Article:

Roberts, J., Tubeuf, S. and Tyler, P. (2024) Evaluating area-based policies using secondary data: the neighbourhood management pathfinders programme. Housing Studies, 39 (7). pp. 1787-1812. ISSN: 0267-3037

https://doi.org/10.1080/02673037.2022.2146065

Housing Studies

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/chos20

Evaluating area-based policies using secondary data: the neighbourhood management pathfinders programme

Jennifer Roberts, Sandy Tubeuf & Peter Tyler

To cite this article: Jennifer Roberts, Sandy Tubeuf & Peter Tyler (2022): Evaluating area-based policies using secondary data: the neighbourhood management pathfinders programme, Housing Studies, DOI: 10.1080/02673037.2022.2146065

To link to this article: https://doi.org/10.1080/02673037.2022.2146065

Published online: 23 Nov 2022.

HOUSING STUDIES
https://doi.org/10.1080/02673037.2022.2146065

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Evaluating area-based policies using secondary data: the neighbourhood management pathfinders programme

Jennifer Roberts[a], Sandy Tubeuf[b] and Peter Tyler[c]

[a]University of Sheffield, Sheffield, UK; [b]Université catholique de Louvain, Ottignies-Louvain-la-Neuve, Belgium; [c]University of Cambridge, Cambridge, UK

ABSTRACT
This study explores the feasibility of evaluating neighbourhood-based initiatives using secondary data. As a case study, we consider the Pathfinders programme, an area-based UK intervention and evaluate its impact on a broad range of outcomes in both short and longer term. We use grid reference data from a household panel survey to identify individuals living in 'treated' areas before and after and appropriate control individuals living in 'untreated' areas. Using a difference-in-difference approach complemented with matching, we find that the programme had positive effects on reported neighbourhood problems as well as on local social interaction, which was not an intended outcome. We also show the practical usefulness of combining secondary data and geographical identifiers to evaluate area-based policies. Using data not collected for this purpose enables the consideration of a broad range of intended and unintended outcomes over the long-run. A drawback of the approach is to require large scale geographical initiatives to ensure a sufficient number of targeted units.

ARTICLE HISTORY
Received 14 December 2021
Accepted 4 November 2022

KEYWORDS
Evaluation; neighbourhood-based policies; propensity score matching

JEL CODES
I28; R23; R58

1. Introduction

Over the past twenty years the UK government has been investigating the scope for local authorities to plan and deliver more of the services that they are responsible for at the local level. Under the Localism Bill, the Coalition Government granted new freedoms and flexibility to local communities in how they seek to deliver services at the neighbourhood level (DCLG, 2011). The aim was 'to help people and their locally elected representatives to achieve their own ambitions. This is the essence of the Big Society.' (Rt Hon Greg Clark p.2 in DCLG, 2011). More recently, the government has announced its intention to adopt a 'levelling-up agenda' that is seeking to help tackle pronounced spatial inequalities across local authorities in England and close the gaps between the country's richest and poorest

CONTACT Sandy Tubeuf s.tubeuf@leeds.ac.uk, sandy.tubeuf@uclouvain.be University of Leeds, Leeds, UK; Université catholique de Louvain, Ottignies-Louvain-la-Neuve, Belgium.

communities.[1] Policies to target resources to assist local areas thus look set to remain, partly prompted by the way that the Covid-19 pandemic has brought renewed attention to area-based inequality.

The UK government has deployed a number of neighbourhood-based initiatives, particularly under New Labour's National Strategy for Neighbourhood Renewal (HM Government (1998). These initiatives provide an extensive evidence base on which to draw upon, which is of value in guiding future policy. However, despite their enduring popularity as a policy tool, there is remarkably little convincing evidence of the effectiveness of these area-based initiatives (ABIs). Rhodes *et al.* (2005) point out that 'too little is known about how successful these ABIs are in helping to turn around depressed areas'. (p.1920). More recently, Gibbons *et al.* (2021) state that 'despite their popularity, the economic (and broader) impacts of such programmes are uncertain'. (p.1)

One of the main reasons for the paucity of reliable evidence is the poor quality of some of the evaluations that have been implemented to assess the effectiveness of ABIs. A National Audit Office (2013) report on evaluation in the UK government identifies a large amount of variation in the quality of evaluation and points to particular weaknesses with the evaluation of spatial policy. The accompanying independent review concluded that none of the spatial policy evaluations provided convincing evidence of policy impacts (Gibbons *et al.*, 2013).[2]

There is no doubt that ABIs are difficult to evaluate. They are generally complicated, broad ranging approaches to bringing about change; Rhodes *et al.* (2005) describe them as 'holistic' or 'multifaceted'. In medical jargon ABIs would be classified as 'complex interventions' because they involve several interacting components (MRC, 2006). Certain features combine to create this complexity. Firstly, it may not be appropriate to classify ABIs as a single intervention, since they involve many different components, each targeting different outcomes. For example, as Rhodes *et al.* (2003) point out in relation to the Single Regeneration Budget (SRB), programmes often involve a large number of partnerships that vary considerably in scope and size. Secondly, the interventions often have a long duration, requiring long-term evaluation. This problem is exacerbated by the fact that many outcomes change very gradually, and even if change does occur in the short-term, it may not persist into the long-term, hence evaluation beyond the timeline of the intervention is necessary. For example, as O'Reilly (2007) underlines, the perceptions of residents may get a short-term boost from seeing physical investments in areas that have been long-neglected, but their satisfaction with the area as a place to live may not actually improve in the longer-term. Thirdly, ABIs often have a large spread of objectives that range from specific outcomes for individuals and households who live in the targeted areas to changes to the processes underlying regeneration; these different outcomes may require different evaluation methods. Fourthly, the impact of ABIs may benefit residents who move out of their neighbourhood during the period of observation and it may be difficult to quantify these impacts. Finally, it is not always clear how to define the treatment group in the context of ABIs, because they tend to have important spillover effects, affecting areas and individuals outside of the targeted intervention area, both positively and negatively.

Nevertheless, evidence is needed to guide public policymaking and to assess what ABIs can achieve in terms of the improvements on a broad range of individual and community outcomes. Most evaluations have relied on directly surveying those responsible for delivering the interventions and/or those targeted by the interventions. These interview responses are particularly prone to be biased by focusing effects (Schkade & Kahneman, 1998) and concerns about future funding, and this likely exaggerates any potential positive effects of the intervention. There is, however, little alternative to bespoke surveys since at first glance, few available secondary data sources can provide information about the effects of ABIs (Rhodes *et al.*, 2005).

In this study, we argue that some longitudinal household surveys, such as the British Household Panel Survey (BHPS), can be valuable in this regard. These surveys provide rich longitudinal data available annually since 1991, and crucially they also include (via the Secure Lab)[3] detailed geographical identifiers for each household. This combination has enormous potential for evaluating policies where there is geographical variation in implementation. First, the use of secondary data, not collected specifically for evaluating the intervention in question, reduces the problem of reporting biases and focusing effects. Second, the rich longitudinal information allows for controlling for a broad set of confounding factors, as well as longer follow-up than is generally possible with bespoke evaluation data. Third, large secondary data sets allow for estimating the effectiveness of the ABI on their initially targeted outcomes, but also on broader, non-intended outcomes. Finally, given that these large secondary data sets generally cover the whole country, they can allow for flexible alternative definitions of the treated and untreated areas; for example, facilitating exploration of spillover effects to areas not specifically targeted by the policy.

This study makes two primary contributions. Firstly, it aims to show how ABIs can be evaluated using secondary data. Here we use the Neighbourhood Management Pathfinders Programme of the early 2000s as a case study, illustrating how to implement this form of evaluation in practice. Secondly, it provides valuable empirical evidence from an evaluation of the Pathfinders programme, using comparable control areas. We explore the effects of 'Pathfinders' on a broad range of intended and non-intended outcomes in both the short and longer term. A final secondary aim is to help to persuade researchers that obtaining Secure Lab access to these data is worth the extra steps required.[4] While the household survey data in the BHPS have been widely used within social science research, the very detailed geographical identifiers are under-exploited, possibly because they are only available via the UK Data Service Secure Lab. More generally, this study adds to the development of empirical methods to evaluate housing and neighbourhood-based interventions (Galster & Santiago, 2015; Galster & Hedman, 2013; Galster *et al.*, 2004).

The Pathfinders programme was a flagship Labour government initiative aimed at enabling deprived communities to improve local outcomes, by improving and joining up local services (such as the police, environmental services and local health care providers), making them more responsive to local needs. The idea was to test the potential role of neighbourhood management in promoting local renewal and narrowing the gap between deprived areas and the rest of the country. The original evaluation of the Pathfinders programme carried out for the Department for

Communities and Local Government (DCLG) was based on extensive fieldwork, case studies on particular themes and two waves of bespoke household surveys (DCLG, 2008). However, these surveys considered only a narrow range of benefits accruing only to those residents living in Pathfinders areas (DCLG, 2006). This policy initiative provides a natural experiment that we exploit for the evaluation; it represents an observational study in which we cannot control or withhold the allocation of the intervention to particular areas or communities, but where natural or predetermined variation in allocation occurs (Petticrew *et al.*, 2005).

The outcome data are selected from the BHPS. We identify the intervention areas before and after the Pathfinders programme using postcode identifiers and the BHPS grid reference data (defined below). We identify a control group composed of similar individuals in similar areas that did not implement the Pathfinders programme. We compare changes in outcomes between the treated and matched control individuals over time using a difference-in-difference approach.

The remainder of the paper proceeds as follows. Section 2 provides an overview of the Pathfinders programme. Section 3 presents the methods and the empirical assumptions, while Section 4 describes the data. The main empirical results and robustness checks are presented and discussed in Section 5, and in Section 6 we summarise and highlight our main conclusions.

2. Evaluation of area-based interventions and the pathfinders programme

The Pathfinders programme was a £100 million investment aimed at bringing the local community and local services together to tackle problems in the area and improve local services. 'Pathfinders' were not randomly allocated across England, rather they were specifically targeted at deprived areas. Thirty-five Pathfinders programmes were funded between 2001 and 2007; some in inner cities, others in residential estates on the edge of towns, and some in coastal towns.[5] The Pathfinders programme required that all participating local authorities provided matched funding when enrolled in the Pathfinders programme. This gives us increased confidence that the intervention areas were all engaged and motivated to deliver the programme. There were also required elements of common delivery so it is therefore appropriate to assume that the implementation of the programme was similar across areas. However, the specific components of the Pathfinders programme did vary from one area to the other, as they were adapted to local needs. For example, some areas implemented new lighting to improve safety while others empowered the local community by building a community centre. The majority of the areas adopted a mixed approach to delivery combining the improvement of mainstream services and building local community capacity and also delivering at least some projects and services at the neighbourhood level. While not completely homogeneous the programmes all shared a common model: a professional manager who had responsibility for viewing the neighbourhood in its totality rather than simply being concerned with specific services, and who sought to develop a systematic, planned approach to improving the quality of life in that neighbourhood (DCLG, 2008).

The original Pathfinders evaluation carried out for DCLG was not a controlled study, however the results were generally suggestive of positive effects on residents' satisfaction with their area, community safety, fear of crime, and general environmental improvements (DCLG, 2008). The final evaluation report stresses that for this programme it was particularly hard to identify measurable impacts because Pathfinders were small-scale strategic local initiatives, not large-spend, large-scale projects. Here, we intend to evaluate the aggregate results of local programmes with a similar Pathfinders design and implementation.

3. Methods

The aim of this study is to analyse whether there is a causal effect of the Pathfinders programme on a number of outcomes over time; we focus on the outcomes for people living in the Pathfinders areas, but also consider the possibility of spillover effects that extend outside of the targeted areas.

We consider that the Pathfinders programme is comparable to a quasi-experiment, and use a difference-in-difference (DiD) method combined with matching in order to estimate the causal impact of the policy (Blundell & Costa-Dias, 2000). We first use matching to ensure that we identify individuals in the non-treated areas who are as similar as possible to treated individuals in intervention areas considering observed characteristics at baseline. We then use a regression-based difference-in-difference method where time-specific additional variables are included to control for potential confounding factors. This combination of methods with pre-matching is recommended by Heckman *et al.* (1997, 1998) and Stuart (2010); it is an attempt to overcome (as far as possible) any bias that results from the non-random allocation of Pathfinders. The DiD method estimates the effect of the intervention by comparing the before and after outcomes for individuals living in areas where the programme was implemented with similar control individuals living in similar areas where the programme was not implemented.

DiD has been widely applied in many areas of policy evaluation[6] (see Bertrand *et al.*, 2004 for a critical review). In the context of neighbourhood-based policies, similar methods have been used to evaluate the employment outcomes of the New Deal for Communities (Romero & Noble, 2008; Romero, 2009), the Single Regeneration Budget (Gibbons *et al.*, 2021), the impact of area-based interventions on house prices (Aarland et al, 2017) as well as the impact of regeneration programmes on health-related outcomes (Ruijsbroek *et al.*, 2017) and on area-level economic outcomes of run-down industrial sites (Ploegmakers & Beckers, 2015). These empirical studies either used data collected with the sole purpose of evaluating the policy in question, administrative data, such as the UK Census using enumeration district as the smallest unit of observation or firm level micro-data. In this study we use secondary data from a general household survey to evaluate the effects of the Pathfinders policy.

Consider Y_{it} is the outcome of interest for individual i in time t. We compare the change in outcomes before and after the intervention for individuals and households in the intervention areas (IA) with the change for those in the control areas (CA). The naïve DiD estimated impact is given by:

$$\hat{\beta} = \left(Y_{IA}^{Post} - Y_{IA}^{Pre}\right) - \left(Y_{CA}^{Post} - Y_{CA}^{Pre}\right) \tag{1}$$

Where $\hat{\beta}$ is the estimated impact of the Pathfinders programme for the individuals in the IA, or the average effect of the treatment on the treated (ATT) (Cameron & Trivedi, 2005; Section 25.3). The superscripts Pre and Post refer to outcome measurement before and after the intervention. We extend this to a regression-based DiD estimator given by:

$$Y_{it}^{*} = \beta_0 + \beta_1 Path_i + \beta_2 Post_t + \beta_3 (Path_i \times Post_t) + \beta_4 X_{it} + \varepsilon_{it} \tag{2}$$

Where Y_{it}^{*} is the unobserved continuous latent outcome underlying the observed dichotomous outcome Y_{it}, $Path_i$ is a dummy variable which takes the value 1 if the individual lives in a Pathfinders area (the IA), and 0 if living in a CA; $Post_t$ is a dummy variable which takes the value 1 if the observation is after the intervention and 0 otherwise; $(Path_i \times Post_t)$ is an interaction term and the estimated ATT is given by its coefficient β_3; ε_{it} is the individual error term. The regression equation (2) is estimated as a linear probability model; it has the advantage, over equation (1), of estimating a standard error for the ATT and also allowing us to control for potentially any remaining observed differences between the IA and the CA given by X_{it} (Imbens & Wooldridge, 2009). For example, household composition may be different, on average, in the IA compared to the CA.

The logic model underlying the 'Pathfinders' approach predicts a broad range of impacts on the neighbourhood including lower crime and anti-social behaviour, cleaner streets, better housing, higher educational attainment, healthier people, higher levels of employment, higher household incomes and stronger community networks (DCLG, 2008). The original evaluation (DCLG, 2008) did not consider all of these potential outcomes, as some of them were not seen as primary outcomes, but the rich set of variables available in the BHPS enables us to consider most of these possible effects.

Ideally in a DiD evaluation setting, the only difference between the compared individuals should be exposure to the intervention. This is not true by design in the context of the Pathfinders as these interventions were targeted at deprived areas, which are systematically different to many other areas of the country in a number of respects.[7] We therefore combine the regression-based DiD estimation with a matching procedure to compare treated individuals to similar control individuals in similar areas. Matching is done across a chosen set of observed characteristics (individual and area based) at baseline (including, for example, household income and area level deprivation); while DiD (which compares the change in outcomes for both groups from baseline to follow-up) also controls for the effect of any *unobservable factors* that do not change over time (for example, pre-existing levels of communication and coordination among service providers, or pre-existing levels of community based trust). Together they reduce the concern that any difference in outcomes is due to the type of areas where Pathfinders were introduced, and is more a result of the Pathfinders programmes themselves. The underlying assumption is that the matched individuals in the CAs are a good representation of what would have happened to the individuals in IAs in the absence of the intervention i.e. it estimates what would have happened to an individual in a Pathfinders area if they had not had a Pathfinders programme.

Matching can be done in variety of different ways, and the method we use here is single nearest neighbour propensity score matching (PSM) without replacement.[8] This essentially means taking each individual in a Pathfinders area and then finding an individual in a CA who looks as similar as possible according to their propensity score. This score is the predicted probability that any individual will be 'treated', given their observable characteristics at wave 10. It is found by estimating a Probit regression on the pooled sample of the individuals living in IAs and CAs. The predicted probabilities from this regression are used to select the closest matches; in this case the single individual (in a similar area) with the closest propensity score to the IA individual (Rosenbaum and Rubin, 1983). Within the vector of possible confounding variables X in equation (2), we distinguish a subset of variables M, which are believed to affect the probability for an individual to be a Pathfinders area. The unconfoundness assumption requires that the assignment to the treatment is random conditional on the subset of variables M. This is generally achieved by finding a good balance across the set of covariates that we match on. It is, however, worth noting that balancing is not strictly necessary when DiD analysis is employed because we are comparing the *change* in outcomes for the treatment and control groups not the levels. The *parallel trends assumption* is more important than balancing in this context. This is important to ensure the internal validity of the DiD estimates, and requires that in the absence of treatment, any difference between outcomes for the IA and the CA is constant over time, rather than already diverging or converging.

The parallel trends assumption cannot be tested directly because we cannot know what would have happened to an individual in an IA had they not been treated. However, we can get an idea of the validity of this assumption by comparing trends in outcomes for both groups in the pre-intervention period. Ideally, we would use time series observations on the outcomes of interest before the intervention. In our analysis we consider outcomes one year before and several years after the intervention, to consider short-term and longer-term outcomes and to explore any time varying confounding factors and anticipatory effects occurring purely as a result of the announcement of the policy.

4. Data

The BHPS is a rich longitudinal data set available from 1991 to 2008. Originally, the BHPS consisted of 5000 households, and all eligible adults (aged 16 years old and more) are followed up and interviewed annually. The data contain rich information on the social and economic circumstances of individuals and households. We considered year 2000 (wave 10) as the most appropriate year to measure pre-treatment characteristics in order to ensure exogeneity as individuals may anticipate the programme before it is implemented (from 2001/2). The analysis was undertaken on an unbalanced sample of individuals who must be observed in the year 2000/1 and whose post intervention outcomes were observed at any time-point over 7 years between 2001/2 and 2007/8. We do not account for movers in the analysis. However, the number of movers was very small. Across the waves we only have 20 movers. This small number is not unexpected given the socio-economic profile of the residents (Papoutsaki *et al.*, 2020). Furthermore, among this sample of movers, we cannot always distinguish between those who moved and those who might have dropped out of the survey for other reasons.

4.1. Identification of the intervention areas

Detailed geographical identifiers, including grid reference, are available for each household in the BHPS via the Secure Lab. The British National Grid is the common referencing format for all geographic data in Great Britain. A location is described using a pair of Easting and Northing coordinates, which gives its East and North distance from the origin (a fixed point to the west of the Scilly Isles) and provides the exact location for each household to a 1-metre resolution.

'Pathfinders' areas were identified by six-digit postcode;[9] there were a total of 7386 'treated' postcodes within the 35 Pathfinders areas.[10] We therefore retrieved the Easting and Northing coordinates of each of the treated postcodes using the online geography matching and conversion tool GeoConvert.[11]

Consider that the Pathfinders programme funded $j = 1,.,J$ different postcodes and P_j is a postcode with the corresponding Easting and Northing coordinates $\left(E_j; N_j\right)$. Given the initiatives included in Pathfinders, for example improving communication with local police and environmental services, it is unrealistic to hypothesise that the effects would be strictly confined to households within the 6-digit IA postcodes themselves; rather there are highly likely to be spillover effects to adjacent areas. In our analysis we define the treated areas in a number of different ways in order to take account of these effects. Our primary definition of the treatment area is based on the overlap of targeted Pathfinders postcodes and lower layer super output areas (LSOA). An LSOA is a geographical area with an average of 1,500 residents and 650 households; it is the UK equivalent of the US 'neighbourhood' area that is often utilised in urban and regional research. In total there are over 32,000 LSOAs in the UK. In this primary definition of the treatment area any postcode situated within the same LSOA as a Pathfinders postcode, P_j, was defined as treated in the analysis. This area was chosen as being sufficiently broad to account for any potential spillover effects into the local neighbourhood.

We also explore two possible alternative definitions of treatment area. Firstly, we assume that the intervention area captured a minimal radius of 800 m (about half a mile) around each Pathfinders postcode P_j; this distance was chosen as a reasonable approximation to the immediate vicinity of the Pathfinders. Secondly, we extend this radius to 1 km (1000 m) to capture effects across a slightly larger area. Both of these areas are slightly smaller than that captured by our primary neighbourhood (LSOA) definition.

We used the geographic information system ArcGIS to identify samples of BHPS respondents living in households in a treated area by overlapping two maps: a map of the intervention areas (one of the three alternative definitions) and a map of the

BHPS surveyed households. Most of the Pathfinders are within urban settlements with a population of 10,000 or more and the wider surrounding area is less sparsely populated.[12] Hence, we only considered urban areas as candidate control areas. The sample sizes of BHPS households living in the alternative 'treated' areas are presented in Table 1. It is important to stress that these sample sizes are not very large. While they may be large enough to be reliable for statistical inference, the tests we perform have only limited power to detect differences.

4.2. Outcomes of interest

The BHPS contains good proxy indicators for the intended outcomes of the Pathfinders programme via a question on housing problems that asks: '*Does your accommodation have any of the following problems*: (i) shortage of space, (ii) noise from neighbours, (iii) other street noise (traffic, businesses, factories etc), (iv) condensation, (v) damp walls, floors, foundations, etc., (vi) pollution, grime or other environmental problem caused by traffic or industry, (vii) vandalism or crime in the area. To utilise these questions, we construct a binary outcome for each of the problems, taking the value 1 if the survey participant reported problems and 0 otherwise.

As well as these housing and neighbourhood outcomes we also consider a health measure (viii) and a social interaction measure (ix) as potential outcomes of interest. The logic model underlying the Pathfinders programme suggests that these outcomes might have been impacted even though they were not the primary targeted outcomes. For health, we use the General Health Questionnaire (GHQ), which is a 12-item questionnaire measuring psychological health; it covers feelings of strain, depression, inability to cope, anxiety-based insomnia, and lack of confidence. The GHQ is a widely recognised instrument that has been adopted by the World Health Organisation as a screening tool for psychological disorders and has been validated in a number of international studies (Goldberg et al, 1997). Responses to GHQ questions are made on a four-point scale of frequency (*not at all*, *no more than usual*, *rather more than usual*, and *much more than usual*), that is then summarised as a score between 0 (the least distressed) and 36 (the most distressed). The measure of social interaction is captured by the question asking '*How often do you talk to any of your neighbours?*' to which participants respond with '*On*

Table 1. Sample size of treated individuals using alternative geographical identifications.

	Sample	Within LSOA (%)	Within 1000 m (%)	Within 800 m (%)
Year - Wave				
2000/1 – 10 (base year)	6,399	174 (2.72)	156 (2.44)	129 (2.02)
2002/3 – 12	5,411	140 (2.59)	134 (2.48)	115 (2.13)
2003/4 – 13	5,336	135 (2.53)	126 (2.36)	103 (1.93)
2004/5– 14	5,185	128 (2.47)	122 (2.35)	105 (2.03)
2005/6 – 15	5,174	110 (2.13)	106 (2.05)	87 (1.68)
2006/7 – 16	5,131	119 (2.32)	113 (2.20)	96 (1.87)
2007/8 – 17	4,936	115 (2.33)	110 (2.23)	94 (1.90)
2008/9 – 18	4,843	112 (2.31)	107 (2.21)	95 (1.96)

% represents the share of individuals within the BHPS sample within urban settlements.

most days, Once or twice a week, Once or twice a month, Less often than once a month, Never'. This measure can be considered as a proxy of the level of trust in the community or social isolation. Both of these are related to the more general construct of social capital, which has been used to describe the quality of the neighbourhood context and living conditions (Forrest and Kearns 2001, Wickes *et al.*, 2019,). Descriptive statistics for the outcomes of interest for treated (IA) and candidate control areas (CA) are presented in Table A1 in the appendix. These statistics are reproduced as line graphs in Figure 1. These line graphs are an important way of exploring how the outcomes have evolved over time, and give an indication of pre-treatment trends, which can help to inform the validity of the parallel trends assumption. These statistics are for all individuals in the potential CAs, before matching has taken place.

The DiD method rests on the assumption that the outcomes for an individual living in an IA and an individual living in a CA would share a common trend in the absence of the intervention. The graphs in Figure 1 suggest that this is certainly not the case for most outcomes in the raw (unmatched) data. Looking first at the housing problems variables shown in Figure 1 (i) to (vii), the proportion of people reporting these problems in the potential control group is lower than for the treatment group in the base year 2000 (wave 10) and in most case remains lower throughout the period that we observe the outcomes. This is as expected given the Pathfinders were focused on particularly deprived areas. For (i) shortage of space, (ii) noise from neighbours and (iii) noise from the street, the proportions experiencing these problems in the treatment and potential control groups seemed to be diverging before the intervention, with the Pathfinders areas getting worse and the potential control areas improving. However, after 2002 (wave 12) there are some clear changes. For (i) shortage of space, these outcomes continued to diverge, reaching a peak in the Pathfinders areas in wave 16 before falling rapidly to wave 18. For both (ii) noise from neighbours and (iii) noise from the street, these fell steeply in the Pathfinders areas to wave 14 before increasing again. The outcomes (iv) condensation, (v) damp and (vi) pollution follow similar trends. They are already converging prior to 2002 and this convergence accelerated to 2003 with rapid improvement in the Pathfinders areas, which start to deteriorate again from wave 14. Vandalism and crime (vii) seem to be following similar upward trends in the Pathfinders areas. After wave 14 this outcome continues to improve in the potential control areas but worsens in the Pathfinders areas. Average levels of psychological health (viii) are worse in the potential control areas at baseline (as shown by the higher GHQ scores) and these outcomes are diverging to 2002. The line graph for the Pathfinders areas is quite noisy, reflecting the small sample size, but despite some fluctuations the overall average level of psychological health in the Pathfinders areas remains better than in the control areas throughout. The frequency of talking to neighbours (ix) starts lower in the Pathfinders areas but rapidly increases and overtakes the potential control areas in 2001. This divergence seems to continue throughout the period we observe the outcomes.

While many of the outcomes do not share pre-treatment trends, it is reasonable to assume that if successful matching brings the levels of outcomes for the treatment

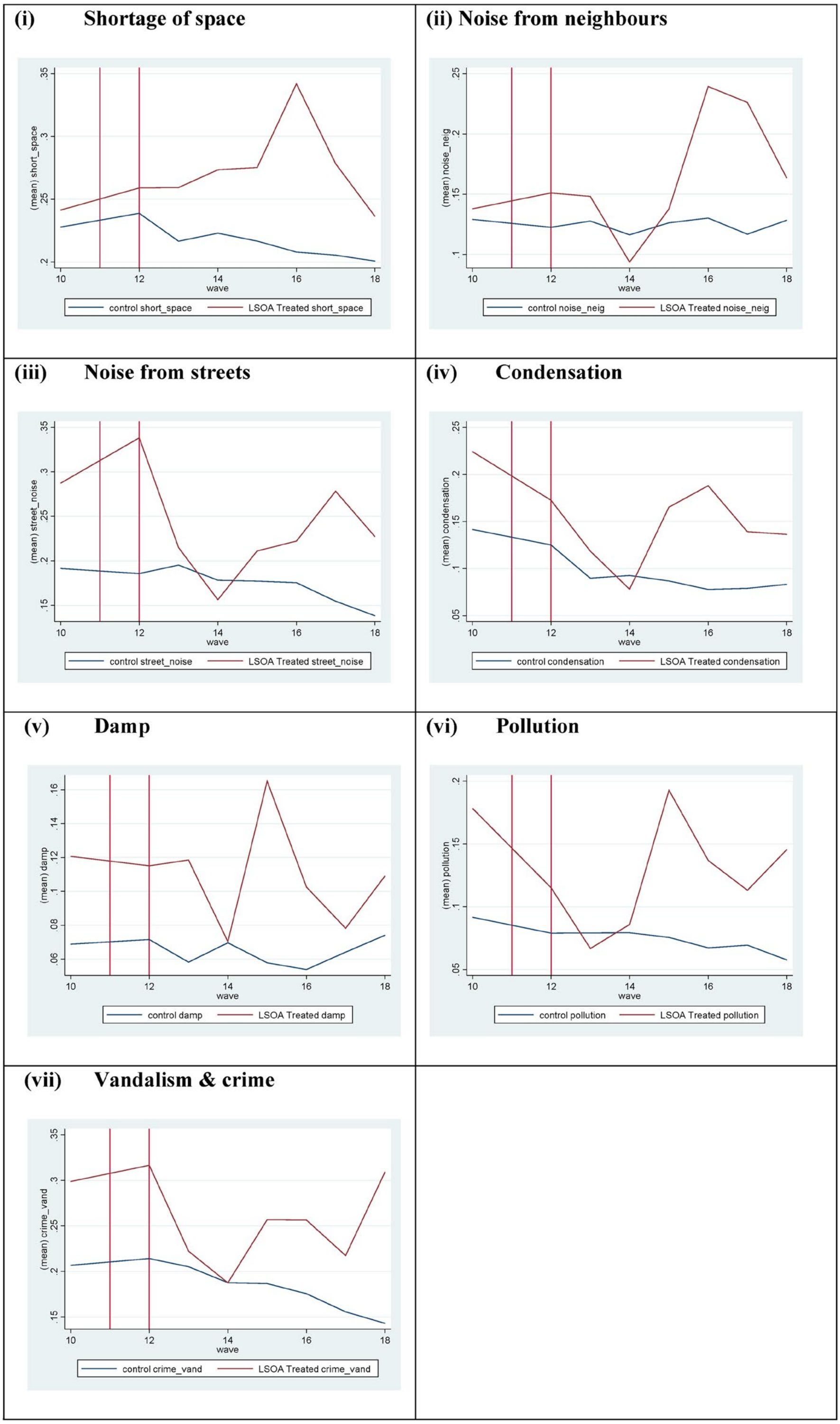

Figure 1. Means outcomes for treatment groups and potential controls.
Note: The two vertical lines show the start of the Pathfinders in 2001/2 (between waves 11 and 12).

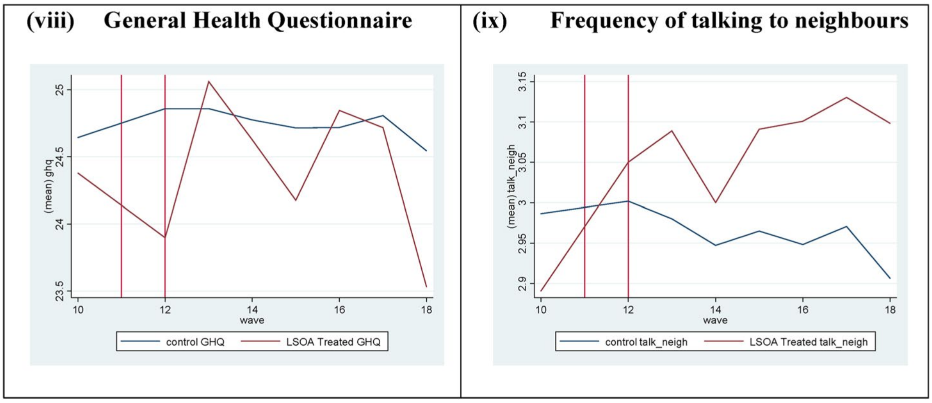

Figure 1. Continued.

and control groups together, the same process will also have some influence on the trends. While Pathfinders were focused on deprived areas, similar deprived areas are available in the potential control areas. For example, the first round of 20 Pathfinders were selected from 72 competitive bids from local authorities which all had more than one ward among the 10% most deprived in England. Indeed, we show below that we are able to achieve a good balance between the treatment and control groups across all chosen matching variables. However, in the discussion of the results we also consider further the implication of these pre-treatment trends for any bias in our DiD estimates.

4.3. Vector of controls

The estimated impact of Pathfinders could naïvely be estimated using an extended vector $\boldsymbol{X}_{it}$ as proposed in equation 2. We initially considered in this vector characteristics of the head of household, namely age, gender, marital status (living as a couple or not), log of income, education level (higher than A-level, A-level, O-level, none), employment status (employed, unemployed/out of labour market) as well as self-assessed health status in five categories (from very good to very poor) and self-reported satisfaction with life, with housing conditions and with neighbourhood (all three measured via a 7-point Likert scale). It also included characteristics of the household such as the number of children 'in different age groups', and the area such as the region of residence (northwest, London-South east, rest of England).

Furthermore, since the Pathfinders programme was mainly targeted at the most deprived areas of the country, we considered the deprivation level of the LSOA of residence as measured by the Townsend index (Townsend et al, 1988). This area level measure of material deprivation incorporates unemployment, car ownership, home ownership and household overcrowding. A higher Townsend score implies a greater degree of deprivation and areas can be ranked according to their score. Differences between the treated (IA) and candidate control areas (CA) are particularly marked (Table A1 in appendix). For example, there is a significant difference between the average Townsend score for the treated areas (4.05, SD: 3.03) compared to the large number of potential control areas, (0.127, SD: 3.37), with t-test p-value < 0.000).

Since we had a large number of candidate controls observations (relative to treated), we preceded the regression-based DiD estimation with a matching procedure to remove those differences by matching, selecting a set of pre-treatment controls and ensuring we only compared treated individuals to similar control individuals in similar areas.

4.4. Pre-treatment controls

The final choice of the variables to include in the propensity score matching (the vector $\boldsymbol{M}$) was guided by the use of stepwise Probit regressions. This suggested the inclusion of the following best predictors of the propensity score: the number of children in the household and the log of income of the head of household both observed at baseline (wave 10), the region of residence along with the Townsend score for the LSOA of residence. Additionally, we added all studied outcomes observed at baseline as well as reported housing problems, namely lack of light and leaky roof. The inclusion of baseline outcomes in the matching is a way of picking up unobserved characteristics that are not controlled for by the other covariates, hence we expect this to result in better matches between the control and treatment areas.

The scores were generated separately for the three alternative definitions of the treatment area (see Table A2 in Appendix); and all three scores were deemed to be of good quality, passing the unconfoundedness and overlap condition (also known as common support).[13]

The wave-specific difference-in-difference estimations after matching additionally included as potential confounding factors the following wave-specific characteristics of the head of household age, age squared, gender, marital status, education level, and employment status.

5. Results

Table 2 presents the estimates of the average effect of the treatment on the treated (ATT) from each of the linear probability model regressions (2) for the set of outcomes of interest, along with the corresponding standard errors and associated level of significance after matching; these are presented for all three alternative definitions of the treatment area. Focusing on the ATT estimates with matched controls, we mainly observe reductions in reported problems within 4 years after the Pathfinders programme was implemented with most significant results observed at wave 14, 3 years after the initial implementation. Individuals living in a Pathfinders area report significantly less problems with noise from the neighbours, noise from the streets, condensation, damp walls, and pollution. The ATT related to reported crime and vandalism are also negative however not statistically significant at standard levels. The largest magnitude of reduction is observed for reported noise, either from neighbours when intervention individuals are considered within 800 m of Pathfinders

Table 2. Difference-in-differences estimates by outcome and treatment radius after matching.

Variables	2002/3 Wave 12	2003/4 Wave 13	2004/5 Wave 14	2005/6 Wave 15	2006/7 Wave 16	2007/8 Wave 17	2008/9 Wave 18
Shortage of space							
LSOA	0.088	0.038	0.124**	0.117*	0.110$^{\$}$	0.152**	0.087
1000 m	0.275**	0.157	0.133	0.037	0.305***	0.342***	0.162*
800 m	0.242	0.394**	−0.014	0.000	0.217**	0.242	0.196*
Noise from neighbours							
LSOA	−0.029	−0.099	−0.026	−0.002	−0.081	0.108$^{\$}$	−0.011
1000 m	0.147	0.147	−0.063	0.056	0.105	0.140$^{\$}$	0.154$^{\$}$
800 m	−0.364***	−0.182	−0.225***	0.075	0.025	0.144	0.203$^{\$}$
Noise from the streets							
LSOA	−0.058	0.044	−0.099$^{\$}$	−0.012	−0.100$^{\$}$	−0.037	−0.011
1000 m	−0.147	0.255*	−0.204***	0.042	−0.029	0.009	0.077
800 m	−0.136	−0.045	−0.297***	−0.133	−0.042	0.265	−0.123
Condensation							
LSOA	0.018	−0.015	−0.016	−0.020	−0.014	0.145***	0.011
1000 m	−0.206$^{\$}$	−0.147	−0.042	−0.028	0.014	0.219***	0.077
800 m	−0.076	0.106	−0.152***	−0.092**	0.025	0.083	0.188*
Damp							
LSOA	−0.018	−0.035	−0.070$^{\$}$	−0.010	0.095**	0.064$^{\$}$	0.037
1000 m	0.039	0.000	−0.021	0.019	0.057	0.009	0.073
800 m	0.000	−0.091	−0.123***	−0.008	0.058	0.083	0.087
Pollution							
LSOA	−0.096	0.009	0.028	−0.035	0.076$^{\$}$	0.034	0.018
1000 m	0.118	0.059	−0.075	−0.120$^{\$}$	−0.038	0.000	0.043
800 m	−0.061	−0.030	−0.138***	−0.083	−0.008	0.068	0.109
Crime and vandalism							
LSOA	−0.123	0.012	−0.056	0.007	−0.055	0.012	−0.062
1000 m	−0.020	0.157	−0.096	0.069	0.043	0.000	0.026
800 m	−0.591***	−0.076	−0.065	0.158	0.192$^{\$}$	0.341	−0.007
General Health Score (0 (the least distressed) and 36 (the most distressed))							
LSOA	−0.190	−0.178	0.951	0.239	−0.300	0.221	0.283
1000 m	−2.461	−2.588	1.683$^{\$}$	−2.083$^{\$}$	−0.952	−0.702	0.154
800 m	−1.773	−2.409	1.652$^{\$}$	−2.125	−0.658	1.598	1.630
Frequency of talking to neighbours (4 – most days to 0 – never)							
LSOA	0.228	0.447**	0.148	0.040	0.107	−0.037	0.121
1000 m	0.186	−0.020	0.200	0.083	0.081	−0.075	0.385**
800 m	−0.136	0.121	0.283	0.233	0.492**	0.288	0.312

Significance levels: ***1%, **5%, $^{\$}$10%.
Models were estimated as Linear Probability Models for all outcomes.
Matched areas were identified via Propensity Score Matching based on the number of children in the household (kids_02, kids_34, kids_511, kids_1618), the log of income, the region of residence (northwest, London-South east, rest of England), the 2001 Townsend score for the LSOA of residence and the studied outcome at baseline.
The DiD regression additionally controlled for age, age2, female, married, education (a level, o level, none), and employment status (employed, unemployed/out of labour market) of the head of the household.

(-36 percentage points (pp.) one year after the intervention to −22 pp. three years after) or from the streets three 3 years after the implementation (-30 percentage points (pp.) when intervention individuals are considered within 800 m of Pathfinders postcodes to −10 pp. when within the LSOA of Pathfinders postcodes). The closer an individual lives to the intervention postcodes the largest is the observed reduction of reporting problems.

These housing-related outcomes were the targeted outcomes of the intervention and improvements were mainly observed for two to four years only. These benefits appear to be short-lived since five to six years after the programme, individuals living in intervention areas reported significantly more problems with noise, condensation, and dampness. The programme also appeared to increase reported problems with shortage of space over the following six years, with reported problems being mostly significant at each time point and progressively increasing to between +15 and +34 percentage points higher six years after the programme.

General health status was only significantly impacted by the Pathfinders programme after three years and four years. While after three years, the GHQ score increased by 1.6 signifying worse psychological health, it shows a significantly better score the following year with a decrease of 2.1. While some housing outcomes improved in response to the Pathfinders programme three years later, the gap in general health (which was at the advantage of Pathfinders areas) also closed between treated and controlled areas.

The ATT estimates related to the frequency of talking with neighbours are mostly positive across the years and when significant the ATT show an increase of +39 pp. to +49 pp. The Pathfinders programme shows long-term positive impact on the frequency of talking to neighbours, which was not an intended outcome.

We present the naïve ATT estimates for each outcome when only time, treatment (Pathfinders) and the time/treatment interaction were included in the model (labelled No Control) and the same ATT estimates when a set of controls variables were additionally included in the estimation in appendices A.3 to A.5; these sets of results are estimated on the original sample and do not include any matching procedure. The alternative treatment definitions (LSOA and radiuses of 800 and 1000 m) essentially tell the same story. The inclusion of controls typically produces an ATT estimate with similar magnitude and significance as compared to the naive specification without controls.

6. Discussion

Regarding the Pathfinders programme, we show that it had positive effects on some of the targeted outcomes including reported street noise, noise from the neighbours, pollution, house condensation and damp walls, and to a lesser extent crime and vandalism. The estimated effects were relatively large changes of between 10 and 40 percentage points. Most impact on the neighbourhood and lower reported housing problems were short-lived and not observed beyond 4 years. On the contrary, five to six years after the programme, people living in Pathfinders areas reported significantly higher problems with noise from the street and the neighbours, condensation, and damp walls. Furthermore, over the six years following the programme, individuals living in Pathfinders areas reported significantly higher problems with shortage of space. While the programme might have helped with overall living improvement related to cleaner houses and streets, it appears that the housing conditions remained relatively poor. The limited reports of improvement in the set of outcomes we considered could therefore be explained by insufficient programme investment to make sustainable changes. Similar results were observed for the New Deal for Communities where the programme was found to be effective in the early years but the investment was not sufficient to ensure sustainable changes (Beatty *et al.*, 2010). The absence of long-term impact could also illustrate adaptation effects to the expectations of the population since a neighbourhood renewal programme can raise expectations and exacerbate reports of problems from individuals.

We find that the programme also had positive effects on some non-intended outcomes such as frequency of talking with neighbours, which can be thought of as a measure of local social cohesion. The Pathfinders programme relied on community development and involved local communities and the voluntary sector to facilitate the neighbourhood renewal and this dimension may underlie our finding of increased frequency of communication with neighbours over the years following the programme. The impact of the Pathfinders programme on health outcomes was rather limited and this is in line with previous studies showing that housing refurbishment programmes are often less effective on health outcomes than person-targeted programmes (Thomson *et al.*, 2006).

The available data did not allow us to carry out an evaluation using extended pre- and post- intervention periods. Regarding the pre- period, the outcomes we consider are only added to later waves of the BHPS, and not observable earlier than wave 10. Regarding the post- period, the evolution of BHPS into Understanding Society led to a number of changes in the data collection and we could not extend the post-intervention period and investigate longer-term outcomes.

This causal evaluation of the Pathfinders programme also provides a case-study to illustrate the practical usefulness of combining secondary data and geographical identifiers for the evaluation of this type of ABI. There is potential transferability of the methods to other policy evaluations where geographically identifiable interventions could be combined with longitudinal individual and household surveys. The use of secondary data that were not collected for this purpose has several advantages for the evaluation of neighbourhood-based interventions. First, longitudinal household and individual surveys allow the investigation of long-run effects. Second, these data enable us to consider a broad range of individual outcomes that were intended as well as non-intended by the programme. They also allow the consideration of spill-over effects into an area wider than that initially targeted. Finally, secondary household survey data also generally includes a rich set of information to facilitate matching the treatment group with appropriate controls.

A disadvantage of the data is that the geographical identifiers required are very precise, and as a result data access can be challenging and will usually necessitate user licence applications and working via a secure lab environment. Furthermore, there might be a trade-off to make with sample size; our analysis here showed that we struggled with statistical power and sometimes lacked stable estimates since the number of individuals living in intervention areas remained quite small. A larger sample of individuals living in the intervention areas would have additionally allowed us to undertake an evaluation at other relevant aggregated levels (for example regions and population sub-groups) to shed more light on exactly who gains from neighbourhood-based interventions, as effects could be heterogeneous within areas or population sub-groups. It is worth stressing that the BHPS successor, the UK Household Longitudinal Study (UKHLS), has a much larger sample size (approximately 80,000 individuals) and thus it can provide more power for spatial analysis; these data may be very valuable in future evaluations of this type. Still, the use of such household surveys for an evaluation is limited to initiatives that are large enough to ensure that the survey includes a sufficient number of people in the intervention areas.

Another limitation of the data we used is to provide only individual- or household-level effects while area-based interventions may also seek wider system change via impact on local businesses, neighbourhood activities and institutions.

7. Conclusion

In this paper we have sought to illustrate the practical usefulness of combining secondary data and geographical identifiers for the evaluation of area-based interventions by evaluating the Pathfinders programme, which had never been causally assessed, as a case-study.

We find that the flagship £100 million Pathfinders programme, aimed at bringing the local community and local services together to tackle problems in the area and improve local services had positive effects on several of the targeted outcomes, including reduced reported noise from neighbours, noise from the streets, pollution, and to a lesser extent house condensation and damp. However, most of these positive effects were short-lived. We did however find a positive impact on the frequency of talking with neighbours which persisted into the longer term; and while this was not an intended programme outcome it is consistent with the logic model underlying the Pathfinders approach in relation to the predicted effects on community networks.

We also provide a number of key findings concerning the methodological implications of this case-study. Numerous area-based interventions are implemented across communities; however, their impacts are often multidimensional and multisectoral, and uncovering whether the intervention worked and for whom might be challenging. This case-study illustrates how secondary data that were not collected for this purpose may allow policy-makers to evaluate the direct and indirect impacts of an intervention using geographical identifiers and a quasi-experimental setting. From our experience, three main preconditions must be met prior to considering a similar methodological approach. First, both the interventions and the survey data include geographical identifiers. Second, the intervention can be interpreted as a quasi-experimental setting where the intervention was non-randomised. Third, intervened individuals or areas are sufficiently numerous to provide the analysis with an appropriate sample size.

Notes

1. https://www.gov.uk/government/collections/new-levelling-up-and-community-investments
2. Gibbons et al (2013) classified the evaluations according to the 5-point Maryland scale, which rates the strength of evidence (https://whatworksgrowth.org/resources/the-scientific-maryland-scale/). None of the spatial policy evaluations scored level 3 (or above), which is deemed the minimum standard for claiming that the programme caused the reported impact; the basic requirement for level 3 is that the study has a robust comparison group.
3. https://www.ukdataservice.ac.uk/get-data/how-to-access/accesssecurelab
4. Access to the data via the Secure Lab essentially involves registration for the UK Data Service, Secure Lab training, application for a licence to use the data for the stated purpose, and data access and analysis via a remote desktop setting.
5. Originally 20 Pathfinders were funded in Round 1 in 2001/2, with a further 15 added in Round 2 in 2003/4. Round 1 Pathfinders received the highest levels of funding. This study evaluates the Pathfinders as a whole.
6. These quasi-experimental methods have become popular to evaluate the effects of a policy change in the absence of a randomized controlled trial, and also to capture the impact of exogeneous events. See Cai (2021) for a recent summary of new studies using DiD methods.
7. Pathfinders were selected from competitive bids from local authorities who had been invited to express interest. Invitations were targeted at those authorities that had more than one ward among the 10% most deprived in England.
8. Matching without replacement means that the same CA observation can only be used once as a match.
9. The UK uses a system of alphanumeric postcodes to identify postal delivery areas. A postcode is made up of four components (area, district, sector and unit). The area code is the letter prefix that denotes city, for example B for Birmingham and LS for Leeds. There are 124 postcode areas in the UK, and approximately 1.7 million postcodes; each on average covers about 14 houses.
10. One of the co-authors of this paper was a lead investigator on the original Pathfinders evaluation for DCLG and provided the postcodes that identify the intervention areas.
11. http://geoconvert.mimas.ac.uk/
12. Only 4 Pathfinders are within the ‘Small Town and Fringe areas’ category.
13. The unconfoundness assumption signifies that the assignment to the treatment is random conditional on the variables M. The overlap condition ensures that there is sufficient overlap in the characteristics of the treated and untreated units to find adequate matches (or a *common support*).

Acknowledgments

We thank participants at the Leeds Institute for Data Analytics (LIDA) research seminars and two anonymous reviewers for useful comments on previous version of this paper. We are grateful to Rita Santos, Effie Kesidou and Luke Munford for help with geo-spatial coding and remote data access. Most of this work was carried out while Sandy Tubeuf visited the University of Sheffield and the University of Cambridge; the research visits were made possible thanks to the Leverhulme Research Fellowship RF-2015-270 and funding from InstEAD at the University of Sheffield.

BHPS is an initiative by the Economic and Social Research Council, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by the National Centre for Social Research. National Grid Reference information was accessed via the Secure Data Service. Neither the original data creators, depositors or funders bear responsibility for the further analysis or interpretation of the data presented in this study. The data used for this research were provided by the University of Essex, Institute for Social and Economic Research. (2018). British Household Panel Survey, Waves 1-18,

1991–2009: Secure Access, National Grid Reference (Easting, Northing, OSGRDIND). [data collection]. 4th Edition. UK Data Service. SN: 6340, http://doi.org/10.5255/UKDA-SN-6340-3

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The Leverhulme Trust Research Fellowship RF-2015-270.

ORCID

Jennifer Roberts http://orcid.org/0000-0003-2883-7251
Sandy Tubeuf http://orcid.org/0000-0001-9001-1157
Peter Tyler http://orcid.org/0000-0002-9299-5396

References

Aarland, K., Osland, L. & Gjestland, A. (2017) Do area-based intervention programs affect house prices? A quasi-experimental approach, *Journal of Housing Economics*, 37, pp. 67–83.

Beatty, C., Brennan, A., Mi, F., Lawless, P., Tyler, P., Warnock, C. & Willson, I. (2010) *The New Deal for Communities Programme: Assessing impact and value for money. Final report* – Volume 6. Project report. Communities and Local Government.

Bertrand, M., Duflo, E. & Mullainathan, S. (2004) How much should We trust differences-in-Differences estimates?, *Quarterly Journal of Economics*, 119, pp. 249–275.

Blundell, R. & Costa-Dias, M. (2000) Evaluation methods for Non-Experimental *data*, *Fiscal Studies*, 21, pp. 427–468.

Cai, C. (2021) Literature on Recent Advances in Applied Micro Methods. Github resources. https://christinecai.github.io/PublicGoods/applied_micro_methods.pdf

Cameron, C. & Trivedi, P. K. (2005) *Microeconometrics: Methods and Applications* (Cambridge, UK: Cambridge University Press).

Department for Communities and Local Government (2006) *Neighbourhood Management: An Overview of the 2003 and 2006 Round 1 Pathfinder Household Survey. Research Report 28*

Department for Communities and Local Government (2008) *Neighbourhood Management Pathfinders: Final Evaluation Report*.

Department for Communities and Local Government (2011) *A plain English guide to the Localism Bill*.

Forrest, R. & Kearns, A. (2001) Social cohesion, social capital and the neighbourhood, *Urban Studies*, 38, pp. 2125–2143.

Galster, G. C. & Hedman, L. (2013) Measuring neighbourhood effects non-experimentally: How much do alternative methods matter?, *Housing Studies*, 28, pp. 473–498.

Galster, G. C. & Santiago, A. M. (2015) Evaluating the potential of a natural experiment to provide unbiased evidence of neighborhood effects on health, *Health Services and Outcomes Research Methodology*, 15, pp. 99–135.

Galster, G., Temkin, K., Walker, C. & Sawyer, N. (2004) Measuring the impacts of community development initiatives: a new application of the adjusted interrupted time-series method, *Evaluation Review*, 28, pp. 502–538.

Gibbons, S., McNally, S. & Overman, H. G. (2013) Review of government evaluations: A report for the UK National Audit Office. http://www.nao.org.uk/wpcontent/uploads/2013/12/LSE-Review-of-selection-of-evaluations-withappendices1.pdf.

Gibbons, S., Overman, H. & Sarvimäki, M. (2021) The local economic impacts of regeneration projects: Evidence from UL's single regeneration budget, *Journal of Urban Economics*, 122, pp. 103315.

Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O. & Rutter, C. (1997) The validity of two versions of the GHQ in the WHO study of mental illness in general health care, *Psychological Medicine*, 27, pp. 191–197.

Heckman, J., Ichimura, H. & Todd, P. (1997) Matching as an econometric evaluation estimator: evidence from evaluating a job training programme, *Review of Economic Studies*, 64, pp. 605–654.

Heckman, J., Ichimura, H., Smith, J. & Todd, P. (1998) Characterizing selection bias using experimental data, *Econometrica*, 66, pp. 1017–1098.

HM Government (1998) 'Bringing Britain together: a national strategy for neighbourhood renewal', Report by the Social Exclusion Unit Stationery Office, Cm 4045, ISBN 0101404522.

Imbens, G. W. & Wooldridge, J. M. (2009) Recent developments in the econometrics of program evaluation, *Journal of Economic Literature*, 47, pp. 5–86.

MRC (Medical Research Council) (2006) Developing and evaluating complex interventions: new guidance. www.mrc.ac.uk/complexinterventionsguidance

O'Reilly. (2007) Comment on Rhodes et al. (2005): Some further thoughts on assessing the effects of area-based initiatives on local outcomes, *Urban Studies*, 44, pp. 1145–1153.

Papoutsaki, D., Buzzeo, J., Gray, H., Williams, M., Cockett, J., Akehurst, G., Alexander, K., Newton, B. & Pollard, E. (2020) Moving out to move on - Understanding the link between migration, disadvantage and social mobility. Social Mobility Commission. Research Report. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/902943/Moving_out_to_move_on_report.pdf

Petticrew, M., Cummins, S., Ferrell, C., Findlay, A., Higgins, C., Hoy, C., Kearns, A. & Sparks, L. (2005) Natural experiments: an underused tool for public health?, *Public Health*, 119, pp. 751–757.

Ploegmakers, H. & Beckers, P. (2015) Evaluating urban regeneration: an assessment of the effectiveness of physical regeneration initiatives on run-down industrial sites in The Netherlands, *Urban Studies*, 52, pp. 2151–2169.

Rhodes, J., Tyler, P. & Brennan, A. (2005) Assessing the effect of area based initiatives on local area outcomes: Some thoughts based on the national evaluation of the single regeneration budget in England, *Urban Studies*, 42, pp. 1919–1946.

Romero, R. G. & Noble, M. (2008) Evaluating england's 'new deal for communities' programme using the difference-in-difference method, *Journal of Economic Geography*, 8, pp. 759–778.

Romero, R. G. (2009) Estimating the impact of england's area-based intervention 'new deal for communities' on employment, *Regional Science and Urban Economics*, 39, pp. 323–331.

Ruijsbroek, A., Wong, A., Kunst, A. E., van den Brink, C., van Oers, H. A. M., Droomers, M. & Stronks, K. (2017) The impact of urban regeneration programmes on health and health-related behaviour: Evaluation of the dutch district approach 6.5 years from the start, *PloS One*, 12, pp. e0177262.

Schkade, D. A. & Kahneman, D. (1998) Does living in California make people happy? A focusing illusion in judgments of life satisfaction, *Psychological Science*, 9, pp. 340–346.

Stuart, E. A. (2010) Matching methods for causal inference: a review and look forward, *Statistical Science : A Review Journal of the Institute of Mathematical Statistics*, 25, pp. 1–21.

Thomson, H., Atkinson, R., Petticrew, M. & Kearns, A. (2006) Do urban regeneration programmes improve public health and reduce health inequalities? A synthesis of the evidence from UK policy and practice (1980–2004). *Journal of Epidemiology and Community Health*, 60, pp. 108–115.

Townsend, P., Phillimore, P. & Beattie, A. (1988) *Health and Deprivation: Inequality and the North* (London: Routledge).

Wickes, R., Zahnow, R., Corcoran, J. & Hipp, J. R. (2019) Neighbourhood social conduits and resident social cohesion, *Urban Studies*, 56, pp. 226–248.

Table A1. Descriptive statistics for each outcome for the treated (LSOA) and potential control areas.

	LSOA treated areas		Rest of the sample (potential control areas)	
	N	Mean (SD)	N	Mean (SD)
Shortage space				
Wave 10	174	0.24 (0.43)	6,207	0.23 (0.42)
Wave 11	158	0.25 (0.44)	6,165	0.23 (0.42)
Wave 12	139	0.26 (0.44)	5,230	0.24 (0.43)
Wave 13	135	0.26 (0.44)	5,158	0.22 (0.41)
Wave 14	128	0.27 (0.45)	5,028	0.22 (0.42)
Wave 15	109	0.28 (0.45)	5,031	0.22 (0.41)
Wave 16	117	0.34 (0.48)	4,993	0.21 (0.41)
Wave 17	115	0.28 (0.45)	4,786	0.21 (0.40)
Wave 18	110	0.24 (0.43)	4,683	0.20 (0.40)
Noise from neighbours				
Wave 10	174	0.14 (0.35)	6,207	0.13 (0.34)
Wave 11	158	0.13 (0.33)	6,165	0.13 (0.34)
Wave 12	139	0.15 (0.36)	5,230	0.12 (0.33)
Wave 13	135	0.15 (0.36)	5,158	0.13 (0.33)
Wave 14	128	0.09 (0.29)	5,028	0.12 (0.32)
Wave 15	109	0.14 (0.35)	5,030	0.13 (0.33)
Wave 16	117	0.24 (0.43)	4,993	0.13 (0.34)
Wave 17	115	0.23 (0.42)	4,786	0.12 (0.32)
Wave 18	110	0.16 (0.37)	4,683	0.13 (0.33)
Street noise				
Wave 10	174	0.29 (0.45)	6,207	0.19 (0.39)
Wave 11	158	0.23 (0.42)	6,165	0.20 (0.40)
Wave 12	139	0.34 (0.47)	5,230	0.19 (0.39)
Wave 13	135	0.21 (0.41)	5,158	0.20 (0.40)
Wave 14	128	0.16 (0.36)	5,026	0.18 (0.38)
Wave 15	109	0.21 (0.41)	5,031	0.18 (0.38)
Wave 16	117	0.22 (0.42)	4,993	0.18 (0.38)
Wave 17	115	0.28 (0.45)	4,788	0.15 (0.36)
Wave 18	110	0.23 (0.42)	4,683	0.14 (0.35)
Condensation				
Wave 10	174	0.22 (0.42)	6,205	0.14 (0.35)
Wave 11	158	0.16 (0.37)	6,161	0.14 (0.34)
Wave 12	139	0.17 (0.38)	5,230	0.13 (0.33)
Wave 13	135	0.12 (0.32)	5,155	0.09 (0.29)
Wave 14	128	0.08 (0.27)	5,027	0.09 (0.29)
Wave 15	109	0.17 (0.37)	5,030	0.09 (0.28)
Wave 16	117	0.19 (0.39)	4,989	0.08 (0.27)
Wave 17	115	0.14 (0.35)	4,782	0.08 (0.27)
Wave 18	110	0.14 (0.34)	4,683	0.08 (0.28)
Damp				
Wave 10	174	0.12 (0.33)	6,203	0.07 (0.25)
Wave 11	158	0.10 (0.30)	6,163	0.09 (0.28)
Wave 12	139	0.12 (0.32)	5,227	0.07 (0.26)
Wave 13	135	0.12 (0.32)	5,156	0.06 (0.23)
Wave 14	128	0.07 (0.26)	5,026	0.07 (0.25)
Wave 15	109	0.17 (0.37)	5,030	0.06 (0.23)
Wave 16	117	0.10 (0.30)	4,992	0.05 (0.23)
Wave 17	115	0.08 (0.27)	4,787	0.06 (0.25)
Wave 18	110	0.11 (0.31)	4,683	0.07 (0.26)
Pollution				
Wave 10	174	0.18 (0.38)	6,205	0.09 (0.29)
Wave 11	158	0.15 (0.36)	6,165	0.08 (0.28)
Wave 12	139	0.12 (0.32)	5,229	0.08 (0.27)
Wave 13	135	0.07 (0.25)	5,156	0.08 (0.27)
Wave 14	128	0.09 (0.28)	5,026	0.08 (0.27)
Wave 15	109	0.19 (0.40)	5,030	0.08 (0.26)
Wave 16	117	0.14 (0.35)	4,984	0.07 (0.25)
Wave 17	115	0.11 (0.32)	4,787	0.07 (0.25)

(*Continued*)

Table A1. (Continued)

	LSOA treated areas		Rest of the sample (potential control areas)	
	N	Mean (SD)	N	Mean (SD)
Wave 18	110	0.15 (0.35)	4,683	0.06 (0.23)
Crime, vandalism				
Wave 10	174	0.30 (0.46)	6,203	0.21 (0.40)
Wave 11	158	0.24 (0.43)	6,158	0.22 (0.42)
Wave 12	139	0.32 (0.47)	5,227	0.21 (0.41)
Wave 13	135	0.22 (0.42)	5,153	0.21 (0.40)
Wave 14	128	0.19 (0.39)	5,028	0.19 (0.39)
Wave 15	109	0.26 (0.44)	5,025	0.19 (0.39)
Wave 16	117	0.26 (0.44)	4,991	0.18 (0.38)
Wave 17	115	0.22 (0.41)	4,774	0.16 (0.36)
Wave 18	110	0.31 (0.46)	4,679	0.14 (0.35)
General Health Score (0 (the least distressed) and 36 (the most distressed))				
Wave 10	172	24.38 (6.26)	6,108	24.64 (5.46)
Wave 11	156	25.01 (5.75)	6,091	24.76 (5.38)
Wave 12	137	23.90 (6.23)	5,179	24.86 (5.40)
Wave 13	132	25.06 (4.92)	5,100	24.86 (5.48)
Wave 14	123	24.62 (5.46)	4,965	24.77 (5.45)
Wave 15	103	24.17 (5.90)	4,953	24.71 (5.42)
Wave 16	116	24.84 (5.15)	4,910	24.74 (5.47)
Wave 17	113	24.72 (5.74)	4,731	24.81 (5.60)
Wave 18	109	23.53 (6.34)	4,603	24.54 (5.54)
Frequency of talking to neighbours (4 – most days to 0 – never)				
Wave 10	174	2.89 (1.05)	6,218	2.99 (1.01)
Wave 11	158	3.04 (1.06)	6,216	3.01 (1.02)
Wave 12	140	3.05 (1.05)	5,270	3.00 (1.01)
Wave 13	135	3.09 (0.97)	5,199	2.98 (1.01)
Wave 14	128	3.00 (0.99)	5,057	2.95 (1.04)
Wave 15	110	3.09 (0.98)	5,064	2.97 (1.03)
Wave 16	119	3.10 (1.08)	5,012	2.95 (1.06)
Wave 17	115	3.13 (1.04)	4,821	2.97 (1.02)
Wave 18	112	3.10 (0.95)	4,731	2.91 (1.08)

Note: outcomes are for individuals in the respective treated and potential control areas.

Table A2. Probability to be in a treated area at wave 10.

	Within LSOA			Within 1000 m			Within 800 m		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Variables	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
Age	−0.008	−0.007	−0.003	−0.009	−0.008	−0.003	0.003	−0.002	0.006
Age squared	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Female	−0.009	−0.015	−0.019	0.031	0.041	0.027	−0.090	−0.049	−0.075
Married	0.045	0.065	0.036	0.045	0.081	0.049	0.125	0.234*	0.191
Kids_02	−0.202	−0.217	−0.199	−0.150	−0.154	−0.170	−0.438	−0.476	−0.498
Kids_34	0.183	0.186	0.193	0.129	0.137	0.124	.	.	.
Kids_511	−0.059	−0.059	−0.056	−0.113	−0.098	−0.109	−0.273	−0.299	−0.341$
Kids_1215	−0.232$	−0.226$	−0.264**	−0.355**	−0.350**	−0.403**	−0.185	−0.163	−0.205
Kids_1618	0.446***	0.457***	0.503***	0.517***	0.523***	0.529***	0.548**	0.554**	0.594***
Log of income	0.167***	0.169***	0.175***	0.099	0.097	0.093	0.208**	0.185*	0.190*
Education A level	−0.077	−0.049	−0.050	−0.142	−0.098	−0.080	−0.334	−0.281	−0.254
Education O level	−0.131	−0.114	−0.122	−0.086	−0.062	−0.041	−0.196	−0.152	−0.099
Education none	−0.033	0.014	0.013	0.030	0.079	0.093	−0.149	−0.084	−0.035
Unemployed	−0.189	−0.241	−0.271	−0.133	−0.217	−0.272	−0.221	−0.233	−0.323
Inactive	−0.089	−0.059	−0.063	−0.083	−0.059	−0.047	0.129	0.152	0.181
North East	0.494***	0.519***	0.571***	−0.183	−0.166	−0.092	−0.487	−0.494**	−0.434*
North West	−0.116	−0.136	−0.110	−0.177	−0.204	−0.151	−0.391***	−0.403***	−0.366**
Yorks & Humber	−0.555***	−0.554***	−0.535***	−0.609***	−0.601***	−0.547***	−0.999***	−1.002***	−0.931***
East Midlands	−0.820***	−0.830***	−0.766***	−0.716***	−0.727***	−0.650**	−0.651**	−0.657**	−0.571**
West Midlands	−0.600***	−0.608***	−0.560***	−1.194***	−1.210***	−1.119***	.	.	.
South East	−0.068	−0.069	−0.025	−0.274$	−0.266$	−0.188	.	.	.
South West	0.340***	0.318***	0.377***	0.604***	0.592***	0.697***	0.136	0.104	0.246
Shortage of space			−0.098			0.116			0.140
Noise from neighbours			−0.232*			−0.211			−0.037
Street noise			0.098			0.072			0.073
Condensation			0.192*			0.046			−0.178
Damp			0.086			−0.044			0.117
Lack of light			−0.102			−0.213			−0.099
Leaky_roof			0.096			0.399**			0.419
Pollution			0.276**			0.363***			0.532
Crime & vandalism			−0.007			0.059			−0.037
GHQ		−0.012	−0.012		−0.001	−0.001		0.010	0.010
Self-assessed health		0.085$	0.083$		0.052	0.043		0.076	0.075
Life satisfaction		0.016	0.017		−0.040	−0.036		−0.098$	−0.096$
Like neighbourdhood		0.139	0.150		0.123	0.108		0.108	0.087
Satisfaction with house		−0.012	−0.011		0.028	0.049		−0.013	0.019
IMD_2010	0.024***	0.025***	0.024***	0.026***	0.028***	0.026***	0.033***	0.034***	0.033***
Number of obs.	6,318	6,109	6,070	6,318	6,109	6,070	4,307	4,157	4,132
Pseudo R2	0.1357	0.1409	0.1512	0.1737	0.1761	0.1927	0.1898	0.2003	0.2258

Coeff.: Estimated coefficient with Probit models. Significance levels: ***1%, **5%, $10%.

Table A3. Estimated average treatment effects for neighbourhood (LSOA) treated areas using a naïve DiD regression model before matching.

Variables	Wave 12	Wave 13	Wave 14	Wave 15	Wave 16	Wave 17	Wave 18
Shortage of space							
No control	0.007	0.029	0.037	0.045	0.120***	0.059	0.022
With controls	0.015	0.037	0.031	0.026	0.124***	0.060	0.029
Noise from neighbours							
No control	0.020	0.012	−0.031	0.002	0.100**	0.100**	0.026
With controls	0.007	0.008	−0.039	−0.008	0.088**	0.081**	0.015
Noise from the streets							
No control	0.057	−0.076$	−0.118**	−0.062	−0.049	0.028	−0.007
With controls	0.047	−0.084$	−0.135***	−0.069	−0.062	0.003	−0.017
Condensation							
No control	−0.035	−0.054	−0.097***	−0.004	0.028	−0.023	−0.030
With controls	−0.043	−0.057	−0.105***	−0.017	0.025	−0.027	−0.032
Damp							
No control	−0.008	0.008	−0.051$	0.055	−0.003	−0.038	−0.017
With controls	−0.016	0.006	−0.055$	0.037	−0.003	−0.042	−0.017
Pollution							
No control	−0.050	−0.099***	−0.080**	0.030	−0.017	−0.043	0.001
With controls	−0.053	−0.102***	−0.090***	0.032	−0.019	−0.044	0.003
Crime and vandalism							
No control	0.010	−0.075$	−0.092**	−0.022	−0.011	−0.030	0.074
With controls	0.005	−0.077**	−0.102***	−0.042	−0.032	−0.055	0.067
General Health Score (0 (the least distressed) and 36 (the most distressed))							
No control	−0.695	0.467	0.116	−0.275	0.392	0.175	−0.747
With controls	−0.739	0.335	0.223	−0.146	0.473	0.268	−0.750
Frequency of talking to neighbours (4 – most days to 0 – never)							
No control	0.144$	0.204**	0.148	0.222**	0.248**	0.255**	0.288***
With controls	0.149$	0.233**	0.118	0.235**	0.193$	0.233**	0.285***

Models were estimated as Linear Probability Models of binary outcomes.
Significance levels: ***1%, **5%, $10%.
No controls specification only includes time, treatment and treatment × time.
With controls specification includes age, age2, female, married, number of children (kids_02, kids_34, kids_511, kids_1618), log of income, education (a level, o level, none), employment status (employed, unemployed/out of labour market), and region (northwest, London-South east, rest of England), all as observed at the same wave of the studied outcome, as well as the Townsend index in 2001.

Table A4. Estimated average treatment effects for 1000 m treated areas using a naïve DiD regression model before matching.

Variables	Wave 12	Wave 13	Wave 14	Wave 15	Wave 16	Wave 17	Wave 18
Shortage of space							
No control	0.026	0.019	−0.060	0.051	0.121**	0.034	0.007
With controls	0.021	0.013	−0.081	0.027	0.114$^{\$}$	0.032	0.003
Noise from neighbours							
No control	0.032	0.064	0.031	0.048	0.190***	0.128**	0.072
With controls	0.008	0.057	0.024	0.034	0.172***	0.111**	0.061
Noise from the streets							
No control	0.068	−0.109**	−0.112**	−0.001	−0.022	0.057	0.042
With controls	0.041	−0.132**	−0.142**	−0.017	−0.041	0.024	0.026
Condensation							
No control	−0.030	−0.034	−0.076$^{\$}$	0.008	0.067	0.045	0.044
With controls	−0.053	−0.048	−0.093**	−0.007	0.054	0.035	0.031
Damp							
No control	0.015	−0.002	0.029	0.118**	0.018	0.021	0.018
With controls	−0.002	−0.012	−0.036	0.100*	0.011	−0.027	0.011
Pollution							
No control	−0.110***	−0.122***	−0.102**	−0.032	−0.009	−0.091**	−0.006
With controls	−0.120	−0.128***	−0.117***	−0.033	−0.011	−0.089**	−0.004
Crime and vandalism							
No control	0.064	−0.102**	−0.072	−0.031	0.045	−0.011	0.015
With controls	0.030	−0.133***	−0.097*	−0.069	0.010	−0.046	−0.002
General Health Score (0 (the least distressed) and 36 (the most distressed))							
No control	−1.318	1.257**	−0.260	−0.305	−0.382	0.208	−1.237
With controls	−1.185	1.233$^{\$}$	0.071	−0.050	−0.137	0.383	−1.151
Frequency of talking to neighbours (4 – most days to 0 – never)							
No control	0.120	0.414***	0.229$^{\$}$	0.318**	0.262$^{\$}$	0.310**	0.481***
With controls	0.096	0.391***	0.134	0.243$^{\$}$	0.145	0.258$^{\$}$	0.467***

Models were estimated by linear probability models of binary outcomes.
Significance levels: ***1%, **5%, $^{\$}$10%.
No controls specification only includes time, treatment and treatment × time.
With controls specification includes age, age2, female, married, number of children (kids_02, kids_34, kids_511, kids_1618), log of income, education (a level, o level, none), employment status (employed, unemployed/out of labour market), and region (northwest, London-South east, rest of England), all as observed at the same wave of the studied outcome, as well as the Townsend index in 2001.

Table A5. Estimated average treatment effects for 800 m treated areas using a naïve DiD regression model before matching.

Variables	Wave 12	Wave 13	Wave 14	Wave 15	Wave 16	Wave 17	Wave 18
Shortage of space							
No control	0.011	0.052	−0.081	0.108	0.145$	0.058	−0.007
With controls	0.010	0.061	−0.066	0.123$	0.181**	0.085	0.024
Noise from neighbours							
No control	0.050	0.077	0.018	0.045	0.145**	0.090	0.112
With controls	0.025	0.071	0.026	0.043	0.158**	0.093	0.125$
Noise from the streets							
No control	0.041	−0.170***	−0.130$	−0.007	−0.067	−0.041	−0.052
With controls	0.019	−0.188***	−0.153**	−0.008	−0.072	−0.051	−0.052
Condensation							
No control	−0.002	−0.043	−0.100**	−0.022	0.101	0.003	0.014
With controls	−0.033	−0.061	−0.116**	−0.035	0.097	−0.001	0.014
Damp							
No control	0.052	−0.009	−0.055	0.103	0.085	−0.048	0.008
With controls	0.029	−0.022	−0.063	0.073	0.087	−0.050	0.011
Pollution							
No control	−0.169	−0.236***	−0.140**	−0.045	0.045	−0.154**	−0.014
With controls	−0.167	−0.234***	−0.144**	−0.029	0.059	−0.145**	−0.002
Crime and vandalism							
No control	0.001	−0.130**	−0.050	0.007	0.166**	−0.084	0.013
With controls	−0.012	−0.154**	−0.045	0.000	0.149$	−0.088	0.019
General Health Score (0 (the least distressed) and 36 (the most distressed))							
No control	−0.883	1.477$	0.084	0.188	−0.207	0.519	0.130
With controls	−0.726**	1.424$	0.121	0.192	−0.101	0.688	0.238
Frequency of talking to neighbours (4 – most days to 0 – never)							
No control	0.079	0.397**	0.150	0.252	0.271	0.329	0.494**
With controls	0.067	0.387**	0.108	0.235	0.166	0.257	0.404**

Models were estimated by linear probability models of binary outcomes.
Significance levels: ***1%, **5%, $10%.
No controls specification only includes time, treatment and treatment × time.
With controls specification includes age, age2, female, married, number of children (kids_02, kids_34, kids_511, kids_1618), log of income, education (a level, o level, none), employment status (employed, unemployed/out of labour market), and region (northwest, London-South east, rest of England), all as observed at the same wave of the studied outcome, as well as the Townsend index in 2001.