2.77	2.62
		$\overline{ARE}\%$	7.70	7.98	8.35	8.06	7.97	8.22
$\rho = 0.25$	$\hat{Y}(pst)$	$\overline{RRMSE}\%$	11.18	11.58	13.84	11.78	12.77	12.92
		$\overline{ARB}\%$	0.27	0.31	0.44	0.25	0.31	0.33
		$\overline{ARE}\%$	7.77	8.04	9.60	8.16	8.84	8.95
	$\hat{\delta}^{EBLUP}$	$\overline{RRMSE}\%$	9.99	9.96	12.39	10.29	11.48	11.32
		$\overline{ARB}\%$	2.26	2.04	3.29	2.44	2.68	2.67
		$\overline{ARE}\%$	7.20	7.20	8.91	7.41	8.22	8.16
	$\hat{\delta}^{SEBLUP}$	$\overline{RRMSE}\%$	10.15	10.12	12.59	10.30	11.45	11.29
		$\overline{ARB}\%$	2.29	2.12	3.35	2.45	2.68	2.66
		$\overline{ARE}\%$	7.29	7.30	9.01	7.41	8.21	8.15
$\rho = 0.5$	$\hat{Y}(pst)$	$\overline{RRMSE}\%$	11.25	15.13	12.92	15.23	12.26	12.97
		$\overline{ARB}\%$	0.23	0.39	0.29	0.37	0.31	0.32

Table 1 (continued)

			$m = 16$	$m = 25$	$m = 36$	$m = 64$	$m = 144$	$m = 225$
$\rho = 0.75$	$\hat{\delta}^{EBLUP}$	$\overline{ARE}\%$	7.76	10.54	8.99	10.53	8.48	8.98
		$\overline{RRMSE}\%$	9.85	13.24	11.81	14.12	10.73	11.50
		$\overline{ARB}\%$	2.23	2.97	2.64	3.03	2.45	2.66
	$\hat{\delta}^{SEBLUP}$	$\overline{ARE}\%$	7.04	9.58	8.48	9.99	7.72	8.27
		$\overline{RRMSE}\%$	10.01	13.36	11.66	13.99	10.63	11.26
		$\overline{ARB}\%$	2.28	3.02	2.68	3.04	2.44	2.65
	$\hat{Y}(pst)$	$\overline{ARE}\%$	7.13	9.65	8.41	9.95	7.67	8.13
		$\overline{RRMSE}\%$	12.81	11.02	13.06	11.15	15.71	15.06
		$\overline{ARB}\%$	0.21	0.27	0.29	0.29	0.34	0.39
	$\hat{\delta}^{EBLUP}$	$\overline{ARE}\%$	8.88	7.65	9.08	7.69	10.88	10.42
		$\overline{RRMSE}\%$	11.81	10.36	11.62	10.50	14.61	13.94
		$\overline{ARB}\%$	2.64	2.11	2.84	1.97	2.95	2.84
	$\hat{\delta}^{SEBLUP}$	$\overline{ARE}\%$	8.41	7.33	8.37	7.39	10.35	9.90
		$\overline{RRMSE}\%$	11.96	10.07	11.33	9.98	13.69	13.02
		$\overline{ARB}\%$	2.66	2.13	2.86	1.98	2.95	2.82
		$\overline{ARE}\%$	8.51	7.19	8.22	7.04	9.86	9.41

SEBLUP estimates outperform EBLUP estimates, in terms of \overline{RRMSE} , when the spatial correlation parameter is large, while EBLUP estimates tend to be more precise than the SEBLUP when ρ is close to 0. The SEBLUP is thus preferred over the EBLUP to examine social issues that spatially cluster together, as is the case of crime rates (Baller et al. 2001) and perceptions about crime and the police (Jackson et al. 2013; Williams et al. 2019). Second, the relative difference between EBLUP and SEBLUP estimates' \overline{RRMSE} shows that the benefit obtained by borrowing strength from neighboring areas is larger as the number of areas increases. For example, for $m = 25$ the relative difference of \overline{RRMSE} shows that SEBLUP estimates are more precise than the EBLUPs only when the spatial correlation parameter is very large ($\rho = 0.75$), while the SEBLUP outperforms the EBLUP in all cases for $m = 255$, even when $\rho = 0$. In other words, the EBLUP is expected to outperform the SEBLUP in studies producing estimates for a small

Table 2 Relative difference between EBLUP and spatial EBLUP's RRMSE ($\times 100$)

	$m = 16$	$m = 25$	$m = 36$	$m = 64$	$m = 144$	$m = 255$
$\rho = -0.75$	2.00	0.35	-0.28	-0.80	-3.22	-3.20
$\rho = -0.5$	2.56	1.23	0.36	-1.11	-0.09	-1.26
$\rho = -0.25$	1.21	1.05	0.00	-0.70	-0.81	-0.79
$\rho = 0$	3.51	1.09	0.69	0.36	0.36	-0.09
$\rho = 0.25$	1.60	1.61	1.61	0.10	-0.26	-0.27
$\rho = 0.5$	1.62	0.91	-1.27	-0.92	-0.93	-2.09
$\rho = 0.75$	1.27	-2.80	-2.50	-4.95	-6.30	-6.60

Table 3 Relative difference between EBLUP and spatial EBLUP's ARB ($\times 100$)

	$m = 16$	$m = 25$	$m = 36$	$m = 64$	$m = 144$	$m = 255$
$\rho = -0.75$	1.36	1.20	0.52	-0.39	-0.71	-0.78
$\rho = -0.5$	2.70	2.50	2.16	-0.38	-0.75	-1.03
$\rho = -0.25$	2.14	1.70	0.39	0.00	-0.78	-0.72
$\rho = 0$	8.82	2.69	1.69	0.77	0.36	-0.38
$\rho = 0.25$	1.33	3.92	1.82	0.41	0.00	-0.37
$\rho = 0.5$	2.24	1.68	1.52	0.33	-0.41	-0.38
$\rho = 0.75$	0.76	0.95	0.70	0.51	0.00	-0.70

number of areas (e.g. estimates of drug use for ten local authorities; Gemmell et al. 2004); while the SEBLUP produces more reliable estimates when the number of areas under study is large (e.g. estimates of perceived disorder for 282 neighborhoods; Buil-Gil et al. 2019). Therefore, both ρ and m need to be taken into account to explain SEBLUP estimates increased precision in terms of \overline{RRMSEs} , and SEBLUP estimates perform better as the number of areas under study increases.

Table 3 shows the relative difference between EBLUP and SEBLUP estimates' \overline{ARB} and Table 4 shows the relative difference between their \overline{ARE} . Looking at Table 3, it is clear that SEBLUP estimates perform better than EBLUPs, in terms of \overline{ARB} , when the number of areas is large (especially $m = \{144, 255\}$), but not in cases of $m = \{16, 25, 36\}$. For $m = 64$, SEBLUP estimates' \overline{ARB} is only improved when $\rho = -0.5$ and $\rho = -0.75$. Again, while the \overline{ARB} of SEBLUP estimates was not improved in any case for $m = \{16, 25, 36\}$, such quality measure shows that SEBLUP estimates outperform EBLUPs, in terms of \overline{ARB} , in all simulations performed for $m = 255$.

Table 4 also shows that both ρ and m have a large impact to improve SEBLUP estimates' precision, now in terms of \overline{ARE} . For example, for $m = 25$ the relative

Table 4 Relative difference between EBLUP and spatial EBLUP's ARE ($\times 100$)

	$m = 16$	$m = 25$	$m = 36$	$m = 64$	$m = 144$	$m = 255$
$\rho = -0.75$	1.52	0.37	-0.30	-0.49	-2.48	-2.00
$\rho = -0.5$	1.97	1.04	0.38	-0.59	-0.37	-0.94
$\rho = -0.25$	1.02	0.74	-0.12	-0.36	-0.25	-0.37
$\rho = 0$	3.22	0.63	0.72	0.37	0.25	-0.12
$\rho = 0.25$	1.25	1.39	1.12	0.00	-0.12	-0.12
$\rho = 0.5$	1.28	0.73	-0.83	-0.40	-0.65	-1.69
$\rho = 0.75$	1.19	-1.91	-1.79	-4.74	-4.73	-4.95

difference between EBLUP and SEBLUP's \overline{ARE} shows that EBLUP estimates outperform SEBLUPs in all cases except for $\rho = 0.75$; while for $m = 144$ such value shows a better precision of SEBLUP estimates except when $\rho = 0$, and for $m = 255$ the SEBLUP estimator produces better estimates than the EBLUP in every single case.

Empirical Evaluation and Application: Confidence in Police Work in London

In this section we assess and apply the SEBLUP in a real case scenario. We produce direct, EBLUP and SEBLUP estimates of confidence in police work at ward level in Greater London from Metropolitan Police Service Public Attitudes Survey (MPSPAS) 2012 data. Such an application provides further evidence about the SEBLUP performance when applied to policing data. Moreover, this application produces a reliable map of the confidence in police work in London and deepens the meso-level explanatory mechanisms of confidence in policing, by which we mean the proportion of citizens who think the police do a good job (Jackson and Bradford 2010; Stanko and Bradford 2009). We then draw the map of the distribution of confidence in policing in London.

There are various reasons why this research has been conducted using London survey data instead of any other city. First, London is one of the few cities with an available local survey designed to measure the confidence in police work. Second, the Greater London Authority website provides information about many auxiliary variables that are relevant for this research and may be used as covariates. Third, London is a well-researched city (Hutt et al. 2018; Jackson et al. 2013; Stanko and Bradford 2009) and thus it is easier to exclude the possibility of drawing spurious associations due to uncontrolled variables. And fourth, during preliminary conversation with Greater London Authority's officers it was acknowledged that this research's potential insights may be of great value for decision-making purposes.

Data and Methods

Data from the MPSPAS 2012 have been used to produce estimates of confidence in police work. MPSPAS is an annual survey conducted by the Metropolitan Police Service since 1983, which records information about perceptions of policing needs, worry about victimization and perceived security and disorder. It consists on a face-to-face questionnaire conducted at the homes of respondents, and it obtains responses from a random probability sample of residents in each of the 32 boroughs in Greater London. Household addresses are selected randomly in each borough, and then the person in each household whose next birthday is closest to the date of the interview is asked to answer the questionnaire. The sample is representative of residents aged 15 or over and it should be large enough to allow analyses at borough level but not at smaller scales. Access to the low level geographies of the MPSPAS was only granted for the 2012 edition, and thus small area estimates of confidence in policing are only produced for this year.

Small area estimates will be produced at the ward level for the five London sub-regions. Each sub-region contains a different number of wards: Central London is composed of 114 wards in six boroughs, North London is composed of 61 wards in three boroughs, South London is composed of 120 wards in six boroughs, East London is composed of 192 wards in ten boroughs, and West London is composed of 140 wards within seven boroughs. The average sample size per borough is $\bar{n} = 401.03$ ($sd = 3.82$) and the average sample size per ward is similar in all sub-regions: in Central London $\bar{n} = 20.23$, in the North $\bar{n} = 19.02$, in the South $\bar{n} = 19.37$, in the East $\bar{n} = 20.6$, and in West London $\bar{n} = 19.44$. On average, there are 19.85 citizens sampled per ward. Note that three wards in Central London and fourteen in East London suffered from zero sample sizes, and thus were not included in our analyses. Regression-based synthetic estimates are used in these seventeen areas.

The variable used to measure confidence in police work has been obtained from the question “Taking everything into account, how good a job do you think the police in this area are doing?”, as suggested by Stanko and Bradford (2009). In order to produce more easily interpretable results, responses were dichotomised to a 0–1 measure, where 1 refers to “Excellent” or “Good”, while “Very poor”, “Poor” and “Fair” responses were recoded as 0. “Don’t know” answers were coded as missing data. We then produce estimates of the proportion of people who think the police are doing a good or excellent job in local area (defined in the survey as the area within about 15 min’ walk from home). Based on the literature review, we fitted EBLUP and SEBLUP models using the following area-level covariates: proportion of black and minority ethnic groups 2011, mean household income 2011–12, crime rate 2011–12, proportion of residents born outside the UK 2011, and proportion of citizens unemployed 2011. All covariates are recorded by the Greater London Authority’s Ward Profiles and Atlas (<https://data.london.gov.uk/dataset/ward-profiles-and-atlas>). We found no available or reliable estimates at the ward level of other covariates explored by previous literature, such as residential instability, perceived disorder and collective efficacy, and thus these are subject of future research.

Direct estimates of the proportion of residents who think that police services do a good or excellent job are produced from the following estimator (Horvitz and Thompson 1952):

$$\hat{Y}_i(dir) = N_i^{-1} \sum_{j \in s_i} w_{ij} y_{ij}, \quad (16)$$

where w_{ij} corresponds to the survey weight of unit j from area i (provided by the original survey), and y_{ij} is the score of unit j from area i . Original survey weights are computed as the proportional distribution by borough of all citizens aged 15 or more across London (derived from Census data) divided by the proportional distribution of the unweighted sample by borough. In order to produce the SEBLUP estimates, a first-order ‘Queen Contiguity’ structure is used to define neighboring areas.

Estimates Reliability Measures

In order to assess the estimates produced in each sub-region, Table 5 shows direct, EBLUP and SEBLUP estimates’ average *RRMSE*, as well as the average Relative

Table 5 Estimates' quality measures

	<i>Central</i>	<i>North</i>	<i>South</i>	<i>East</i>	<i>West</i>	<i>All areas</i>
$\bar{\rho}$	0.74	0.03	0.38	0.06	0.60	0.46
m	111	61	120	178	140	610
$\overline{RRMSE}\%[\hat{Y}(dir)]$	18.31	20.30	17.93	20.64	19.55	19.40
$\overline{RRMSE}\%[\hat{\delta}^{EBLUP}]$	11.05	14.38	11.24	13.44	14.04	12.21
$\overline{RRMSE}\%[\hat{\delta}^{SEBLUP}]$	10.31	14.45	11.18	13.69	13.97	11.91
$\overline{RD}\%[\hat{\delta}^{EBLUP}, \hat{\delta}^{SEBLUP}]$	-7.25	-0.28	-2.05	-1.89	-5.27	-4.76

Difference ($\overline{RD}\%$) between EBLUP and SEBLUP's estimates \overline{RRMSE} . The direct estimates' $RRMSE$ is the Coefficient of Variation (Rao and Molina 2015), while the EBLUP estimates' $RRMSE$ is obtained from Prasad-Rao analytical approximation (Prasad and Rao 1990) and SEBLUPs' $RRMSE$ s have been produced using an analytical approximation as in Molina et al. (2009).

Table 5 shows that direct estimates are the least precise (larger \overline{RRMSE}) in all cases, as expected. SEBLUP estimates are more reliable than EBLUPs, in terms of \overline{RRMSE} , in all six scenarios. The $\overline{RD}\%$ shows that the averaged increased precision of SEBLUP estimates compared to EBLUPs is larger as both the ρ and m increase. First, although ρ is similar in North ($\rho=0.03$) and East London ($\rho=0.06$), the $\overline{RD}\%$ shows better results in the East ($\overline{RD}\%=-1.89$) compared to the North ($\overline{RD}\%=-0.28$) partly due to the larger m in East London ($m=178$). Then, even though the low ρ partly explains the small increased precision of SEBLUP estimates when compared to EBLUPs, the spatial autocorrelation parameter cannot be used on its own to explain why such increased precision is higher in the case of $m=178$ than $m=61$. Second, although m is slightly larger in South London ($m=120$) compared to Central London ($m=111$), the $\overline{RD}\%$ is higher in Central London ($\overline{RD}\%=-7.25$) due to the high spatial autocorrelation parameter ($\rho=0.74$). Finally, the best relative results of the SEBLUP estimator have been obtained in Central London, where both m (111) and ρ (0.74) are large, and West London for the same reason ($m=140$ and $\rho=0.60$). In the case of all areas, m is large (610) and ρ is equal to 0.46, and thus the averaged Relative Difference between EBLUP and SEBLUP's estimates \overline{RRMSE} is quite high ($\overline{RD}\%=-4.76$). These results provide empirical evidence to support the simulation study results: the SEBLUP should be used in those studies producing estimates of geographically concentrated phenomena (Baller et al. 2001) for a large number of areas; while the EBLUP is preferred when producing estimates of non-geographically concentrated phenomena with a small spatial autocorrelation coefficient for a small number of domains.

Table 5 also shows that the level of spatial clustering of the public confidence in police work is much larger in Central and Western London than in the North and East, and there is a medium level of spatial concentration in the South. In other words, while neighboring areas tend to show similar values of confidence in the police in Central and Western London, and thus policing interventions may be planned for groups of areas, in

Table 6 Goodness-of-fit indices of EBLUP and SEBLUP models of confidence in police work

		<i>Central</i>	<i>North</i>	<i>South</i>	<i>East</i>	<i>West</i>	<i>All areas</i>
<i>EBLUP</i>	<i>Log-likelihood</i>	60.54	23.34	58.23	77.46	52.63	266.55
	<i>AIC</i>	-121.43	-35.12	-107.78	-140.79	-98.56	-519.11
	<i>BIC</i>	-98.23	-25.99	-90.32	-117.34	-80.02	-488.22
<i>SEBLUP</i>	<i>Log-likelihood</i>	69.06	27.40	60.56	80.28	56.80	275.58
	<i>AIC</i>	-126.12	-40.81	-109.12	-148.56	-101.61	-535.58
	<i>BIC</i>	-109.86	-30.14	-92.39	-129.47	-83.96	-499.86

the North and East place-based policing strategies should be adjusted to the characteristics and needs of each small area.

Mapping the Confidence in Police Work

Goodness-of-fit indices are analyzed to assess the models used in this application. Log-likelihood, AIC and BIC measures show that the SEBLUP model has a better goodness of fit than the EBLUP, and thus we focus on its results (see Table 6).

Table 7 shows the results of the EBLUP and SEBLUP models fitted to produce estimates of confidence in police work for all London wards. All covariates but the crime rate show significant relations with the confidence in police work (Kwak and McNeeley 2017; Wu et al. 2009). The proportion of citizens unemployed is the most important covariate introduced in our area-level SEBLUP model, followed by the concentration of ethnic minorities and the proportion of immigrants (Dai and Johnson 2009; Kwak and McNeeley 2017; Sampson and Bartusch 1998; Wu et al. 2009). The mean income also shows a significant but smaller positive relation with the confidence in the police.

Table 7 EBLUP and SEBLUP models of confidence in police work (all areas)

	EBLUP				SEBLUP			
	Coeff.	SE	t-value	p value	Coeff.	SE	t-value	p value
(Intercept)	0.615	0.05	10.24	0.000	0.588	0.06	9.32	0.000
Proportion minorities	-0.114	0.08	-1.37	0.049	-0.112	0.09	-1.17	0.048
Mean income	0.001	0.00	3.20	0.001	0.001	0.00	3.09	0.002
Crime rate	-0.001	0.00	-0.69	0.132	-0.001	0.00	-0.99	0.123
Proportion immigrants	-0.027	0.09	-1.13	0.037	-0.031	0.09	-1.12	0.036
Proportion unemployed	-0.317	0.15	-1.90	0.004	-0.293	0.17	-1.70	0.009
AIC			-519.11				-535.17	
BIC			-488.22				-499.86	
Spatial correlation							0.46	

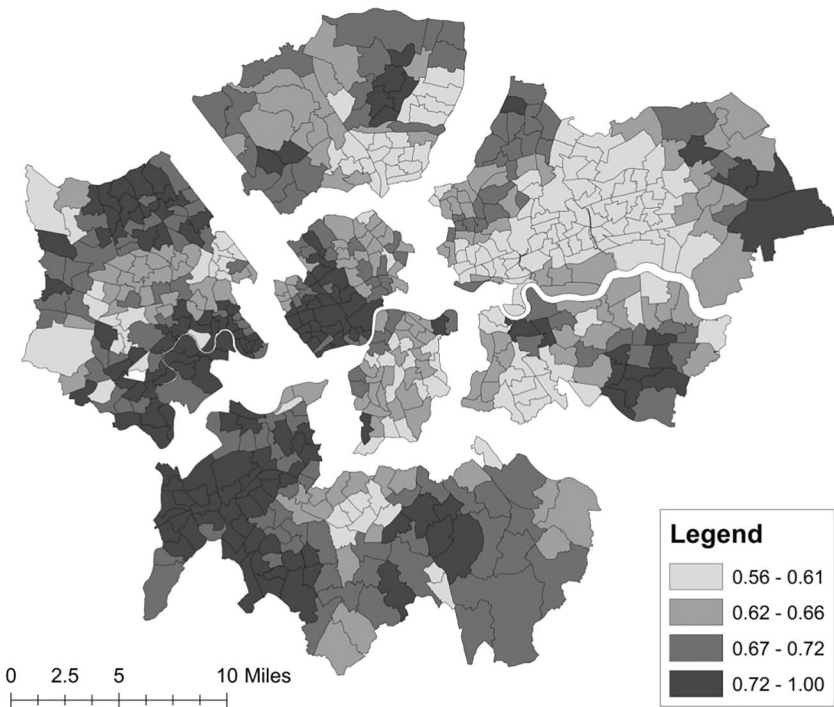


Fig. 2 Proportion of citizens who think the police do a good or excellent job (SEBLUP estimates). Division based on quartiles

Figure 2 shows the geographical distribution of SEBLUP estimates of confidence in police work at ward level in Greater London, where lighter scales of grey indicate a lower proportion of citizens who think that police do a good or excellent job, and darker scales of grey shows higher confidence in police work. The highest estimates of confidence in police work have been found in eight wards located in Central London, six of which are in Kensington and Chelsea (Chelsea Riverside (97.3%), Campden (89.99%), Earl's Court (86.66%), Courtfield (86.28%), Queen's Gate (85.47%) and Brompton and Hans Town (84.62%)) and two in Westminster (Lancaster Gate (88.46%) and Marylebone High Street (88.38%)). There are also high proportions of citizens who think that police do a good job in some western areas of Harrow, Richmond upon Thames and Hammersmith and Fulham. The lowest proportions have been estimated in Alexandra, located in Haringey (43.79%), followed by 27 eastern wards distributed among Lewisham, Newham, Barking and Dagenham, Redbridge, Tower Hamlets, Barking and Dagenham and Greenwich. From a broader perspective, these results add evidence to the estimates produced by the London Mayor's Office for Policing and Crime (<https://maps.london.gov.uk/NCC/>) at a larger geographical scale, which show the highest levels of trust in policing in Central and Southwest London and lower trust in the police in East and North London.

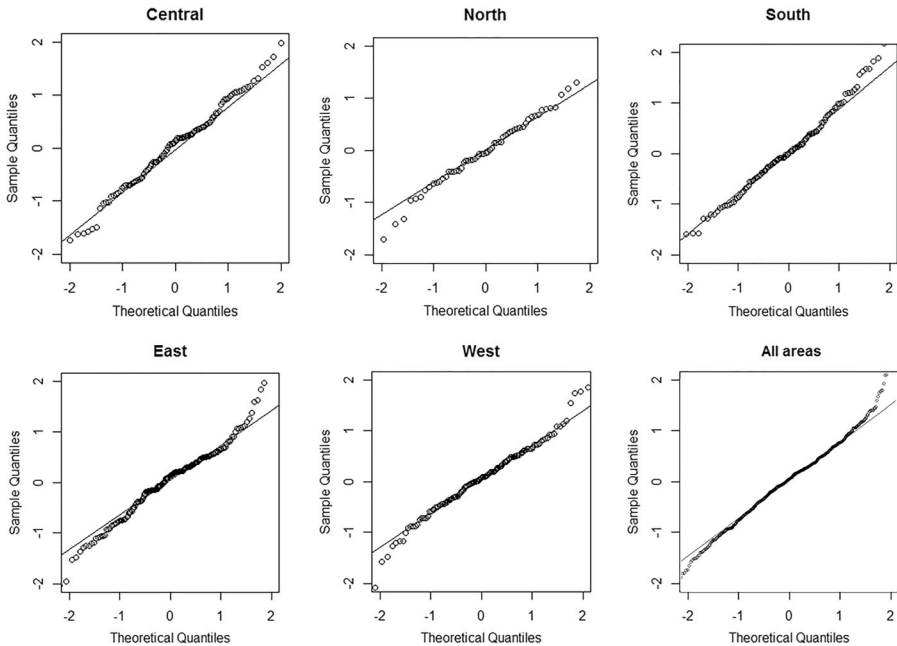


Fig. 3 Normal q - q plots of standardised residuals of SEBLUP estimates

Model Diagnostics

We provide diagnostics of our spatial models by analysing the normality of SEBLUP standard residuals. Residuals are produced as suggested by Petrucci and Salvati (2006:178) and normal q - q plots are shown in Fig. 3. Most residuals show no important deviations. The Shapiro-Wilk test for normality fails to reject the null hypothesis of normal distribution in all five cases: $W=0.984$ and p -value=0.204 in the case of Central London, $W=0.969$ and p -value=0.128 in the model fitted for North London, $W=0.967$ and p -value=0.089 for South, $W=0.939$ and p -value=0.079 in the case of East London, and $W=0.975$ and p -value=0.098 for West London. We also fail to reject the null hypothesis of normal distribution for the model fitted with all areas: $W=0.964$ and p -value=0.121.

Discussion and Conclusions

Place-based policing requires the incorporation of SAE when producing maps of confidence in police work at small geographical levels. By producing reliable small area estimates of confidence in policing, we allow for advanced spatial analyses to explain its distribution and provide precise maps to develop place-based interventions to enhance confidence in police work and reduce crime and disorder. While police records are easily geocoded and mapped, advanced statistical analyses are required to produce reliable estimates of survey-recorded confidence in the police. Small

geographical areas are unplanned domains in most crime surveys, and thus model-based SAE is needed to produce estimates of adequate precision (Rao and Molina 2015). Due to the typically high levels of spatial autocorrelation of confidence in policing, we propose making use of the SEBLUP to increase the reliability of estimates produced from crime surveys. The simulation study and application results allow examining the cases in which the SEBLUP produces better estimates than traditional model-based estimators when applied to policing data. Our estimates of confidence in police work not only have tactical and strategic value to design place-based policing interventions, but they also are important from an accountability point of view: government and auditors' inspections into the police expect that police forces enhance their public confidence and legitimacy (HMICFRS 2017).

We have assessed the SEBLUP performance under different scenarios with unequal number of areas and spatial correlation parameters. Our results show that the SEBLUP tends to outperform the EBLUP not only when ρ moves away from zero and is close to 1 and -1 , but also when m is large. The SEBLUP performs better as the number of areas under study increases, while the EBLUP estimator outperforms the SEBLUP both when $\rho \cong 0$ and m is small. Future work will investigate the SEBLUP using different simulation parameters with smaller intra-class correlations and more complex contiguity matrices, such as second-order 'Queen Contiguity' and distance weighted matrices. Furthermore, future research will examine whether small area estimators that borrow strength from temporal series, such as the Rao and Yu (1994) model, provide more reliable estimates in policing research, since confidence in policing is known to be quite stable over time and thus temporally correlated random effect can be used in this field.

From a substantive perspective, our estimates show that citizens are more confident in policing in most Central and Southwestern London neighborhoods, while estimates show a lower confidence in the police in East and North London. Unlike previous research, our estimates are produced at a ward level and thus allow not only for mapping the distribution of confidence in police work at a large scale, but these also bring to light internal heterogeneity in the levels of confidence at a neighborhood level. In Central London, for example, estimates are significantly higher in the northern part of the River Thames, where Westminster and Kensington and Chelsea are located, than in the Southern part of the river. Although crime rates are higher in the northern part of the river, these do not appear to be as significant as the unemployment rate and concentration of minorities, which are more prominent in the southern part of Central London, to explain the distribution of confidence in policing. Our estimates also allow distinguishing clear differences within West London, where confidence in police is clearly higher in most Hounslow wards than in the majority of Ealing neighborhoods, where unemployment and deprivation is more common. These estimates are useful to develop more accurate explanations of the distribution of confidence in police work and to design place-based policing strategies to increase the public confidence in policing and their cooperation with police services.

The unemployment rates, concentration of minorities and immigrants and average income have shown to be good area-level predictors of the confidence in police work (Bradford et al. 2017; Dai and Johnson 2009; Jackson et al. 2013; Kwak and McNeeley 2017; Sampson and Bartusch 1998; Wu et al. 2009). The two most important covariates (among those included in our models) to explain the geographies of confidence in police work in London are the unemployment rates and concentration of ethnic

minorities. As argued by Sampson and Bartusch (1998:801): “perhaps we should not be surprised that those most exposed to the numbing reality of pervasive segregation and economic subjugation become cynical about human nature and legal systems of justice”. High levels of unemployment and ethnic segregation, as forms of deprivation, might explain that neighbors’ local identities shaped by deprivation are less willing to trust and cooperate with police services (Kwak and McNeeley 2017), but also with other government services (Dai and Johnson 2009). Other researchers argue that this might also be due to an excessive police control and use of force on certain communities with larger concentration of minorities (Dai and Johnson 2009). Open access to Metropolitan Police stop and search data was available only after 2015 and the spatial information about police use of stop and search was available only since mid-2016, and thus we could not include this covariate in our analyses (based on survey data from 2012). However, our area-level estimates of confidence in police work from 2012 show a significant negative Spearman correlation with the proxy measure of stop and search in 2017 (stop and search count: $\rho = -0.22$, $p \text{ value} < 0.01$; stop and search per resident: $\rho = -0.16$, $p \text{ value} < 0.01$). Thus, future research with newer survey data should incorporate this covariate to explore the effect of stop and search on the confidence in police work. Similar mechanisms are used to explain the effect of the concentration of immigrants and average income in the confidence in police work, although these show smaller coefficients in our study. Immigrants and citizens with low income tend to cluster in areas with large levels of concentrated disadvantage -and possibly higher police control and use of force- where social attitudes of distrust towards the police are likely to emerge.

Future research with newer survey data will focus on scoping for other available covariates (e.g. residential instability, collective efficacy, stop and search) to estimate confidence in the police at a ward or smaller spatial levels; and to examine causal mechanisms between economic deprivation, ethnic segregation and confidence in police work (Dai and Johnson 2009). Further research will also replicate similar analyses in other cities and countries with different social and demographic characteristics (and available survey data) to assess the generalizability of the current study’s findings. In addition, new SAE methods are needed that deal with semicontinuous zero-inflated skewed data in policing data (see Dreassi et al. 2014). By expanding the body of research that makes use of SAE techniques in policing research and practice, these methods may become a core tool in survey-based crime analysis and place-based policing.

Acknowledgments The authors thank the Metropolitan Police Service and the Mayor’s Office for Policing and Crime for providing the data used in this research.

Compliance with Ethical Standards

Conflict of Interests The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

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/>.

References

- Asfar, A. K., & Sadik, K. (2016). Optimum spatial weighted in small area estimation. *Global Journal of Pure and Applied Mathematics*, 12(5), 3977–3989.
- Baller, R. D., Anselin, L., Messner, S. F., Deane, G., & Hawkins, D. F. (2001). Structural covariates of U.S. county homicide rates: Incorporating spatial effects. *Criminology*, 39(3), 561–590.
- Bennett, S., Davis, J., & Mazerolle, L. (2014). Police-led interventions to enhance police legitimacy. In G. Bruinsma & D. Weisburd (Eds.), *Encyclopedia of criminology and criminal justice* (p. 37533765). New York: Springer.
- Bradford, B., Sargeant, E., Murphy, K., & Jackson, J. (2017). A leap of faith? Trust in the police among immigrants in England and Wales. *British Journal of Criminology*, 57(2), 381–401.
- Braga, A., Papachristos, A. V., & Hureau, D. M. (2014). The effects of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly*, 31(4), 633–663.
- Buelens, B., & Benschop, T. (2009). Small area estimation of violent crime victim rates in the Netherlands. EUROSTAT. <https://ec.europa.eu/eurostat/documents/1001617/4398369/SIPI-SMALL-AREA-ESTIMATION-BUELENS-BENSCHOP.pdf>. Accessed 27 March 2019.
- Buil-Gil, D., Medina, J., & Shlomo, N. (2019). The geographies of perceived neighbourhood disorder. A small area estimation approach. *Applied Geography*, 109, 102037.
- Chandra, H., Salvati, N., & Chambers, R. L. (2007). Small area estimation for spatially correlated populations – A comparison of direct and indirect model-based methods. *Statistics in Transition*, 8(2), 331–450.
- Cressie, N. (1993). *Statistics for spatial data*. New York: Wiley.
- Dai, M., & Johnson, R. (2009). Is neighborhood context a confounder? Exploring the effects of citizen race and neighborhood context on satisfaction with the police. *Policing: An International Journal of Police Strategies & Management*, 32(4), 595–612.
- Dreassi, E., Petrucci, A., & Rocco, E. (2014). Small area estimation for semicontinuous skewed spatial data: An application to the grape wine production in Tuscany. *Biometrical Journal*, 56(1), 141–156.
- Fay, R., & Diallo, M. (2012). Small area estimation alternatives for the National Crime Victimization Survey. In Proceedings of the survey research methods section, joint statistical meetings, American Statistical Association (pp. 3742–3756). Washington, DC: Westat.
- Fay, R., & Herriot, R. (1979). Estimation of income from small places: An application of James-stein procedures to census data. *Journal of the American Statistical Association*, 74(366), 269–277.
- Gemmell, I., Millar, T., & Hay, G. (2004). Capture-recapture estimates of problem drug use and the use of simulation based confidence intervals in a stratified analysis. *Journal of Epidemiology and Community Health*, 58(9), 758–765.
- Groves, R. M., & Cork, D. L. (Eds.). (2008). *Surveying victims: Options for conducting the National Crime Victimization Survey*. Washington, DC: National Academies Press.
- HMICFRS. (2017). PEEL: Police legitimacy (including leadership) 2017. An inspection of Metropolitan Police Service. HMICFRS. <https://www.justiceinspectors.gov.uk/hmicfrs/wp-content/uploads/peel-police-legitimacy-2017-metropolitan.pdf>. Accessed 21 March 2019.
- Horvitz, D. G., & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(260), 663–685.
- Hutt, O., Bowers, K., Johnson, S., & Davies, T. (2018). Data and evidence challenges facing place-based policing. *Policing: An International Journal*, 41(3), 339–351.
- Jackson, J., & Bradford, B. (2010). What is trust and confidence in the police? *Policing: A Journal of Policy and Practice*, 4(3), 241–248.
- Jackson, J., Bradford, B., Stanko, B., & Hohl, K. (2013). *Just authority? Trust in the police in England and Wales*. Abingdon: Routledge.
- Kongmuang, C. (2006). *Modelling crime: A spatial microsimulation approach*. PhD thesis, University of Leeds.

- Kwak, H., & McNeeley, S. (2017). Neighbourhood characteristics and confidence in the police in the context of South Korea. *Policing and Society*, 29(5), 599–612.
- Molina, I., & Marhuenda, Y. (2015). Sae: An R package for small area estimation. *The R Journal*, 7(1), 81–98.
- Molina, I., Salvati, N., & Pratesi, M. (2009). Bootstrap for estimating the MSE of the spatial EBLUP. *Computational Statistics*, 24, 441–458.
- Petrucchi, A., & Salvati, N. (2006). Small area estimation for spatial correlation in watershed erosion assessment. *Journal of Agricultural, Biological, and Environmental Statistics*, 11(2), 169–182.
- Pierce, G. L., Spaar, S., & Briggs, L. R. (1988). *The character of police work: Strategic and tactical implications*. Boston: Northeastern University.
- Prasad, N. G. N., & Rao, J. N. K. (1990). The estimation of the mean squared error of small-area estimators. *Journal of the American Statistical Association*, 85, 163–171.
- Pratesi, M., & Salvati, N. (2008). Small area estimation: The EBLUP estimator based on spatially correlated random effects. *Statistical Methods and Applications*, 17(1), 113–141.
- Rao, J. N. K., & Molina, I. (2015). Small area estimation. *Second edition*. Hoboken: Wiley.
- Rao, J. N. K., & Yu, M. (1994). Small area estimation by combining time series and cross-sections data. *The Canadian Journal of Statistics*, 22, 511–528.
- Rosenbaum, D. P. (2006). The limits of hot spots policing. In D. Weisburd & A. A. Braga (Eds.), *Police innovation: Contrasting perspectives* (pp. 245–264). New York: Cambridge University Press.
- Salvati, N. (2004). Small area estimation by spatial models: The spatial empirical best linear unbiased predictor (spatial EBLUP). Working Paper 2004/03, Dipartimento di Statistica “Giuseppe Parenti”, Università degli Studi di Firenze.
- Salvati, N., Giusti, C., & Pratesi, M. (2014). The use of spatial information for the estimation of poverty indicators at the small area level. In G. Betti & A. Lemmi (Eds.), *Poverty and social exclusion. New methods of analysis* (pp. 261–282). New York: Wiley.
- Sampson, R. J., & Bartusch, D. J. (1998). Legal cynicism and (subcultural) tolerance of deviance: The neighborhood context of racial difference. *Law & Society Review*, 32, 777–804.
- Sherman, L. W., Gartin, P. R., & Buerger, M. E. (1989). Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, 27(1), 27–55.
- Stanko, E. A., & Bradford, B. (2009). Beyond measuring ‘how good a job’ police are doing: The MPS model of confidence in policing. *Policing: A Journal of Policy and Practice*, 3(4), 322–330.
- Tankebe, J. (2012). Viewing things differently: The dimensions of public perceptions of police legitimacy. *Criminology*, 51(2), 103–135.
- Taylor, J. (2013). *Small area synthetic estimation of perceptions of alcohol and drug-related anti-social behaviour*. PhD thesis, University of Portsmouth.
- Tyler, T. R. (2004). Enhancing police legitimacy. *The Annals of the American Academy of Political and Social Science*, 592(1), 83–99.
- Tyler, T. R., & Bies, R. J. (1990). Beyond formal procedures: The interpersonal context of procedural justice. In J. S. Carroll (Ed.), *Applied social psychology and organizational settings* (pp. 77–98). Hillsdale: Erlbaum.
- Weisburd, D. (2015). The law of crime concentration and the criminology of place. *Criminology*, 53(2), 133–157.
- Weisburd, D. (2018). Hot spots of crime and place-based prevention. *Criminology & Public Policy*, 17(1), 5–25.
- Weisburd, D., Bushway, S., Lum, C., & Yang, S. (2004). Trajectories of crime at place: A longitudinal study of street segments in the city of Seattle. *Criminology*, 42, 283–322.
- Wheeler, A., Silver, J., Worden, R., & Mclean, S. (2017). *Mapping attitudes towards the police at micro places*. <https://doi.org/10.2139/ssrn.3079674>. Accessed 23 March 2019.
- Whitworth, A. (2012). Sustaining evidence-based policing in an era of cuts: Estimating fear of crime at small area level in England. *Crime Prevention and Community Safety*, 14(1), 48–68.
- Williams, D., Haworth, J., Blangiardo, M., & Cheng, T. (2019). A spatiotemporal Bayesian hierarchical approach to investigating patterns of confidence in the police at the neighborhood level. *Geographical Analysis*, 51, 90–110.
- Wu, Y., Sun, I. Y., & Triplett, R. A. (2009). Race, class or neighborhood context: Which matters more in measuring satisfaction with police? *Justice Quarterly*, 26(1), 125–156.