

International Journal of Population Data Science

Journal Website: www.ijpds.org



Swansea University
Prifysgol Abertawe

Describing the geographical linkage between the UK Millennium Cohort Study and crime incidents in England and Wales

Nicolás Libuy^{1,*}, Emla Fitzsimons¹, and David Church¹

Submission History

Submitted:	08/12/2023
Accepted:	23/02/2024
Published:	25/04/2024

¹Centre for Longitudinal Studies,
Social Research Institute,
University College London, UK

²School of Government,
Universidad Adolfo Ibáñez,
Chile

Abstract

Background

Enhancing longitudinal cohort studies by linking routine external data to them is increasingly used to evaluate how local environments impact participants' outcomes (e.g. crime on adolescents' perception of security and victimisation).

Objective

To describe the geographical linkage between the UK Millennium Cohort Study (MCS) and street-level crime incidents reported to the Police in England and Wales, and to estimate crime count and rates around MCS participants' residences.

Methods

Eight years of monthly street-level police data were linked to the residential postcodes of MCS participants living in England and Wales in surveys 5, 6 and 7 to create individual-level variables of neighbourhood crime counts and rates (28,724 surveys and 11,365 individuals). Radial buffers around participants' residences were created at ages 11, 14 and 17. Crime counts and rates were created prior to the month of interview (at 1, 3, 6, 9, and 12 months prior). A homogenisation of crime categories reported in the police data was conducted to evaluate changes over time and areas. Multivariate models were used to study the association between MCS participants' demographic characteristics and derived measures of neighbourhood crime.

Results

While total crime rates and counts around MCS participants remain stable over the period, they hide heterogeneous upward and downward trends in specific sub-categories, with violence and sexual offences showing a larger increase. We observe a negative socioeconomic gradient between household income deciles, recorded at age 11, and subsequent exposure to neighbourhood crime.

Conclusion

Linking routine crime data to longitudinal studies, such as the MCS, which follow children and their families through a critical period of development, can provide a new resource to understand how local crime impacts child and adolescent outcomes.

Keywords

linkage; crime; children; adolescence; England; Wales

Highlights

- Enhancing longitudinal cohort studies by linking routine crime data to them provides a valuable resource to understand how local crime may impact child and adolescent outcomes.
- We linked 100% of productive surveys in England and Wales for three surveys of the UK Millennium Cohort Study (surveys 5, 6 and 7) and created local crime counts and rates around survey participants' residences.
- We observed a negative association between deciles of household income, measured at age 11, and future exposure to neighbourhood crime.
- The linked data created in this study have limitations and advantages, and researchers should carefully consider appropriate statistical methods to alleviate potential biases.

*Corresponding Author:

Email Address: nicolas.libuy@ucl.ac.uk (Nicolás Libuy)

Introduction

Growing up in high-crime neighbourhoods can have detrimental effects on children and adolescents, increasing their risk of developing emotional and behavioural problems [1–4]. Data from the Crime Survey for England and Wales reveals that between 2015 and 2019, approximately 10% to 12% of children aged 10 to 15 experienced one or more crime incidents, with a slight decline to 7% in 2020. This 2020 prevalence equates to roughly 600,000 crimes, with 50% classified as violent offences [5]. Moreover, a 2017 study estimated that about 2.2 million 10- to 17-year-olds in the UK are concerned about crime in their local area [6]. These statistics underscore the importance of investigating the impact of residing in high-crime areas on the well-being of children and adolescents.

Effectively characterising crime incidents in the neighbourhoods of adolescents as they grow up requires tracking children over extended periods, carefully documenting any changes in residence, and having access to reliable and objective measures of crime incidents that can be linked to their locations [7]. High-quality longitudinal surveys are rare but are essential to understanding the role of area-level changes in crime over time, with cross-sectional surveys and administrative records falling short [8, 9]. Even when sensitive longitudinal data are available to characterise and study neighbourhoods' crime through secure Trusted Research Environments (TRE), accessing it can be time-consuming, and researchers typically have limited access to geographical information, given its potential for disclosure. Offering a readily accessible set of derived variables for researchers has the potential to significantly lower the barriers to investigating the consequences of exposure to area-level crime.

In this study, we overcome some of these challenges by creating a flexible dataset of area-level crime suitable for a variety of research purposes. We link three surveys of the UK Millennium Cohort Study (MCS) [10, 11] with Crime Data (CD) containing street-level crime incidents reported to the police in England and Wales [12]. Our objectives are twofold. Firstly, we aim to describe the geographical linkage between the MCS and CD. Secondly, we seek to demonstrate the potential utility of the linked data by estimating crime counts and rates around MCS participants' residences, as well as examining the relationship between household income inequalities and exposure to neighbourhood crime during adolescence.

The derived CD-MCS linked data cover respondents residing in England and Wales during the fifth, sixth, and seventh MCS surveys (ages 11, 14, and 17; 28,724 surveys and 11,365 individuals), conducted in 2012, 2015, and 2018, respectively. This dataset is accessible via the UK Data Service under an End User Licence (EUL) agreement, enabling its use by researchers interested in investigating the impacts of neighbourhood crime on individuals' lives.

Methods

Data sources

The millennium cohort study (MCS)

The MCS is a multipurpose longitudinal cohort study, tracking a nationally representative sample of individuals born in the

UK at the beginning of the twenty-first century. The first survey was conducted from 2000 to 2002, and since then, study participants and their parents have taken part in a total of seven surveys, gathering detailed insights into physical, socio-emotional, cognitive, and behavioural development [10]. Details of the MCS surveys and sampling design are described elsewhere [11]. In essence, the MCS sampling design is a clustered and stratified design, deliberately over-representing areas with high proportions of ethnic minorities in England, as well as residents of areas of high child poverty and residents of Scotland, Wales and Northern Ireland [13]. Our focus is on study participants residing in England and Wales during surveys 5, 6, and 7 because the Crime Data for Northern Ireland and Scotland were unavailable at the time of conducting this research.

Crime data (CD)

The CD is a monthly street-level database available at <https://data.police.uk/data/>, a web platform for policing in England, Wales and Northern Ireland. Maintained by the Single Online Home (SOH) National Digital Team of the National Police Chiefs' Council (NPCC), it encompasses crime incidents reported to various police forces across these nations [12]. Police forces report sensitive crime data, which are subsequently validated and geomasked to mitigate privacy risks for victims while preserving an approximate location of the incidents. Geomasking the actual geographic locations of criminal incidents is crucial for safeguarding victims' privacy while ensuring the accuracy of official statistics. Geomasking (or 'location anonymisation') takes the form of moving or 'snapping' the actual locations of crimes to the nearest road centre-line, commercial hub (e.g. shopping centre) or public place (e.g. park). The CD provide data fields including the name of the reporting police force, the latitude and longitude of the geomasked coordinates of the crime, the corresponding Lower Super Output Area (LSOA), the crime type or category (Table 1), and the outcome associated with the crime (e.g. offender sent to prison, under investigation, status update unavailable, investigation complete, no subject identified, among other possible outcomes).

Access to MCS and CD

We obtained access to the CD through the Open Government Licence, and the analysis of sensitive data was conducted via the UCL Data Safe Haven (DSH), which upholds rigorous security standards. To link the CD and MCS, we requested access to residential postcodes and interview dates of MCS participants.¹ The linkage and geographic manipulation were performed using the *sf* and *dplyr* R packages [14, 15], after initial processing of the raw CD in ArcGIS Pro software.

Linkage method and data restrictions

The CD were linked to MCS using the centroid of residential MCS postcodes along with the month of the interview, employing the *sf* R package [15] and following three steps. Firstly, the coordinates of residential postcodes centroids

¹Access to the data was approved by the CLS Data Access Committee (DAC:112).

Table 1: Crime categories in the crime data

	Crime type	Description
1	Anti-social behaviour	Includes personal, environmental and nuisance anti-social behaviour.
2	Bicycle theft	Includes the taking without consent or theft of a pedal cycle.
3	Burglary	Includes offences where a person enters a house or other building with the intention of stealing.
4	Criminal damage and arson	Includes damage to buildings and vehicles and deliberate damage by fire.
5	Drugs	Includes offences related to possession, supply and production.
6	Other crime	Includes forgery, perjury and other miscellaneous crime.
7	Other theft	Includes theft by an employee, blackmail and making off without payment.
8	Public disorder and weapons*	Includes offences which cause fear, alarm, distress or a possession of a weapon such as a firearm.
9	Possession of weapons*	Includes possession of weapons, such as a firearm or knife.
10	Public order*	Includes offences which cause fear, alarm or distress.
11	Robbery	Includes offences where a person uses force or threat of force to steal.
12	Shoplifting	Includes theft from shops or stalls.
13	Theft from the person	Includes crimes that involve theft directly from the victim (including handbag, wallet, cash, mobile phones) but without the use or threat of physical force.
14	Vehicle crime	Includes theft from or of a vehicle or interference with a vehicle.
16	Violent crime	Includes offences against the person such as common assaults, Grievous Bodily Harm
15	Violence and sexual offences	and sexual offences.

Source: Anti-social behaviour (ASB) Incidents, Crime and Outcomes data from data.police.uk.

Notes: *The CD data includes the category 'Public disorder and weapons' until April 2013, and from May 2013 onwards, it was divided into two categories: 'Public order' and 'Possession of weapons'.

for surveys 5, 6 and 7 were projected onto the UK Ordnance Survey National Grid (OSGB36 datum). Secondly, radial buffers of fixed sizes around all postcode centroids were created. Buffer sizes of 500 and 1,600 metres were chosen as they are commonly employed in the literature on neighbourhood crime [2]. Thirdly, we projected the CD onto the same geographical reference system as MCS postcodes centroids (UK Ordnance Survey National Grid) to ensure geographical objects were comparable. We then counted the number of crime incidents reported in each category falling within a predefined radial buffer and a fixed period before the month of the interview.

Crime counts were calculated over five distinct periods: the month preceding the interview month, and the last 3, 6, 9 and 12 months preceding the interview month. There are several reasons for constructing crime rates using various time windows. Firstly, depending on the specific research question at hand, researchers may deem it appropriate to focus solely on a specific fixed period while disregarding others. For instance, MCS questions about victimisation specifically inquire about participants' experiences over the past year, rendering annual crime rates more suitable than other periods. Secondly, within a given MCS survey, the fieldwork extends for several months (usually 15 months), resulting in changes to the reference period of questions. Seasonal variations in exposure to crime incidents should be meticulously considered as they can impact both the cross-sectional association between area-level crime measures and MCS participants' responses, and significantly, the within-participant association

– i.e. how changes in neighbourhood crime around MCS participants' residences over time correlate with changes in MCS participants' responses.

We generate counts for total crimes and also for the categories listed in Table 1. The fixed periods preceding the interview month, used to calculate crime counts and rates, were selected to optimise the overall availability of CD and the comparability of crime categories over time. The total crime counts and rates derived for MCS 5, 6, and 7 are consistent over time; however, comparability for specific categories and certain fixed periods before the interview month may vary. The comparability of crime counts and rates over time was primarily influenced by the fifth survey. This occurred as the fieldwork coincided with changes in CD categories, as described next.

Table 2 shows changes in crime categories in CD, by MCS fieldwork dates. Apart from total crimes, only the following categories have remained unchanged since January 2011: 'Anti-social behaviours', 'Burglary', 'Robbery', and 'Vehicle crime'. Additional categories were introduced in September 2011 ('Criminal damage and arson', 'Drugs', 'Other theft', and 'Shoplifting') while others emerged in May 2013 ('Bicycle theft' and 'Theft from the person'). Two other categories that present changes are 'Violent crime', which existed from January 2011 to April 2013, and since then it became 'Violence and sexual offences.' Finally, from May 2013 onwards, the category 'Public disorder and weapons' was divided into two categories ('Public order' and 'Possession of weapons'). However, between January and September 2011, both categories were included in the category 'Other crimes',

posing challenges for comparison before September 2011. To address this, we create a unique category from September 2011 onwards, retaining the name 'Public disorder and weapons', by consolidating incidents across these three categories.

Crime rates were calculated by using crime counts as the numerator and an appropriate population estimate as the denominator. Typically, the denominator for crime rates is a population estimate within a geographic area, such as small area population estimates derived from census data. However, no such data exist for the circular buffers created around MCS postcode centroids. To address this, we utilised the annual LSOAs population estimates for the years 2012, 2015, and 2018 [16]. In the UK, LSOAs are constructed from groups of 2011 Census Output Areas (OAs), typically comprising 4 to 6 OAs, and designed to have a population ranging from 1,000 to 3,000 persons. We estimate the total population in a specific circular buffer as a weighted average among those LSOAs intersecting the buffer. To illustrate, suppose each LSOA is a perfect $1 \times 1 \text{ km}^2$ square, forming a grid over the UK territory (e.g. dashed grey lines in Figure 1). In our buffer population estimation, each LSOA's population contributes to the estimated population within the buffer based on the percentage of the LSOA area intersecting with the buffer. For instance, if we create a buffer with a radius of 500 metres, and the buffer's centroid is precisely positioned at the corner of the grid, the four segments of the $1 \times 1 \text{ km}$ squares intersecting the circle contribute equally to the population estimate of the circular buffer. Each segment carries the same weight, equal to one-quarter of the circular buffer's area divided by the area of the LSOA, which is $1 \times 1 \text{ km}^2$ (1,000,000) in this scenario. However, as depicted in Figure 1, LSOA geometries are not perfect squares but irregular polygons.²

Statistical analysis

We study the level of crime around MCS participants' residences as follows. Firstly, to validate our derived crime measures, we compare them with well-established and widely used area-level measures, i.e. the Index of Multiple Deprivation (IMD) domains in England and Wales designed to characterise community safety and victimisation risks. For England, we use the Crime domain, which measures the risk of victimisation [18, 19]. In Wales, we rely on the Community Safety Domain, which assesses crime and fire incidents, as well as people's perceptions of outdoor safety [20, 21]. We refer to these two indices as C-IMD to facilitate the exposition.

We linked the rank and deciles of C-IMD to MCS using the residential LSOA of MCS participants. To compare our derived measure with C-IMD, we can only focus on the sixth MCS survey, or when MCS participants were 14 years old, because C-IMD is available for years that are closest to this survey fieldwork and for both England and Wales, respectively in 2015 and 2014 [18, 21]. The periods of other C-IMD are not well suited for comparing with the fieldwork of MCS surveys 5 and 7. We show separate results for England and Wales

because the methodology between the two countries is not comparable. Focusing on crime rates around MCS participants' residences as a dependent variable, we estimate the association with C-IMD (ranks and deciles) controlling for socioeconomic characteristics of MCS participants and Police Force fixed effects. Standard errors are clustered at LSOA level. This analysis focuses on total crime rates using the last 3 months before the month of interview and the 1,600-meter radial buffer (Figure 2, Figure 1 and Supplementary Table 1 in the Supplementary Appendix).

Secondly, we show descriptive statistics of demographic characteristics at age 11 (Table 3) and total counts and crime rates over time and fixed periods before the month of interview for surveys 5, 6 and 7 (Table 4). We compute 95% confidence intervals using the MCS survey design and the *svy* Stata command that considers the stratified and clustered sampling design. Table 4 includes MCS participants linked to the CD at each specific survey. Estimates of counts and rates were weighted using sampling weights provided by the Centre for Longitudinal Studies [22]. We use Violin plots to visualise the overall level of variability of crime counts and rates (Figure 3) and estimate the 80/20 percentile ratio of crime counts to describe cross-sectional inequalities in exposure to crime. We use crime variables with the 1,600-metre buffer and 3-month period because it allows us to characterise overall changes for the three MCS surveys, as well as understand drivers of total crimes trends by specific comparable categories since September 2011 ('Anti-social behaviours', 'Burglary', 'Criminal damage and arson', 'Drugs', 'Public disorder and weapons', 'Robbery', 'Shoplifting', 'Vehicle crime', and 'Violence and sexual offences'). For comparability, all other categories were aggregated under the category 'Other crime' (Figure 5). We additionally explore how seasonality may affect crime rates experienced by MCS participants by plotting crimes by month of interview (Figure 4).

Finally, to demonstrate the use of our derived crime variables, we study how MCS participants' sociodemographic characteristics are associated with the level of crime where MCS participants lived at ages 14 and 17. For this empirical exercise, we focus on a balanced sample of children who participated in the MCS at ages 11, 14 and 17 ($n = 7,568$). Sociodemographic characteristics include the following variables: decile of household income at age 11, ethnicity, maternal education, employment status of MCS participant's parents, housing tenure, number of siblings, country, and neighbourhood crime rates where MCS participants lived at age 11 (Table 5 and Figure 6). To translate household income into deciles, we use the equivalised weekly net family income variable derived and validated by the Centre for Longitudinal Studies [22]. It is defined as the total net income reported by household's members divided by number of people in the household considering their weight on an equivalised income scale.³ The sample mean was used to impute household income in 0.2% (15 of 7,568) of observations. Missing information on ethnicity (0.9%), housing tenure (1.6%), working status (0.2%), number of

²As a robustness analysis, instead of using LSOA population estimates, we used gridded data created by the UK Environmental Information Data Centre [17], that contains $1 \times 1 \text{ km}$ gridded population estimates for the UK based on Census 2011 and Land Cover Map 2015 input data. Descriptive statistics for crime rates based using $1 \times 1 \text{ km}$ gridded population estimates are not shown in this paper, but available upon request.

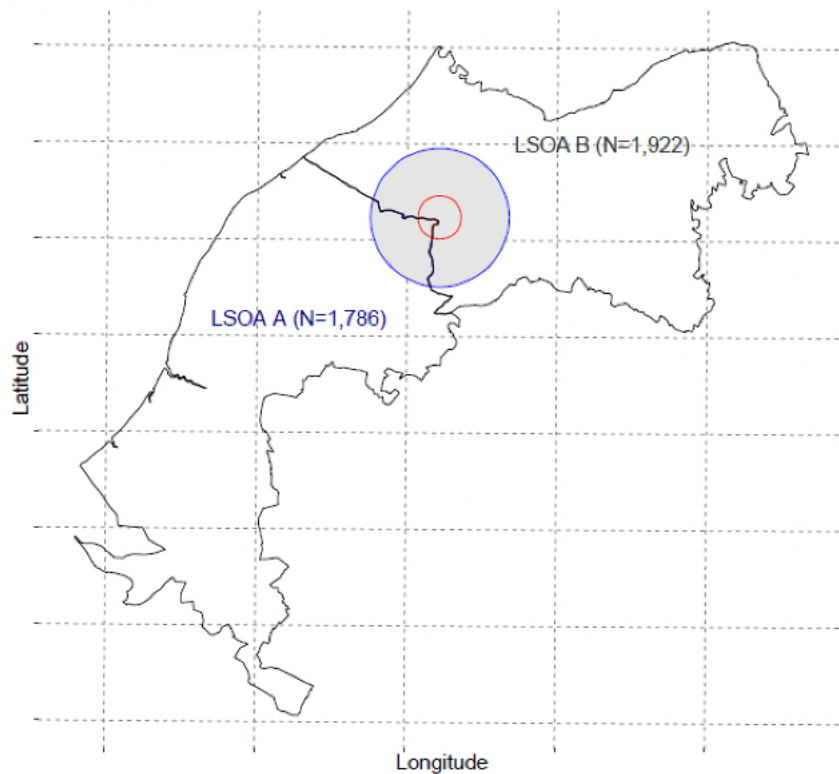
³Equivalence scales are used to adjust the household income by considering the relative needs of a MCS family to those of a couple with no children and are created following the Organisation for Economic Co-operation and Development (OECD) methodology for equivalisation of household incomes [22].

Table 2: Changes in classification of crime incidents and MCS fieldwork

Crime type	2011			2012			2013			2014			2015			2016			2017			2018			2019		
	Jan-Apr	May-Aug	Sep-Dec	Jan-Apr	May-Aug	Sep-Dec	Jan-Apr	May-Aug	Sep-Dec	Jan-Apr	May-Aug	Sep-Dec	Jan-Apr	May-Aug	Sep-Dec	Jan-Apr	May-Aug	Sep-Dec	Jan-Apr	May-Aug	Sep-Dec	Jan-Apr	May-Aug	Sep-Dec	Jan-Apr	May-Aug	Sep-Dec
1 Anti-social behaviour																											
2 Bicycle theft																											
3 Burglary																											
4 Criminal damage and arson																											
5 Drugs																											
6 Other crime																											
7 Other theft																											
8 Public disorder and weapons ^a																											
9 Possession of weapons ^a																											
10 Public order ^a																											
11 Robbery																											
12 Shoplifting																											
13 Theft from the person																											
14 Vehicle crime																											
15 Violent crime ^b																											
16 Violence and sexual offences ^b																											
Millennium Cohort Study Fieldwork				MCS (age 11)									MCS (age 14)									MCS (age 17)					

Notes: Each box represents 4-months windows. Grey boxes indicate the availability of categories in each period. MCS fieldwork of surveys 5, 6, and 7 lasted around 12 months starting in January 2012, 2015, and 2018, respectively. ^a 'Possession of weapons' and 'Public order' were aggregated to facilitate comparisons with category 'Public disorder and weapons'. ^b 'Violent crime' and 'Violence and sexual offences' are combined to facilitate comparisons.

Figure 1: Example of 1,600- (blue line) and 500-meter radial buffer (red line) intersecting with two LSOAs



siblings (0.9%) and maternal education (0.1%) was imputed using the mode of each variable. In Supplementary Table 6 of the Supplementary Appendix, we show results based on a sensitivity analysis using multiple imputation by chained equations that combine 20 imputed datasets.

Results

We successfully linked CD data for all 11,365 individuals residing in England and Wales, who participated in at least one of the three MCS 5-7 surveys. This represents 100% coverage

Table 3: Descriptive statistics at age 11 by quintile of crime rates (1,600-metre buffer and 3-month period, percentages)

	Full sample	Quintile of Crime rates at age 11				
		I	II	III	IV	V
Decile of household income						
1 = lowest income	9.5	2.1	6.7	9.1	13.2	17.9
2	10.7	4.1	7.1	12.0	15.4	16.4
3	10.1	5.6	8.4	10.6	12.6	14.4
4	10.7	7.1	10.5	12.8	10.5	12.9
5	9.7	9.7	11.0	9.0	9.0	9.5
6	9.5	10.7	11.2	10.0	8.1	7.0
7	9.4	11.2	11.4	9.4	7.8	6.6
8	9.7	13.4	11.1	10.2	7.7	5.0
9	10.1	17.3	11.7	8.2	7.0	5.2
10 = highest income	10.6	18.5	10.8	8.8	8.6	5.0
White ethnicity	82.5	93.7	85.5	83.7	75.6	71.7
Housing tenure						
Rent house or other	42.7	27.0	34.9	45.1	51.1	58.7
Own house	57.3	73.0	65.1	54.9	48.9	41.3
Working status						
Both in work	46.0	61.5	51.4	44.0	37.8	31.7
Main in work, partner not	3.4	2.4	3.0	3.1	3.5	5.4
Partner in work, main not	17.3	15.4	16.9	17.3	18.1	19.1
Both not in work	6.2	2.3	3.6	7.5	8.2	10.1
Main in work or on leave, no partner	14.8	13.3	15.1	15.3	14.4	16.2
Main not in work nor on leave, no partner	12.3	5.0	10.0	12.7	17.9	17.4
Maternal Education						
NVQ level 1	9.7	6.8	9.8	11.9	9.6	10.6
NVQ level 2	30.5	27.8	31.8	31.7	31.4	29.9
NVQ level 3	13.0	14.9	15.2	13.6	10.6	10.3
NVQ level 4	24.8	37.9	26.6	21.4	20.1	15.9
NVQ level 5	2.7	4.2	2.2	2.3	2.4	2.1
Overseas qualification	3.1	1.4	2.0	2.9	4.4	5.5
Other	16.2	7.1	12.4	16.2	21.6	25.7
Number of siblings	1.6	1.5	1.5	1.6	1.7	1.9
Country						
England	94.3	94.4	95.0	95.1	93.4	93.6
Wales	5.7	5.6	5.0	4.9	6.6	6.4
Observations	10,638	2,128	2,128	2,127	2,128	2,127

Notes: NVQ = National Vocational Qualifications.

of MCS productive surveys in England and Wales across surveys 5, 6, and 7. Specifically, there were 10,638 participants from MCS survey 5, and 9,457 and 8,629 participants in surveys 6 and 7, respectively.

Figure 2 shows the predicted crime rates for the sixth MCS survey, derived from a regression model incorporating deciles of C-IMD, along with additional controls for deciles of household income, maternal education, ethnicity, housing tenure, and Police Force fixed effects. The figure reveals that MCS participants residing in less deprived LSOAs, as indicated by lower safety and victimisation risks according to C-IMD, experience lower crime rates, measured by incidents reported to the police. The downward trend in Figure 2 is reassuring as it indicates that our derived crime variables capture the between-areas crime differences in neighbourhoods

where MCS participants live. Supplementary Table 1 in the Supplementary Appendix presents OLS estimates, indicating that a one standard deviation increase in C-IMD rank is associated with a reduction of 6.5 points in crime rates in England (p -value < 0.01 ; $R^2 = 0.41$), and 6.1 points in Wales (p -value < 0.01 ; $R^2 = 0.24$). We observe a similar downward trend when employing ranks of C-IMD instead of deciles (Supplementary Figure 1 in the Supplementary Appendix), with more affluent areas (higher ranks) showing lower crime rates. Supplementary Figure 1 in the Supplementary Appendix further underscores the considerable variability of crime rates within similar C-IMD ranks, indicating that linked C-IMD at the LSOA level hide the heterogeneity of local crime rates generated considering the residential location and interview date.

Table 4: Total crime rates and counts by buffer size

	500 metres			1,600 metres		
	Survey 5 (Age:11)	Survey 6 (Age:14)	Survey 7 (Age:17)	Survey 5 (Age:11)	Survey 6 (Age:14)	Survey 7 (Age:17)
Rates						
1 month	12.1 [11.4,12.9]	10.1 [9.5,10.8]	11.9 [11.1,12.7]	9.0 [8.6,9.3]	8.3 [7.9,8.6]	8.8 [8.5,9.2]
3 months	36.7 [34.5,39.0]	31.1 [29.3,33.0]	35.5 [33.2,37.8]	26.8 [25.7,27.8]	24.3 [23.3,25.2]	26.4 [25.3,27.5]
6 months	77.0 [72.2,81.7]	63.4 [59.6,67.2]	70.7 [66.3,75.0]	55.2 [53.0,57.3]	48.1 [46.3,50.0]	53.2 [51.0,55.4]
9 months	122.2 [114.7,129.7]	97.3 [91.3,103.2]	108.0 [101.4,114.7]	86.2 [82.9,89.5]	72.9 [70.1,75.7]	81.1 [77.7,84.5]
12 months	168.9 [158.7,179.2]	131.2 [123.3,139.2]	144.8 [136.0,153.5]	118.3 [113.8,122.7]	97.9 [94.1,101.6]	108.6 [104.1,113.1]
Counts						
1 month	30 [27,32]	25 [23,27]	29 [26,32]	239 [210,267]	222 [198,247]	232 [207,257]
3 months	88 [80,96]	78 [72,84]	87 [79,95]	708 [624,791]	659 [586,732]	695 [620,770]
6 months	182 [166,199]	159 [146,172]	175 [158,192]	1452 [1,283,1,621]	1314 [1,167,1,461]	1405 [1,251,1,559]
9 months	286 [261,311]	243 [223,263]	266 [241,292]	2264 [2,002,2,526]	1987 [1,765,2,209]	2138 [1,903,2,373]
12 months	393 [358,427]	327 [300,355]	356 [322,390]	3101 [2,745,3,457]	2664 [2,367,2,961]	2859 [2,546,3,173]
Total	10,638	9,457	8,629	10,638	9,457	8,629

Notes: Table shows means with 95% confidence intervals estimated using the MCS survey design (in parenthesis). Rates rows show crime rates per 1000 population, and Counts rows indicate the total counts of crime incidents.

Table 3 shows demographic characteristics of MCS participants at age 11 by quintiles of crime rates, with higher quintiles indicating higher crime rates. The data show that families of MCS participants residing in high-crime areas exhibit lower household income, are less likely to have both parents employed, are more likely to be of an ethnic minority group, and to have achieved lower levels of education compared to families residing in areas with fewer crimes.

Table 4 presents estimates of crime counts and rates for the two buffer sizes and five fixed periods across MCS surveys. As previously stated, we focus on the 1,600-metre buffer and a 3-month period preceding the interview. During survey 5 (age 11), around 708 crime incidents were reported to the Police near the participants' residences. However, this figure decreased in the following two surveys to 659 and 695 incidents, respectively. We observed a decrease in crime rates between surveys 5 and 6, where the number of crimes per 1000 population dropped from around 27 to 24. However, there was an increase in crime rates in survey 7, where the number of crimes per 1000 population rose to 26.

Violin plots in Figure 3 show the nonparametric estimation of the empirical distribution of the crime counts and rates by MCS surveys. It shows that while the distribution of crime rates is unimodal around the sample mean, the distribution of crime counts is not, reflecting high inequalities in exposure

to crime incidents between MCS participants. The 80/20 percentile ratio of crime counts for MCS surveys 5, 6 and 7 are 9.6, 9.2 and 10.3 respectively, indicating that MCS participants in the top quintile of the crime counts distribution experience between 6 to 8 times higher rates of crime incidents than those in the bottom quintile of the distribution (Supplementary Table 2 in the Supplementary Appendix).

Figure 4 shows predictive 3-month crime rates by interview month. It also shows the percentage of surveys conducted each month. As the number of surveys decreased in the last four months, the accuracy of the predictions decreased, resulting in wider confidence intervals. Overall, the crime rates for February and March, which reference periods are respectively Nov-Jan and Dec-Jan, were slightly higher than those for May-July.

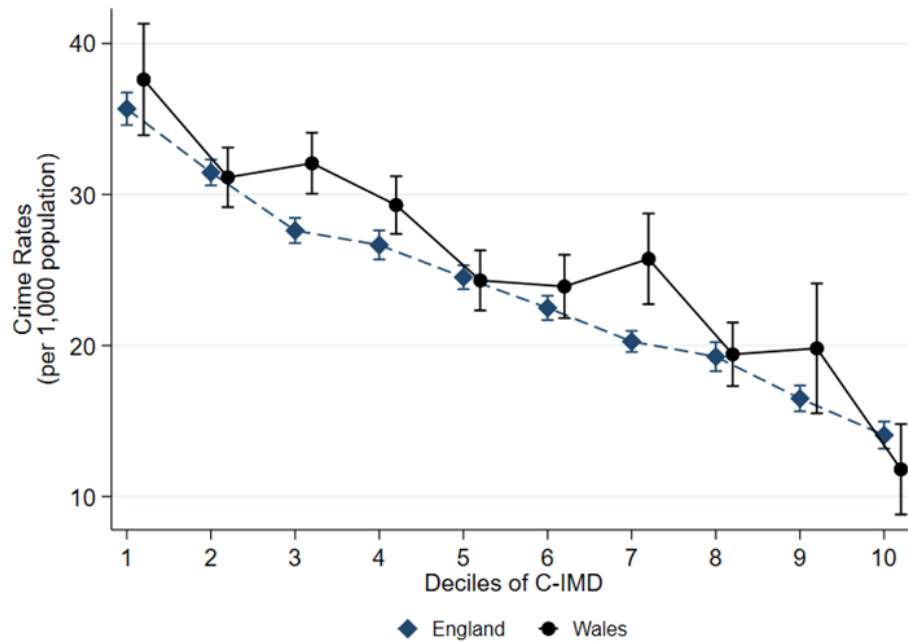
Although the survey-to-survey changes in total crime rates around MCS participants' residences show an increase between survey 5 and subsequent surveys, the underlying composition of crime types shows a much more heterogeneous picture. Figure 5, and Supplementary Table 3 in the Supplementary Appendix, reveal a sharp rise in crime rates and counts of categories 'Violence and sexual offences' and 'Public disorder and weapons'. 'Anti-social behaviour' incidents decrease over time, from around 256 at age 11 to 147 crime incidents at age 17, with a similar pattern observed for crime rates. The

Table 5: OLS estimates for the association between household income deciles (age 11) and crime rates (age 14 and 17)

	Crime rate (age 14)		Crime rate (age 17)	
	Coef.	95% CI	Coef.	95% CI
Decile of household income (age 11)				
1 = lowest income	4.47***	[2.16,6.79]	5.21***	[2.47,7.96]
2	2.89***	[0.92,4.86]	3.56***	[1.12,6.00]
3	2.97***	[1.30,4.64]	3.51***	[1.37,5.64]
4	1.20	[-0.32,2.72]	1.76*	[-0.10,3.62]
5	1.79***	[0.46,3.13]	1.56**	[0.16,2.96]
6	0.50	[-0.61,1.61]	0.36	[-0.94,1.65]
7	0.73	[-0.31,1.77]	0.96	[-0.44,2.35]
8	0.75*	[-0.04,1.54]	0.73	[-0.23,1.68]
9	-0.19	[-0.89,0.50]	-0.41	[-1.16,0.35]
10 = highest income	Ref.		Ref.	
Ethnicity				
White	-0.70	[-1.68,0.29]	-0.12	[-1.09,0.84]
Other	Ref.		Ref.	
Housing tenure				
Rent house or other	Ref.		Ref.	
Own house	0.12	[-0.78,1.01]	-0.65	[-1.72,0.41]
Working status				
Both in work	Ref.		Ref.	
Main in work, partner not	-1.39	[-3.67,0.89]	-2.79***	[-4.86,-0.73]
Partner in work, main not	-0.97***	[-1.62,-0.32]	-0.54	[-1.29,0.22]
Both not in work	-0.41	[-1.71,0.89]	-0.87	[-2.49,0.75]
Main in work or on leave, no partner	-0.72	[-1.69,0.25]	-0.54	[-1.66,0.57]
Main not in work nor on leave, no partner	-1.85***	[-3.11,-0.60]	-2.75***	[-4.32,-1.19]
Maternal Education				
NVQ level 1	-0.02	[-1.22,1.18]	0.97	[-0.62,2.56]
NVQ level 2	0.02	[-0.94,0.98]	0.64	[-0.61,1.89]
NVQ level 3	0.19	[-0.90,1.29]	0.77	[-0.57,2.11]
NVQ level 4	-0.24	[-1.10,0.61]	0.17	[-0.98,1.32]
NVQ level 5	Ref.		Ref.	
Overseas qualification	0.01	[-1.71,1.72]	0.12	[-1.69,1.94]
Other	-0.31	[-1.48,0.87]	0.07	[-1.41,1.56]
Number of siblings	0.04	[-0.22,0.30]	-0.29*	[-0.62,0.04]
Country				
England	Ref.		Ref.	
Wales	0.24	[-3.80,4.28]	2.84	[-4.67,10.35]
Crime Rate (age 11)	0.67***	[0.63,0.72]	0.69***	[0.64,0.75]
Constant	2.92***	[1.09,4.74]	3.96***	[1.51,6.41]
Police Force Fixed Effects	Yes		Yes	
R ²	0.659		0.607	
Observations	7586		7586	

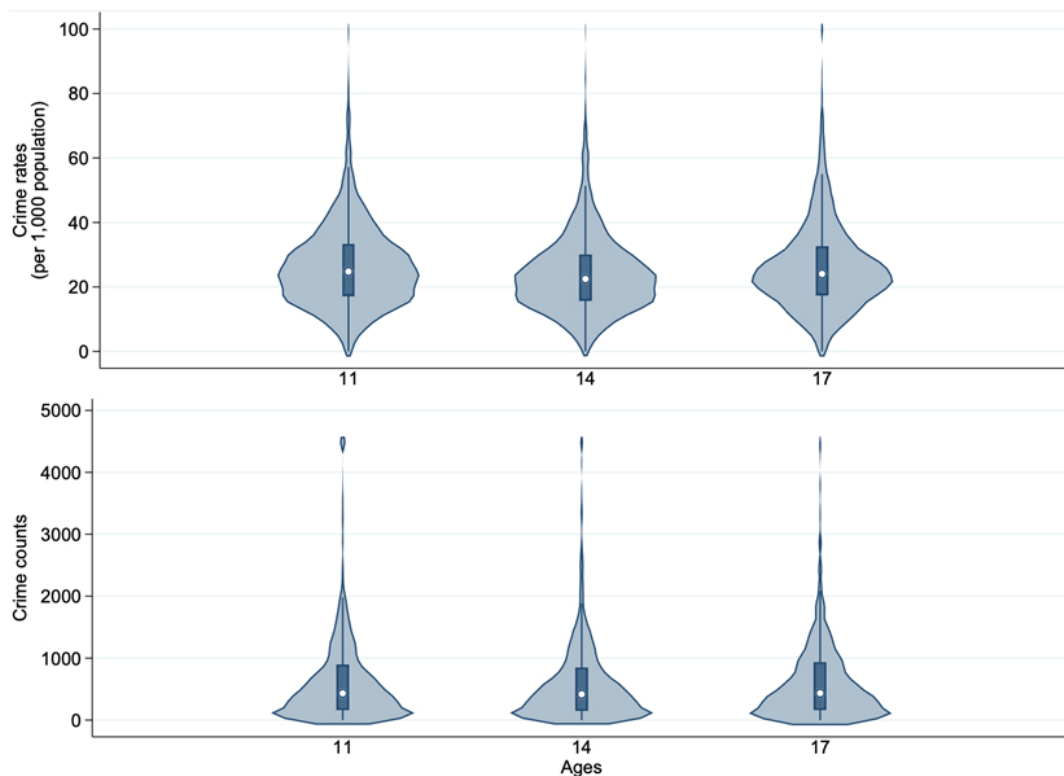
Notes: 95% confidence interval in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are estimated considering the MCS sampling design, and estimates are weighted using MCS derived sampling weights. NVQ = National Vocational Qualifications. Missing income was imputed using the mean of household income at sweep 6. Missing data on ethnicity, housing tenure, working status, and maternal education were imputed using the mode. Binary variables for missing data were included (but not reported) in the table.

Figure 2: Predicted Margins, Crime rates (1,600-metre buffer and 3-month period) by Deciles of C-IMD at LSOA level. MCS survey 6 (age 14), by country



Notes: Figure shows predicted margins of a linear regression of crime rates (1,600-metre buffer and 3-month period) as dependent variable and deciles of C-IMD, adjusted by deciles of household income, maternal education, ethnicity, house tenure, and Police Force fixed effects. Estimated coefficients for deciles of C-IMD are statistically significant at 1% for both England and Wales, and R^2 are 0.40 and 0.26, respectively.

Figure 3: Distribution of crime rates and counts (1,600-metre buffer and 3-month period)



Notes: Violin plots show the median (white circles), interquartile range (dark blue boxes) and the empirical distribution by univariate kernel density estimation (light blue areas).⁴ remaining crime categories show moderate changes over the period.

⁴High crime rates ($n = 43$, 0.15%) were truncated to 100 and crime counts were truncated at 4500 ($n = 260$, 0.9%) to accommodate the scale size in Figure 3.

In term of level of crime incidents, the top five most prevalent categories are 'Anti-social behaviour', 'Violence and sexual offences', 'Criminal damage and arson', 'Burglary' and 'Vehicle crime'. Together they account for 71.2%,

72.6%, 69.8%, of all crime incidents reported to the Police around MCS participants' residences in surveys 5, 6 and 7 respectively.⁵

In Table 5, we show estimates of the association between household income and neighbourhood crime rates, controlling for area-level crime rates at age 11, sociodemographic characteristics and Police Force fixed effects. We find evidence of a socioeconomic gradient between household income, recorded at age 11, and future exposure to neighbourhood crime (Figure 6). We estimate that MCS participants at the bottom of the income distribution during childhood are exposed to higher crime incidents during adolescence (ages 14 and 17) and that this association is slightly larger as they reach 17 years old. As expected, living in neighbourhoods with higher crime rates at the age of 11 is strongly linked to being exposed to local crime at ages 14 and 17. Supplementary Tables 4, 5 and 6 in the Supplementary Appendix show a robustness analysis (train/test split) to mitigate concerns about overfitting.

Regarding the associations between other household characteristics and crime rates at ages 14 and 17, we do not find any evidence that maternal education, housing tenure or White ethnicity are associated with living in neighbourhoods with higher crime rates, once Police Force fixed effects are controlled for. We find some evidence that more stable parental working conditions, i.e. when both parents of MCS participants work, is associated with higher neighbourhood crime rates (Table 5). While interesting, these associations should not be given any causal interpretation.

Discussion and conclusion

In this study, we linked eight years of monthly street-level police data to three surveys of the MCS in England and Wales (2012–2018; ages 11, 14 and 17) to create individual-level exposure variables of neighbourhood crime counts and rates. The linkage methods developed in this study allow us to create highly local crime exposure variables that account for changes in MCS participants' residential locations over time; by linking on the basis of interview date we are also able to obtain the crime measures most closely aligned with the timing of survey responses. The complete availability of postcode centroid of MCS residence allowed us to link 100% of MCS productive surveys in England and Wales, enhancing the already rich individual-level data contained in MCS. The novel data linked to the MCS created as part of this research will be made available to researchers by the Centre for Longitudinal Studies through the UK Data Service (UKDS).⁶

This new data resource opens a range of possibilities for researchers to study how the level of crime in one's area affects the multi-faceted lives of children and adolescents born at the turn of the millennium. There is evidence that exposure to higher rates of crime is detrimental for children

and adolescents, particularly when they grow up, impacting the likelihood of offending and antisocial behaviours [1, 3, 4], child-parent conflict and parental monitoring [23], and mental health [2, 24]. It has also been associated with adolescent victimisation [25], and perceptions of neighbourhood security [26, 27]. Another interesting debate within criminology that can be informed by the linkage of crime data to MCS is around whether actual crime or perceptions of crime matter more regarding adolescents' behaviours [28]. All of these domains are covered in the MCS surveys, and along with the uniquely rich data collected from participants' early lives, provide much potential for understanding how crime affects individuals' lives.

We find that at age 11 (2012), the average crime rate, derived using the 1,600-meters buffer and 3-month period before the interview date, was around 27 per 1,000 population, equating to approximately 708 reported crime incidents. This remained relatively stable across the following surveys, a trend similar to the stagnation in total national crimes reported to the Police over the same period [29]. However, in any given survey, we observe high variation across areas in exposure to crime incidents and crime rates, with 80/20 percentiles ratios that reflect the well-documented disparities in crime exposure for individuals living in more deprived areas [30, 31].

Against this overall stable trend, we provide evidence that MCS participants have experienced a sharp increase in area-level exposure to violence and sexual offences over the eight-year period we consider. Whether the increase in reported incidents of violence and sexual offences to the Police reflects changes in reporting attitudes over time or an increase in actual crime incidents is an area that warrants further research. Understanding the relationship between increased exposure to violence and sexual offences and MCS participants' reports of mental health and wellbeing, alongside other domains of their lives, is also an important topic for future study.

In contrast to the figures pertaining to violence and sexual offences, we documented a decrease in the reported incidents of anti-social behaviour in the vicinity of MCS participant residences during the analysed period spanning from 2012 to 2018. This trend is consistent with the official statistics of anti-social behaviour incidents reported to the Police in England and Wales. According to Office for National Statistics, the number of incidents from the years ending March 2012 to March 2018 has decreased by 12 points, from 41 to 28 incidents per 1,000 population [32]. However, it is important for researchers to be careful when interpreting these numbers as they exclude anti-social behaviour incidents reported to other agencies. Additionally, these data are not designated as National Statistics and may not necessarily indicate the true extent of victimisation in a particular location [33].

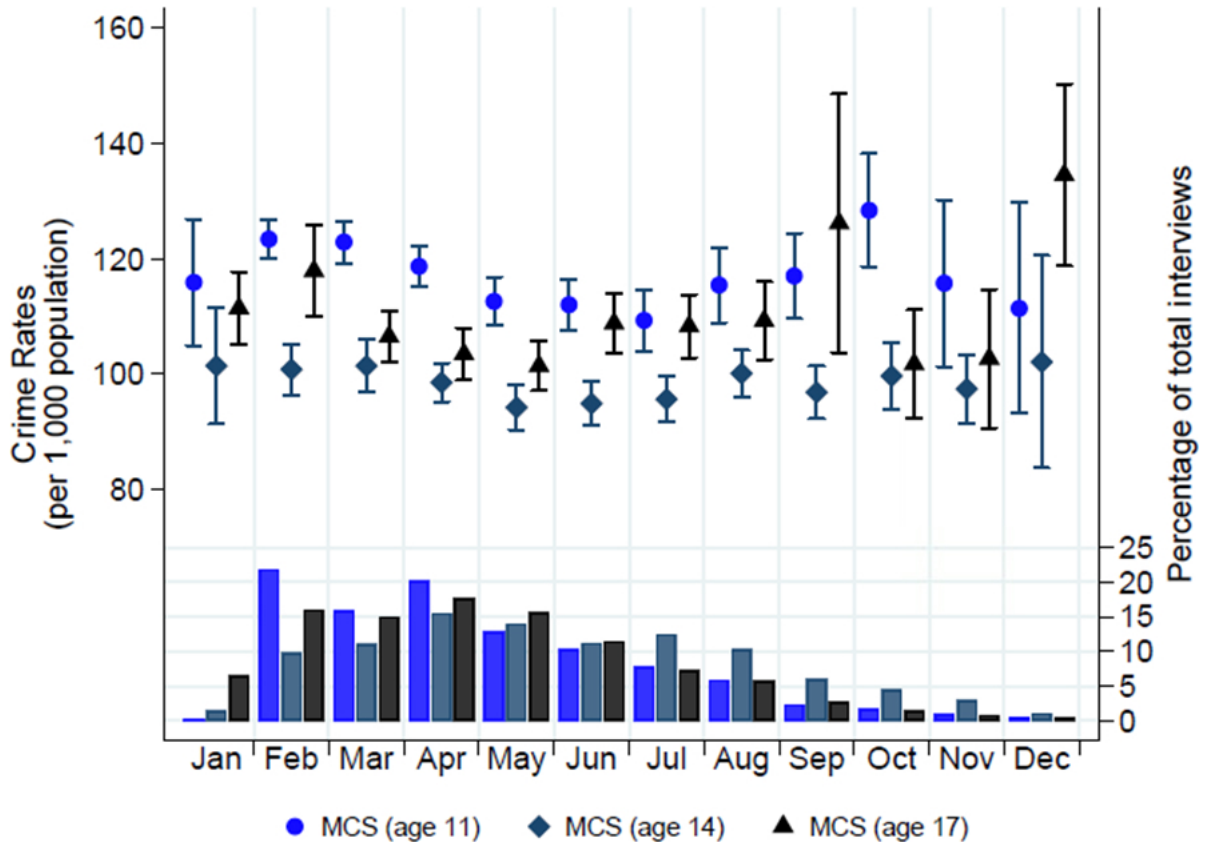
The downward gradient between deciles of household income at age 11 and exposure to area-level crime at ages 14 and 17 highlights the potential of our approach to improve our understanding of the impact of longitudinal changes in crime exposure during childhood and adolescence. Our multivariate analysis aims to showcase the potential utility of derived crime variables and results should not be interpreted as causal. However, it provides new evidence similar to previous research on inequalities in area-level crime exposure [30, 31, 34–36].

While this paper advances previous linkages between local area crime exposure and longitudinal data, we acknowledge some limitations that should be carefully considered when

⁵For each sweep and all categories in Figure 5, we sum all crime incidents reported to the police around MCS participants' residences, allowing duplicates of incidents reported near MCS participants who live nearby, or treating each MCS participant as a separate observation. We then calculate the percentage contribution of each category to the total crime incidents.

⁶More details on the Access Policies of linked MCS data can be found in <https://cls.ucl.ac.uk/data-access-training/data-access/>

Figure 4: Predictive Margins of crime rates (1,600-metre buffer and 3-month period) by month of interview



Notes: Predicted crime rates (left y-axis) were calculated separately by MCS survey. For each survey, we estimate a linear model with crime rates as dependent variable and month of interview, year of interview and Police Force fixed effects. The percentage of total interview (right y-axis).

Figure 5: Crime rates and counts by crime categories (1,600-metre buffer and 3-month period)

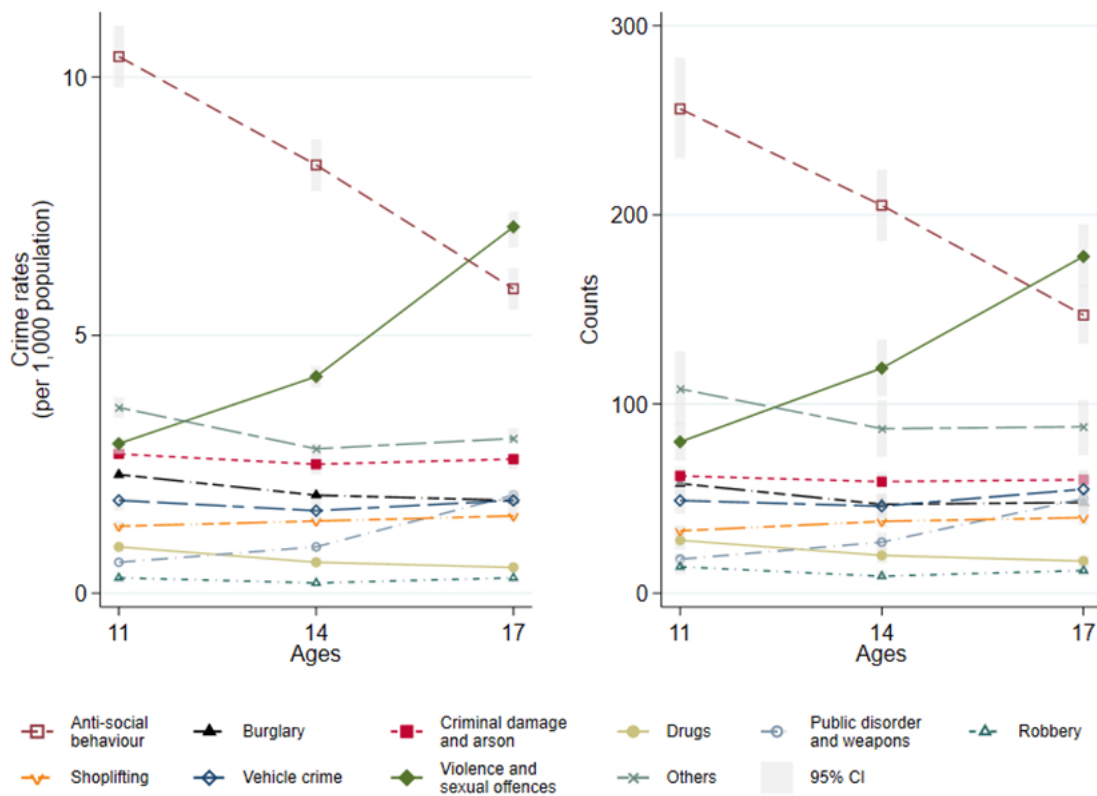
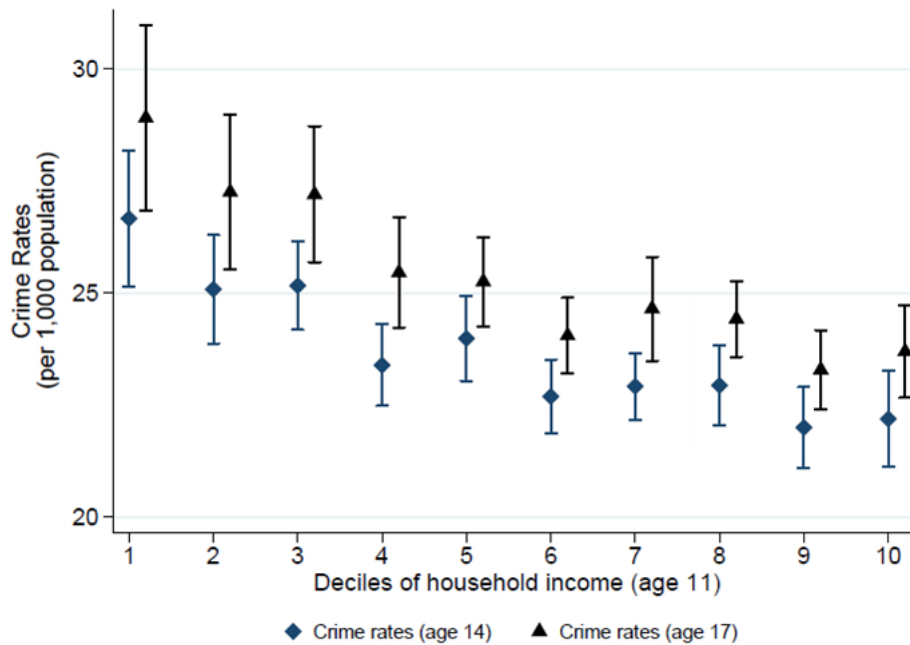


Figure 6: Predictive margins of crime rates at age 14 and 17



Notes: Figure shows the predictive margins for crime rates (1,600-metre buffer and 3-month period) based on OLS estimates reported in Table 5.

using similar methods and derived area-level exposure variables. First, geomasking the location of crime incidents in the CD may introduce measurement errors in the derived crime variables. CD geomasks crime incidents' location by replacing the reported coordinates with the nearest anonymous map point, chosen within a list of street-level points based on the 2012 Ordnance Survey population and housing developments [12]. We consider this an important limitation that researchers should address using adequate methods. That said, its effects on the MCS-linked data are mitigated by the fact that the majority of MCS participants live close to street segments. Potential misclassification can be attenuated by considering buffers of larger sizes, e.g. 1,600m instead of 500m, though at the expense of reduced precision in estimates of crime counts and population. Researchers can use methods to appropriately account for spatial measurement errors due to geolocation errors, aggregations of outcomes and covariates, among others [37–39].

Another limitation arises from the inconsistencies in geocoding policies across various police forces, which may lead to variations in the quality of crime incident records. One approach to help overcome this limitation in multivariate analysis is to control for police force fixed effects or time-varying characteristics of police forces, which should to some extent absorb any observed and unobserved differences between police forces catchment areas.

A third limitation of area-level crime estimates based on Police data is that they omit crime incidents unknown to the Police, i.e. the *dark figure* of crime. As the likelihood of reporting crimes to the police varies geographically, estimates of any area-level crime exposure based only on incidents reported to the police may contain some measurement error. Research for England and Wales shows that the *dark figure* of crime is larger in small urban districts; wealthy and deprived areas; areas with higher immigrant populations; areas where

house prices are lower; and areas with a higher proportion of ethnic minorities, i.e. Black and other ethnicities (excluding Whites and Asians) [40]. In analyses based on multivariate linear models with area-level crime variables as covariates, the presence of spatial measurement error attenuates regression estimates towards the null. Bias-corrected estimates have been developed to attenuate these biases and should be considered by researchers [41, 42].

A fourth limitation that could bias the analysis of changes in area-level exposure to crime using MCS-linked data is attrition. If specific sub-groups, for example, more at-risk populations, disproportionately leave the study, crime rates based on only remaining productive surveys may underestimate the true exposure to area-level crime risks. Future research using MCS-linked data should carefully consider using multiple-imputation to address this limitation [43, 44].

While our article aims to characterise how crime counts and rates have evolved around MCS participants' residences between 2012 and 2018, we do not explore the underlying determinants of area-level crime, which are complex and multidimensional. Several factors have been hypothesised to determine the rise in area-level crimes, such as socioeconomic factors [45], gentrification processes [46], the built environment [47], and policing resources and law-enforcement policies [45, 48]. In the UK, for example, austerity policies that lead to a reduction of police forces have been associated with an increase in crimes [48–50]. Researchers should interpret with caution police recorded data as it may reflect changes in police forces activity due to austerity policies in the UK. Future research with the MCS linked to CD should consider these potential confounders.

Children and adolescents worldwide living in high-crime areas experience daily victimisation risks. Longitudinal studies like the MCS, which track children and their families over a

crucial period of development, offer a distinctive opportunity to comprehend the exposure of children and adolescents to local crime as they progress through life. This study's methods, data and findings provide a first step towards unlocking this potential. We encourage researchers to use the data created as part of this study to explore further the potential of linking survey longitudinal data with routinely collected incidents reported to the police. Further analyses could explore the validation of crime index such as the Cambridge Crime Harm Index, which weights each crime type according to how harmful it is relative to all other crimes [51]. In future research with MCS, it would be beneficial to explore additional measures of area-level crime exposure that consider the interdependence between geographical features (i.e. spatial contagion) or the decline of such features as we move further away from urban centres (i.e. spatial decay).

Conflict of Interest Statement

The Authors have no conflict of interest to declare.

Acknowledgments

We are grateful to the Centre for Longitudinal Studies, Social Research Institute, UCL Institute of Education for the use of the MCS data. The MCS is supported by the Economic and Social Research Council (ESRC) and a consortium of government departments. This research would not have been possible without the important contributions of the MCS cohort members and their families. We acknowledge support from the ESRC (ES/W013142/1, ES/W001179/1).

Ethics statement

The UK Millennium Cohort Study has received ethical approval from the National Health Service (NHS) Research Ethics Committee (REC) system. Ethical approval has been sought for all MCS surveys since the start of the study in 1999.

Data availability statement

The UK Millennium Cohort Study data and the MCS data linked to CD are available at UK Data Service (SN: 2000031, GN 33557 Linked Geographical Data, <http://doi.org/10.5255/UKDA-Series-2000031>).

References

1. Vogel M, Hoeben EM, Bernasco W. Nearby neighbourhood influences on adolescent offending. *British journal of criminology*. 2021;61(1):228–50. <https://doi.org/10.1093/bjc/azaa069>
2. Baranyi G, Di Marco MH, Russ TC, Dibben C, Pearce J. The impact of neighbourhood crime on mental health: A systematic review and meta-analysis. *Social Science & Medicine*. 2021;282:114106. <https://doi.org/10.1016/j.socscimed.2021.114106>
3. Ingoldsby EM, Shaw DS. Neighborhood contextual factors and early-starting antisocial pathways. *Clinical Child and Family Psychology Review*. 2002;5(1):21-55. <https://doi.org/10.1023/A:1014521724498>
4. Damm AP, Dustmann C. Does growing up in a high crime neighborhood affect youth criminal behavior? *American Economic Review*. 2014;104(6):1806-32. <https://www.aeaweb.org/articles?id=10.1257/aer.104.6.1806>
5. Youth Justice Board (YJB). Youth justice statistics 2019/20, england and wales. Youth Justice Board (YJB), Ministry of Justice; 2021. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/956621/youth-justice-statistics-2019-2020.pdf.
6. Pople L, Rees G. Good childhood report 2017. The Children's Society; 2017. Available from: <https://www.childrenssociety.org.uk/the-good-childhood-report-2017>.
7. Shaw DS, Gross HE. What we have learned about early childhood and the development of delinquency. In: Liberman AM, editor. *The long view of crime: A synthesis of longitudinal research*. New York, NY: Springer New York; 2008. p. 79-127. https://doi.org/10.1007/978-0-387-71165-2_3
8. Liberman AM. Synthesizing recent longitudinal findings. In: Liberman AM, editor. *The long view of crime: A synthesis of longitudinal research*. New York, NY: Springer New York; 2008. p. 3-20. https://doi.org/10.1007/978-0-387-71165-2_1
9. Krohn MD, Thornberry TP. Longitudinal perspectives on adolescent street gangs. In: Liberman AM, editor. *The long view of crime: A synthesis of longitudinal research*. New York, NY: Springer New York; 2008. p. 128-60. https://doi.org/10.1007/978-0-387-71165-2_4
10. Fitzsimons E, Haselden L, Smith K, Gilbert E, Calderwood L, Agalioti-Sgompou V, et al. Millennium cohort study: Age 17 sweep (mcs7) user guide. UCL Centre for Longitudinal Studies.; 2020. Available from: <https://cls.ucl.ac.uk/wp-content/uploads/2022/05/MCS7-user-guide-Age-17-ed2.pdf>.
11. Connelly R, Platt L. Cohort profile: Uk millennium cohort study (mcs). *International Journal of Epidemiology*. 2014;43(6):1719-25. <https://doi.org/10.1093/ije/dyu001>
12. Police UK. Data police uk: Single Online Home National Digital Team; 2021 [Available from: <https://data.police.uk/>].
13. Plewis I. The millennium cohort study: Technical report on sampling, 4th edition UK Data Service: Centre for Longitudinal Studies 2007. Available from: https://doc.ukdataservice.ac.uk/doc/4683/mrdoc/pdf/mcs_technical_report_on_sampling_4th_edition.pdf.

14. Wickham H, François R, Henry L, Müller K. Dplyr: A grammar of data manipulation. 2021. Available from: <https://github.com/tidyverse/dplyr>.
15. Pebesma E. Simple features for r: Standardized support for spatial vector data. *The R Journal*. 2018;10(1):439–46. <https://doi.org/10.32614/RJ-2018-009>
16. Office for National Statistics (ONS). Lower layer super output area population estimates (supporting information) Office for National Statistics 2022 [Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/lowersuperoutputareamidyearpopulationestimates>.
17. Reis S, Liska T, Steinle S, Carnell E, Leaver D, Roberts E, et al. Uk gridded population 2011 based on census 2011 and land cover map 2015 [Internet]. NERC Environmental Information Data Centre. 2017. Available from: <https://doi.org/10.5285/0995e94d-6d42-40c1-8ed4-5090d82471e1>.
18. Baljit G. English indices of deprivation 2015 [Internet]. 2015. Available from: <https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015>.
19. Department for Communities and Local Government The english indices of deprivation 2015- statistical release [Internet]. 2015 [cited Date Accessed]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/465791/English_Indices_of_Deprivation_2015_-_Statistical_Release.pdf.
20. Welsh Government. Welsh index of multiple deprivation (wimd) 2014 - report [Internet]. 2014 [cited Date Accessed]. Available from: <https://www.gov.wales/sites/default/files/statistics-and-research/2019-04/welsh-index-of-multiple-deprivation-2014-revised.pdf>.
21. Welsh Government. Welsh index of multiple deprivation - 2014 [Internet]. 2014. Available from: <https://stats.wales.gov.wales/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation/Archive/WIMD-2014>.
22. CLS. Millennium cohort study, user guide (surveys 1 – 5). Centre for Longitudinal Studies, University College London; 2020. Available from: https://doc.ukdataservice.ac.uk/doc/5795/mrdoc/pdf/mcs1-5_user_guide_ed9_2020-08-07.pdf.
23. Zuberi A. Neighborhoods and parenting: Assessing the influence of neighborhood quality on the parental monitoring of youth. *Youth & society*. 2016;48(5):599–627. <https://doi.org/10.1177/0044118X13502365>
24. Baranyi G, Cherrie M, Curtis SE, Dibben C, Pearce J. Changing levels of local crime and mental health: A natural experiment using self-reported and service use data in scotland. *Journal of Epidemiology and Community Health*. 2020;74(10):806. <http://jech.bmj.com/content/74/10/806.abstract>
25. Dong B, Morrison CN, Branas CC, Richmond TS, Wiebe DJ. As violence unfolds: A space–time study of situational triggers of violent victimization among urban youth. *Journal of Quantitative Criminology*. 2020;36(1):119–52. <https://doi.org/10.1007/s10940-019-09419-8>
26. Lorenc T, Clayton S, Neary D, Whitehead M, Petticrew M, Thomson H, et al. Crime, fear of crime, environment, and mental health and wellbeing: Mapping review of theories and causal pathways. *Health and Place*. 2012;18(4):757–65. <https://doi.org/10.1016/j.healthplace.2012.04.001>
27. Mason P, Kearns A, Livingston M. "Safe going": The influence of crime rates and perceived crime and safety on walking in deprived neighbourhoods. *Social Science and Medicine*. 2013;91:15–24. <https://doi.org/10.1016/j.socscimed.2013.04.011>
28. Esteban-Cornejo I, Carlson JA, Conway TL, Cain KL, Saelens BE, Frank LD, et al. Parental and adolescent perceptions of neighborhood safety related to adolescents' physical activity in their neighborhood. *Research quarterly for exercise and sport*. 2016;87(2):191–9. <https://doi.org/10.1080/02701367.2016.1153779>
29. Home Office. Police recorded crime and outcomes open data tables GOV.UK2022 [Available from: <https://www.gov.uk/government/statistics/police-recorded-crime-open-data-tables>.
30. Whitworth A. Inequality and crime across england: A multilevel modelling approach. *Social Policy and Society*. 2012;11(1):27–40. <https://doi.org/10.1017/S1474746411000388>
31. Rennó Santos M, Testa A, Weiss DB. Where poverty matters: Examining the cross-national relationship between economic deprivation and homicide. *The British Journal of Criminology*. 2018;58(2):372–93. <https://doi.org/10.1093/bjc/azx013>
32. Office for National Statistics (ONS). Crime in england & wales, year ending march 2018 - pfa tables (table p7) [Internet]. 2023 [cited Date Accessed]. Available from: <https://www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/crimeandjustice/datasets/policeforceareadatatables/yearendingmarch2018/policeforceareatablesyearendingmarch2018v2.xlsx>.
33. Office for National Statistics (ONS). User guide to crime statistics for england and wales: March 2020 [Internet]. 2021 [cited Date Accessed]. Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/methodologies/userguide-to-crime-statistics-for-england-and-wales#appendix-1-recorded-crime-list>.
34. Kim B, Seo C, Hong Y-O. A systematic review and meta-analysis of income inequality and crime in europe:

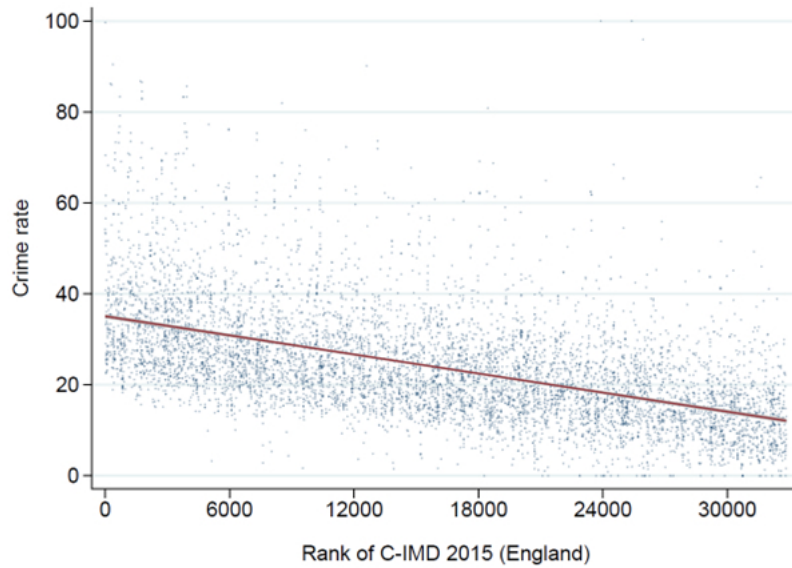
- Do places matter? European Journal on Criminal Policy and Research. 2020. <https://doi.org/10.1007/s10610-020-09450-7>
35. Kelly M. Inequality and crime. The Review of Economics and Statistics. 2000;82(4):530-9. <https://doi.org/10.1162/003465300559028>
36. Lappi-Seppälä T, Lehti M. Cross-comparative perspectives on global homicide trends. Crime and Justice. 2014;43:135-230. <https://doi.org/10.1086/677979>
37. Schennach SM, Starck V. Using spatial modeling to address covariate measurement error. Cemmap working paper. 2020. <https://doi.org/10.47004/wp.cem.2020.5820>
38. Szpiro AA, Sheppard L, Lumley T. Efficient measurement error correction with spatially misaligned data. Biostatistics (Oxford, England). 2011;12(4):610-23. <https://doi.org/10.1093/biostatistics/kxq083>
39. Cressie NAC. Statistics for spatial data, revised edition: Wiley; 2015. <https://doi.org/10.1002/9781119115151>
40. Buil-Gil D, Medina J, Shlomo N. Measuring the dark figure of crime in geographic areas: Small area estimation from the crime survey for england and wales. The British Journal of Criminology. 2021;61(2):364-88. <https://doi.org/10.1093/bjc/azaa067>
41. Huque MH, Bondell HD, Carroll RJ, Ryan LM. Spatial regression with covariate measurement error: A semiparametric approach. Biometrics. 2016;72(3):678-86. <https://doi.org/10.1111/biom.12474>
42. Huque MH, Bondell HD, Ryan L. On the impact of covariate measurement error on spatial regression modelling. Environmetrics (London, Ont). 2014;25(8):560-70. <https://doi.org/10.1002/env.2305>
43. Huque MH, Carlin JB, Simpson JA, Lee KJ. A comparison of multiple imputation methods for missing data in longitudinal studies. BMC Medical Research Methodology. 2018;18(1):168. <https://doi.org/10.1186/s12874-018-0615-6>
44. Spratt M, Carpenter J, Sterne JAC, Carlin JB, Heron J, Henderson J, et al. Strategies for multiple imputation in longitudinal studies. American Journal of Epidemiology. 2010;172(4):478-87. <https://doi.org/10.1093/aje/kwq137>
45. Han L, Bandyopadhyay S, Bhattacharya S. Determinants of violent and property crimes in england and wales: A panel data analysis. Applied Economics. 2013;45(34):4820-30. <https://doi.org/10.1080/00036846.2013.806782>
46. McDonald SC. Does gentrification affect crime rates? Crime and justice (Chicago, Ill). 1986;8:163-201. <https://doi.org/10.1086/449122>
47. De Nadai M, Xu Y, Letouzé E, González MC, Lepri B. Socio-economic, built environment, and mobility conditions associated with crime: A study of multiple cities. Scientific Reports. 2020;10(1):13871. <https://doi.org/10.1038/s41598-020-70808-2>
48. Facchetti E. Police infrastructure, police performance, and crime: Evidence from austerity cuts. 2023. https://www.dropbox.com/s/bzewldmxrs85raf/Facchetti_policestations.pdf?dl=0
49. Giulietti CM, Brendon. Kicking you when you're already down: The multipronged impact of austerity on crime. 2021;746. <https://ideas.repec.org/p/zbw/glodps/746.html>
50. Mirko Draca ML. Law, order and austerity: Police numbers and crime in the 2010s Advantage Magazine: Austerity, 10th Anniversary SpecialLink opens in a new window. 2020. Available from: https://warwick.ac.uk/fac/soc/economics/research/centres/cage/news/03-06-20-advantage_magazine_summer/article-3/3._law_order_and_austerity.pdf.
51. Sherman L, Neyroud PW, Neyroud E. The cambridge crime harm index: Measuring total harm from crime based on sentencing guidelines. Policing: A Journal of Policy and Practice. 2016;10(3):171-83. <https://doi.org/10.1093/police/paw003>



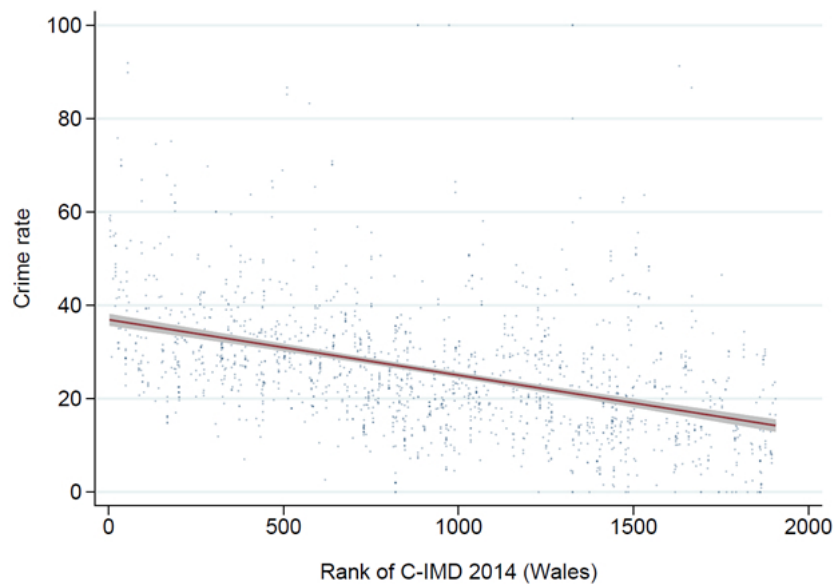
Supplementary Appendix

Supplementary Figure 1: Association between Crime rates (1600-metre buffer and 3-month period) and LSOAs ranks of C-IMD, MCS survey 6 (age 14), by country

Panel A. England



Panel B. Wales



Notes: Panel A and B show a scatter plot with a linear prediction between the Crime rates (1600-metre buffer and 3-month period) per 1,000 population for England and Wales, respectively. The estimated slope, once scaled up by 1,000, in England is -0.7 ($P < 0.000$, $R^2 = 0.29$) and for Wales is -1.2 ($P < 0.000$, $R^2 = 0.19$), after scaled up by 100.



Supplementary Table 1: OLS estimates for total crime rates (3-month, 1600-metre buffer) MCS sweep 6 (age 14), by country

	England		Wales	
	Coef.	95% CI	Wales	95% CI
Crime Domain IMD	−6.52***	[−7.00, −6.04]	−6.06***	[−7.38, −4.75]
White ethnicity	−1.68***	[−2.69, −0.68]	−8.47	[−18.89, 1.96]
Decile of household income				
1 = lowest income	Ref		Ref	
2	−2.37***	[−3.82, −0.92]	−1.85	[−6.57, 2.88]
3	−3.03***	[−4.46, −1.60]	−1.23	[−4.74, 2.27]
4	−3.46***	[−4.97, −1.95]	−2.83	[−6.82, 1.17]
5	−3.23***	[−4.78, −1.68]	−2.88	[−7.92, 2.16]
6	−3.83***	[−5.42, −2.24]	−2.76	[−7.62, 2.10]
7	−4.88***	[−6.46, −3.30]	−2.35	[−7.40, 2.69]
8	−4.26***	[−5.92, −2.59]	−4.92*	[−9.98, 0.14]
9	−5.07***	[−6.73, −3.40]	−5.86**	[−10.50, −1.21]
10 = highest income	−4.54***	[−6.49, −2.59]	−7.67***	[−12.93, −2.41]
Housing tenure				
Rent house or other	Ref		Ref	
Own house	−0.15	[−0.98, 0.68]	1.94*	[−0.19, 4.07]
Maternal Education				
NVQ level 1	−0.03	[−1.59, 1.54]	−0.96	[−6.43, 4.51]
NVQ level 2	0.17	[−1.15, 1.49]	0.31	[−4.65, 5.28]
NVQ level 3	−0.65	[−2.06, 0.76]	1.12	[−4.00, 6.24]
NVQ level 4	−0.34	[−1.54, 0.86]	0.65	[−4.35, 5.64]
NVQ level 5	Ref		Ref	
Overseas qualification	0.99	[−0.83, 2.81]	5.95	[−6.43, 18.33]
Other	−0.26	[−1.76, 1.23]	1.67	[−3.68, 7.01]
Constant	24.93***	[22.84, 27.03]	34.29***	[23.09, 45.48]
Police Force Fixed Effects	Yes		Yes	
Observations	7817		1640	
R ²	0.407		0.238	

Notes: 95% confidence interval in brackets. *p < 0.10, **p < 0.05, ***p < 0.001. Standard errors are clustered at LSOA level, and estimates are weighted using MCS derived sampling weights. Missing income was imputed using the mean of household income at sweep 6. Missing data on ethnicity and housing tenure were imputed using the mode.

Supplementary Table 2: Total crime counts percentiles (20th, 40th, 60th and 80th) by MCS surveys

	MCS 5 (age 11)	MCS 6 (age 14)	MCS 7 (age 17)
P20	113	111	106
P40	318	307	326
P60	581	550	592
P80	1083	1016	1091
P80/P20	9.6	9.2	10.3

Notes: Estimates are weighted using MCS derived weights.

Supplementary Table 3: Crime rates and counts by crime categories (1600-metre buffer and 3-month period)

	Rates			Counts		
	11	14	17	11	14	17
Total	26.8 [25.7,27.8]	24.3 [23.3,25.2]	26.4 [25.3,27.5]	708 [624,791]	659 [586,732]	695 [620,770]
Anti-social behaviour	10.4 [9.8,11.0]	8.3 [7.8,8.8]	5.9 [5.5,6.3]	256 [230,283]	205 [186,224]	147 [132,163]
Burglary	2.3 [2.2,2.4]	1.9 [1.7,2.0]	1.8 [1.7,1.9]	58 [51,64]	47 [42,52]	48 [42,54]
Criminal damage and arson	2.7 [2.6,2.8]	2.5 [2.3,2.6]	2.6 [2.4,2.7]	62 [57,67]	59 [55,64]	60 [55,65]
Drugs	0.9 [0.8,0.9]	0.6 [0.6,0.7]	0.5 [0.5,0.6]	28 [23,33]	20 [16,24]	17 [14,20]
Public disorder and weapons	0.6 [0.6,0.6]	0.9 [0.8,0.9]	1.9 [1.7,2.0]	18 [16,21]	27 [23,32]	50 [44,55]
Robbery	0.3 [0.2,0.3]	0.2 [0.2,0.2]	0.3 [0.2,0.3]	14 [10,18]	9 [7,12]	12 [9,15]
Shoplifting	1.3 [1.2,1.4]	1.4 [1.3,1.5]	1.5 [1.4,1.7]	33 [30,36]	38 [35,42]	40 [36,44]
Vehicle crime	1.8 [1.6,1.9]	1.6 [1.5,1.7]	1.8 [1.7,2.0]	49 [42,56]	46 [40,53]	55 [45,64]
Violence and sexual offences	2.9 [2.7,3.0]	4.2 [4.0,4.4]	7.1 [6.7,7.4]	80 [70,90]	119 [104,134]	178 [162,195]
Others	3.6 [3.4,3.8]	2.8 [2.6,2.9]	3.0 [2.9,3.2]	108 [89,128]	87 [72,102]	88 [73,102]

Notes: The table shows means with 95% confidence interval estimated using the MCS survey design (in parenthesis). Rates columns show crime rates per 1000 population. Counts columns indicate the total counts of crime incidents.



Supplementary Table 4: OLS estimates for the association between household income deciles at age 11, and crime rates at age 14

	Crime rate (age 14)					
	Model 1		Model 2		Model 3	
	Coef.	95% CI	Coef.	95% CI	Coef.	95% CI
Decile of household income (age 11)						
1 = lowest income	14.21*** ^a	[8.49,19.93]	6.86***	[3.98,9.75]	4.47***	[2.16,6.79]
2	10.90***	[6.91,14.89]	4.70***	[2.31,7.08]	2.89***	[0.92,4.86]
3	9.58***	[5.94,13.23]	4.58***	[2.53,6.63]	2.97***	[1.30,4.64]
4	6.62***	[3.87,9.37]	2.33***	[0.59,4.07]	1.2	[-a0.32,2.72]
5	5.57***	[3.12,8.02]	2.84***	[1.30,4.38]	1.79***	[0.46,3.13]
6	3.45***	[1.47,5.43]	1.30**	[0.07,2.53]	0.5	[-0.61,1.61]
7	2.96***	[1.32,4.60]	1.37**	[0.30,2.44]	0.73	[-0.31,1.77]
8	1.89**	[0.45,3.33]	1.31***	[0.43,2.20]	0.75*	[-0.04,1.54]
9	0.46	[-0.65,1.57]	0.31	[-0.39,1.01]	-0.19	[-0.89,0.50]
10 = highest income	Ref.		Ref.		Ref.	
White	-2.28**	[-4.05,-0.52]	0.32	[-0.74,1.37]	-0.7	[-1.68,0.29]
Housing tenure						
Rent house or other	Ref.		Ref.		Ref.	
Own house	0.15	[-1.45,1.75]	0.78	[-0.25,1.82]	0.12	[-0.78,1.01]
Working status						
Both in work	Ref.		Ref.		Ref.	
Main in work, partner not	-0.95	[-3.74,1.85]	-1.95	[-4.41,0.52]	-1.39	[-3.67,0.89]
Partner in work, main not	-1.62***	[-2.66,-0.59]	-1.13***	[-1.83,-0.42]	-0.97***	[-1.62,-0.32]
Both not in work	-0.97	[-3.38,1.43]	-0.82	[-2.38,0.74]	-0.41	[-1.71,0.89]
Main in work or on leave, no partner	-1.89**	[-3.42,-0.36]	-1.09*	[-2.18,0.00]	-0.72	[-1.69,0.25]
Main not in work nor on leave, no partner	-3.46***	[-5.97,-0.94]	-2.64***	[-4.20,-1.07]	-1.85***	[-3.11,-0.60]
Maternal Education						
NVQ level 1	0.17	[-1.96,2.31]	-0.16	[-1.44,1.13]	-0.02	[-1.22,1.18]
NVQ level 2	-0.1	[-1.89,1.69]	-0.1	[-1.09,0.90]	0.02	[-0.94,0.98]
NVQ level 3	-0.74	[-2.46,0.98]	0.14	[-0.99,1.27]	0.19	[-0.90,1.29]
NVQ level 4	-1.35*	[-2.87,0.17]	-0.35	[-1.22,0.53]	-0.24	[-1.10,0.61]
NVQ level 5	Ref.		Ref.		Ref.	
Overseas qualification	0.5	[-2.15,3.14]	-0.39	[-2.14,1.37]	0.01	[-1.71,1.72]
Other	-0.27	[-2.34,1.81]	-0.48	[-1.72,0.75]	-0.31	[-1.48,0.87]
Number of siblings in the household	-0.23	[-0.65,0.20]	-0.09	[-0.37,0.19]	0.04	[-0.22,0.30]
Country						
England	Ref.		Ref.		Ref.	
Wales	1.93*	[-0.31,4.17]	0.26	[-0.90,1.41]	0.24	[-3.80,4.28]
Crime Rate (age 11)			0.69***	[0.64,0.73]	0.67***	[0.63,0.72]
Constant	21.77***	[19.23,24.30]	3.27***	[1.51,5.03]	2.92***	[1.09,4.74]
Police Force Fixed Effect	No		No		Yes	
R ²	0.119		0.629		0.659	
Training (80% sample) MSE=	150.37		56.22		51.76	
Test (20% sample) MSE=	132.29		57.50		52.55	
Observations	7586		7586		7586	

Notes: 95% confidence interval in brackets. *p < 0.10, **p < 0.05, ***p < 0.01. Standard errors are estimated considering the MCS sampling design, and estimates are weighted using MCS derived sampling weights. NVQ = National Vocational Qualifications. Missing income was imputed using the mean of household income at sweep 6. Missing data on ethnicity, housing tenure, working status, and maternal education were imputed using the mode. Binary variables for missing data were included (but not reported) in the table. A train/test split was used to prevent overfitting. We used the splitsample command in Stata and split the sample as 80% for training (n = 6069) and 20% for test (n = 1517). We used the parameters estimated on the training data to compute out-of-sample prediction in the test data, and the Mean Squared Error (MSE) in the training and test data was reported.

Supplementary Table 5: OLS estimates for the association between household income deciles at age 11, and crime rates at age 17

	Crime rate (age 17)					
	Model 1		Model 2		Model 3	
	Coef.	95% CI	Coef.	95% CI	Coef.	95% CI
Decile of household income (age 11)						
1 = lowest income	19.89***	[14.46,25.31]	11.00***	[7.62,14.39]	5.21***	[2.47,7.96]
2	15.31***	[11.34,19.29]	8.24***	[5.24,11.24]	3.56***	[1.12,6.00]
3	13.77***	[9.60,17.93]	7.69***	[4.94,10.44]	3.51***	[1.37,5.64]
4	9.32***	[6.57,12.06]	4.72***	[2.54,6.90]	1.76*	[-0.10,3.62]
5	7.27***	[4.84,9.69]	4.11***	[2.53,5.69]	1.56**	[0.16,2.96]
6	4.74***	[2.53,6.96]	2.22***	[0.74,3.70]	0.36	[-0.94,1.65]
7	4.41***	[2.51,6.31]	2.70***	[1.15,4.25]	0.96	[-0.44,2.35]
8	2.54***	[1.04,4.03]	2.02***	[0.97,3.06]	0.73	[-0.23,1.68]
9	0.89	[-0.19,1.97]	0.73*	[-0.12,1.58]	-0.41	[-1.16,0.35]
10 = highest income	Ref.		Ref.		Ref.	
White	-1.15	[-2.77,0.48]	1.30**	[0.07,2.52]	-0.12	[-1.09,0.84]
Housing tenure						
Rent house or other	Ref.		Ref.		Ref.	
Own house	0.66	[-0.69,2.01]	0.72	[-0.42,1.87]	-0.65	[-1.72,0.41]
Working status						
Both in work	Ref.		Ref.		Ref.	
Main in work, partner not	-3.15***	[-5.22,-1.09]	-4.54***	[-6.80,-2.28]	-2.79***	[-4.86,-0.73]
Partner in work, main not	-1.84***	[-3.16,-0.52]	-1.43***	[-2.33,-0.52]	-0.54	[-1.29,0.22]
Both not in work	-2.72**	[-5.32,-0.12]	-1.95*	[-4.06,0.17]	-0.87	[-2.49,0.75]
Main in work or on leave, no partner	-2.79***	[-4.61,-0.98]	-1.71***	[-3.00,-0.42]	-0.54	[-1.66,0.57]
Main not in work nor on leave, no partner	-5.80***	[-8.43,-3.17]	-4.88***	[-6.69,-3.07]	-2.75***	[-4.32,-1.19]
Maternal Education						
NVQ level 1	0.87	[-1.69,3.44]	0.53	[-1.26,2.31]	0.97	[-0.62,2.56]
NVQ level 2	0.03	[-1.86,1.92]	0.06	[-1.20,1.33]	0.64	[-0.61,1.89]
NVQ level 3	-0.39	[-2.30,1.52]	0.31	[-1.08,1.71]	0.77	[-0.57,2.11]
NVQ level 4	-1.03	[-2.57,0.52]	-0.05	[-1.23,1.13]	0.17	[-0.98,1.32]
NVQ level 5	Ref.		Ref.		Ref.	
Overseas qualification	0.77	[-2.57,4.12]	-0.45	[-2.52,1.62]	0.12	[-1.69,1.94]
Other	0.29	[-2.17,2.76]	-0.54	[-2.20,1.12]	0.07	[-1.41,1.56]
Number of siblings in the household	-1.08***	[-1.59,-0.58]	-0.67***	[-1.03,-0.31]	-0.29*	[-0.62,0.04]
Country						
England	Ref.		Ref.		Ref.	
Wales	2.14	[-0.50,4.77]	0.36	[-1.16,1.88]	2.84	[-4.67,10.35]
Crime Rate (age 11)			0.71***	[0.65,0.76]	0.69***	[0.64,0.75]
Constant	22.55***	[20.09,25.01]	4.13***	[1.64,6.61]	3.96***	[1.51,6.41]
Police Force Fixed Effect	No		No		Yes	
R ²	0.112		0.548		0.607	
Training (80% sample) MSE =	197.49		106.84		101.94	
Test (20% sample) MSE =	175.44		97.56		92.01	
Observations	7586		7586		7586	

Notes: 95% confidence interval in brackets. *p < 0.10, **p < 0.05, ***p < 0.01. Standard errors are estimated considering the MCS sampling design, and estimates are weighted using MCS derived sampling weights. NVQ= National Vocational Qualifications. Missing income was imputed using the mean of household income at sweep 6. Missing data on ethnicity, housing tenure, working status, and maternal education were imputed using the mode. Binary variables for missing data were included (but not reported) in the table. A train/test split was used to prevent overfitting. We used the splitsample command in Stata and split the sample as 80% for training (n=6069) and 20% for test (n=1517). We used the parameters estimated on the training data to compute out-of-sample prediction in the test data, and the Mean Squared Error (MSE) in the training and test data was reported.

Supplementary Table 6: OLS estimates for the association between household income deciles at age 11, and crime rates at age 14 and 17

	Crime rate (age 14)		Crime rate (age 17)	
	Coef.	95% CI	Coef.	95% CI
Decile of household income (age 11)				
1 = lowest income	4.44***	[2.04,6.83]	4.43***	[1.62,7.24]
2	2.88 ^o **	[0.84,4.92]	3.03**	[0.65,5.40]
3	2.92***	[1.19,4.66]	2.70**	[0.58,4.82]
4	1.19	[-0.37,2.74]	1.56	[-0.32,3.45]
5	1.79***	[0.44,3.14]	1.31	[-0.27,2.90]
6	0.56	[-0.57,1.70]	0.48	[-0.80,1.76]
7	0.72	[-0.33,1.78]	0.49	[-0.68,1.66]
8	0.74 ^o	[-0.05,1.54]	0.75	[-0.26,1.77]
9	-0.20	[-0.89,0.50]	-0.56	[-1.37,0.24]
10 = highest income	Ref.		Ref.	
Ethnicity				
White	Ref.		Ref.	
Other	-0.70	[-1.68,0.28]	-0.15	[-1.14,0.84]
Housing tenure				
Rent house or other	Ref.		Ref.	
Own house	0.15	[-0.77,1.07]	-0.14	[-1.01,0.72]
Working status				
Both in work	Ref.		Ref.	
Main in work, partner not	-1.37	[-3.65,0.91]	-1.57**	[-3.03,-0.10]
Partner in work, main not	-0.98***	[-1.63,-0.32]	-0.44	[-1.24,0.35]
Both not in work	-0.37	[-1.67,0.94]	-0.17	[-1.76,1.42]
Main in work or on leave, no partner	-0.70	[-1.68,0.29]	-0.32	[-1.50,0.85]
Main not in work nor on leave, no partner	-1.81***	[-3.08,-0.54]	-1.76**	[-3.27,-0.25]
Maternal Education				
NVQ level 1	-0.01	[-1.22,1.19]	0.57	[-0.93,2.06]
NVQ level 2	0.04	[-0.92,1.01]	0.52	[-0.76,1.80]
NVQ level 3	0.19	[-0.91,1.29]	0.79	[-0.50,2.07]
NVQ level 4	-0.23	[-1.09,0.62]	0.19	[-0.91,1.29]
NVQ level 5	Ref.		Ref.	
Overseas qualification	-0.02	[-1.74,1.70]	0.04	[-1.76,1.84]
Other	-0.31	[-1.49,0.87]	0.28	[-1.21,1.77]
Number of siblings in the household	0.04	[-0.21,0.30]	-0.14	[-0.44,0.15]
Country				
England	Ref.		Ref.	
Wales	0.23	[-3.83,4.28]	2.05	[-3.88,7.98]
Crime Rate (age 11)	0.68***	[0.63,0.72]	0.69***	[0.65,0.74]
Constant	2.87***	[1.04,4.70]	3.00***	[0.91,5.09]
Police Force Fixed Effects	Yes		Yes	
Observations	7586		7586	

Notes: 95% confidence interval in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are estimated considering the MCS sampling design, and estimates are weighted using MCS derived sampling weights. NVQ = National Vocational Qualifications. Multiple imputation by chained equation combining 20 datasets was used to impute missing information on housing tenure (1.6%), working status (0.2%), number of siblings (0.9%) and maternal education (0.1%). A multinomial logit model was used for working status and maternal education, a linear model was used for number of siblings, and a Logistic model was used for housing tenure. Predictors in multiple imputation models included ethnicity, deciles of household income and country.