

Signs of Safety Evaluation Extension

Intervention Developer	MTM (http://munroturnellmurphy.com/)
Delivery Organisations	Bexley, Bristol, Leicestershire, Lincolnshire, London Borough of Brent, Norfolk, Suffolk, West Sussex, Wokingham
Evaluator	Health Social Care Workforce Research Unit - King's College London, What Works for Children's Social Care
Principal Investigator	Michael Sanders
Protocol Author(s)	Vicky Clayton, Michael Sanders, Eva Schoenwald, Alix Leroy
Type of Trial	Matched difference-in-difference
Age or Status of Participants	Treated: <18 years for children and parents
Number of Participating Local Authorities	9 treatment + up to 18 control
Number of Children and Families	Up to 710,000 ¹ children from 9 treatment local authorities and up to 18 control local authorities
Primary Outcome(s)	Duration of assessments, Likelihood of initial child protection conference
Secondary Outcome(s)	Likelihood of cases that received an NFA to be re-referred, likelihood of cases to be re-referred and result in a CPP or CLA, increases in kinship care vs non-kinship care

¹ Number of assessments or referrals between 2011 and 2018 for selected treatment and control local authorities. This number is likely to overestimate the actual treatment and control groups as one child can be assessed several times.

Table of contents

Background and Problem Statement	3
Intervention and Theory of Change	3
Description of Intervention	3
Impact Evaluation	4
Research questions	4
Design	4
Matching	6
Local Authority Matching	6
Individual-level Matching	9
Participants	10
Sample size / MDES calculations	10
Outcome measures	11
Analysis plan	11
Covariates	13
Handling missing data	14
Sensitivity	15
Ethics & Participation	16
Registration	16
Data protection	16
Personnel	21
Appendix	21

Background and Problem Statement

Signs of Safety (SofS) is an approach to child protection that was developed in Australia in the 1990s. It draws on Solution-Focused Brief Therapy and encourages the social worker to work collaboratively with parents / carers, and assess risks and strengths in the family. It has since been adopted across North America, Australasia and Europe and is a popular practice model in England, with a 2017 Health Social Care Workforce Research Unit survey suggesting that at least 42 local authorities used “pure SofS” and 49 used “some form of SofS” as their practice model (26 non-responses to the survey mean that these numbers are conservative estimates).

The Munro, Turnell and Murphy Child Protection Consulting (MTM) was supported by Wave 1 of the Department for Education Innovation Programme to work with 10 pilot local authorities in England to implement SofS. The 10 pilot local authorities were evaluated by Health Social Care Workforce Research Unit - King's College London (HSCWRU)². HSCWRU is currently evaluating the SofS with 8 of the same local authorities plus an additional local authority in Wave 2 of the Innovation Programme. What Works for Children’s Social Care (WWCSC) is complementing the HSCWRU’s evaluation in this phase through conducting a difference-in-difference analysis. The aim of this analysis is to obtain quasi-causal estimates of the effect of SofS on the duration of assessments, the likelihood of an initial child protection conference³, the likelihood of cases designated as NFA to be re-referred, the likelihood of cases to be re-referred and result in a child protection plan (CPP) or the child being looked after (LAC), and the share of kinship care vs non-kinship care. Quasi-causal estimates are useful to get to an accurate picture of the effect of SofS and so can help decision-makers in local authorities make decisions on whether or not to invest in SofS. This work builds on WWCSC’s systematic review of SofS focusing on reducing entry to care⁴.

Intervention and Theory of Change

Description of Intervention

SofS was designed as a whole system intervention for child protection that intends to investigate risks and strengths in the family and child’s environment. It uses a collaborative approach between social workers and families to strengthen and safeguard children in their home. The program uses structured and organised processes such as Solution-Focused Brief Therapy methods (a solution- and future-focused approach to clinical practice developed in 1970) in order to be effective. It is based on 3 key concepts:

- Working relationship - The programme emphasises the importance of a constructive approach to child protection rather than a paternalistic one. The social worker builds and sustains a relationship of trust with the family and through the use of processes and tools, the family and practitioners address in partnership the situations of child abuse and maltreatment.

² Baginsky, M., Moriarty, J., Manthorpe, J., Beecham, J. & Hickman, B. (2017, July). Evaluation of Signs of Safety in 10 pilots: Research report. *Social Care Workforce Research Unit, King’s College London / Department for Education*.

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/625376/Evaluation_of_Signs_of_Safety_in_10_pilots.pdf

³ Although it is possible to have multiple initial child protection conferences in one year, we measure this as a binary because having multiple is unusual, and whether they have an initial child protection conference captures what we are interested in.

⁴ Sheehan, L., O’Donnell, C., Brand, S. L., Forrester, D., Addiss, S., El-Banna, A., Kemp, A., Nurmatov, U. (2018). Signs of Safety: Findings from a mixed-methods systematic review focussed on reducing the need for children to be in care. <https://whatworks-csc.org.uk/evidence/evidence-store/intervention/signs-of-safety/>

- Thinking critically, fostering a stance of inquiry - The practitioner must avoid defining a truth on a given situation, think critically and constantly challenge their view in order to minimise errors. It requires the social worker to regularly review the balance of strengths and dangers to avoid self-reinforcing bias which may perpetuate false beliefs.
- Locating grand aspirations - Capturing learning that arises from lived experience by documenting practitioner and family experience of good practice on the ground with complex and challenging case studies.

For a full overview of the intervention and the theory of change, please see the Wave 1 evaluation report (Baginsky et al., 2017: 9) or the Signs of Safety website (<https://www.signsofsafety.net/signs-of-safety/>).

Impact Evaluation

Research questions

Although there are a range of possible outcomes that could be assessed, we focus on only two primary outcomes to avoid problems associated with multiple comparisons. These two outcome variables were selected on the basis that they have shown promise in the Wave 1 evaluation and Wave 2 interim findings, and in discussions with HSCWRU. These are:

- Does Signs of Safety reduce the duration of assessments?
- Does Signs of Safety reduce the likelihood of an initial child protection conference?

Based on discussions with MTM to better understand SofS's theory of change, in addition to the primary outcomes, we will evaluate three secondary outcomes, namely:

- Does Signs of Safety reduce the rate of re-referrals of cases that have previously been assessed as NFA?
- Does Signs of Safety affect the rate of re-referrals that lead to a child protection plan (CPP) or to a child becoming looked after (LAC)?
- Does Signs of Safety increase the rate of kinship care and lower non-kinship care?

Design

The design of the analysis is a difference-in-difference (DiD) design. Six of the initial 10 pilots had begun SofS prior to the start of the Wave 1 evaluation, and all of the nine pilots in Wave 2 had begun SofS prior to the start of Wave 2. For this reason, a methodology was required that allowed a quasi-causal estimation for interventions already underway. The unit of analysis is at an individual level to optimise the power to detect an effect within the constraints of the project.

One of the primary outcomes is the duration of assessment. Both the Wave 1 report and the interim findings from Wave 2 found statistically significant differences between SofS sites and their Statistical Nearest Neighbours (SNNs) in terms of assessment duration and the proportion of assessments starting and finishing on the same day ($p \leq 0.05$). We note that requirements on the time taken to complete an assessment changed in 2015, and that for this reason comparisons across time are not recommended.⁵ However, in difference-in-difference we are comparing the change over time in the treatment local authority with the

⁵ E. R. Munro, J. Stone (2014). The impact of more flexible assessment practices in response to the Munro Review of Child Protection. August 2014, Research Report, DfE. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/345476/DFE-RR376 - The effect of flexible assessment practices follow up.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/345476/DFE-RR376_-_The_effect_of_flexible_assessment_practices_follow_up.pdf)

change in comparator local authorities (“control group”) and not just the duration of assessment before and after the intervention. We only include assessments completed in the time period (as otherwise we are without duration data).

The other primary outcome is the likelihood of an initial child protection conference (ICPC) , the individual-level version of an outcome analysed in the Wave 1 evaluation (the local authorities’ *rate* of ICPCs). SofS emphasises that it is a safety-focused approach and uses a family’s network to assume more responsibility for the safety of the child. A lower likelihood of an initial child protection conference suggests that there should be less reasonable cause to suspect that a child is suffering or is likely to suffer significant harm.

We also analyse secondary outcomes complementing our primary outcomes mentioned above. This analysis takes account of additional fundamental goals SofS aims to achieve. The secondary outcomes are the proportion of cases which received a NFA designation that are re-referred once or more within a 6-month window, the proportion of re-referrals within a 6-month window that led to a CPP or the child becoming looked after, and the share of looked after children that are in kinship care rather than other forms of care. Lower rates of the first two suggest more accurate assessments when the child was first assessed, while the latter measures changes in the level of family support through SofS.

Identification Strategy		Difference-in-difference
Unit of analysis		Individual (child / young person)
Primary outcome 1	variable	Duration of assessments
	measure (instrument, scale)	In working days
Primary outcome 2	variable(s)	Likelihood of initial child protection conference
	measure(s) (instrument, scale)	0-1
Secondary outcome 1	variable(s)	Likelihood of a case that received a NFA to be re-referred
	measure(s) (instrument, scale)	0-1 (within a 6 month window)
Secondary outcome 2	variable(s)	Likelihood of a case that has been re-referred to lead to a child protection plan/a child becoming looked after
	measure(s)	0-1 (re-referral within a 6 month window)

Secondary outcome 3	variable(s)	Likelihood of a looked after child to be taken into kinship care (as opposed to non-kinship care)
	measure(s)	0-1

Matching

We match intervention local authorities to control local authorities that most closely resemble them in the prior trends in outcome variables, and then analyse individual-level data from the intervention and control local authorities. We also match at an individual level using Coarsened Exact Matching (CEM). We outline the matching procedures in detail below.

Local Authority Matching

Exclusion Criteria

We select control local authorities from all English local authorities with children’s social care services excluding:

- Local authorities also running SofS. HSCWRU surveyed local authorities as to their use of SofS and we exclude local authorities which identified themselves as using “pure” SofS or “some form” of SofS as of Autumn 2017. MTM also provided a list of local authorities they consider to be currently “Active” in their use of SofS (as of 2019). Checks will be done with the selected control sites to ensure they are not using SofS.
- Local authorities which had applied to run SofS with MTM but which were rejected for substantive reasons beyond rationing to avoid policy endogeneity. We understand that the Innovation Programme accepted all applications to be part of the pilot.
- Local authorities which are likely to experience “contamination” from the treatment authorities (those which share geographic borders, or, who are partners⁶ in the Partners in Practice programme⁷).
- We further exclude control local authorities for which we have less than two years of testing periods or less than three training periods because of the difficulties of testing whether the match is sufficiently close.

Matching on parallel trends

The identifying assumption in a DiD analysis is parallel trends in outcome variables prior to the intervention. We match on local authority level trends in primary outcomes (from the child in need and child protection statistics⁸) in the duration of assessments and child protection conference rates for the years prior to the intervention. We set out to find two matches for each primary outcome variable separately. We match on the majority of the available pre-treatment data, testing on the remaining minority of years’ trends before that.⁹ Given that matching on each primary outcome variable implies analysing and contacting up to 36 local authorities serving as comparators and there are only 39 possible comparators, we choose

⁶ From a survey of the pilot local authorities.

⁷ Innovation Programme. Partners in Practice. <https://innovationcsc.co.uk/partners-in-practice/>

⁸ HM Government. *Statistics: children in need and child protection*. <https://www.gov.uk/government/collections/statistics-children-in-need>

⁹ We apply a 70% training, 30% testing rule-of-thumb. When applying the 70% rule to the number of years before SofS, we round to the nearest year when necessary.

the comparator local authorities for the secondary outcomes based on the two best matches of primary outcomes for each pilot local authority. If the primary outcome yielded only one suitable comparator, we draw on the ten best matches of the primary outcome to see whether a parallel trend in secondary outcomes exists. If none of the matches demonstrate sufficient evidence for a parallel trend in secondary outcomes (see below for full exclusion criteria), we exclude the pilot local authority from the analysis of our secondary outcomes.

Specifically, we match on the shortest Euclidean distance between the data for every pilot local authority and its potential comparators. The data is the change in the outcome variable between one year and the next (the outcome variable is standardised to take into account changes in the way that it is measured over the years) for all years in the matching dataset. For each intervention local authority, the two lowest scoring pairs whose trends are also convincing when inspected visually will be first preference for matching.¹⁰ Matching will be done with replacement, such that a single comparator could be used as a match for multiple intervention local authorities.

Two local authorities identified as the closest matches for each outcome for the intervention local authority will then be approached to be the control local authorities. We will request their pseudo-anonymised individual-level administrative data of the outcomes from the year before the pilot local authority they are matched to started Signs of Safety up until the current year's data.

If matches are not identified and the suspected reason for this is a small number of potential neighbours, we relax the exclusion criteria on geographic neighbours which we are using as a proxy for potential contamination, ask pilot local authorities which geographic neighbours they are working with closely and exclude just those geographic neighbours whom they are working with closely. We also relax the constraint around excluding local authorities which are currently "Active" as of 2019 but did not report using SofS in the Autumn 2017 survey as matches for the local authorities who were embedded pre-2016 as there is sufficient data for post-treatment comparison if the potential match local authority started post-2018.

If matches are not identified from a relatively large number of potential neighbours (30+), we accept that we cannot conduct the analysis for that pilot local authority.

Data description and assumptions

This section details the main assumptions we have made due to data inconsistencies over time. More details can be found in the appendix.

The way in which assessment duration has been reported has changed over the time period of interest. The outcome "median duration of (continuous) assessments" was collected from 31/03/2013 - 31/03/2014 to the current day. We use this metric as a matching criteria for the local authorities which start SofS after 31st March 2015 (Bexley). We use two other proxies depending on the available data:

¹⁰ If the visual analysis of the two lowest scoring pairs does not yield sufficient evidence for parallel trends, we visually inspect the ten lowest scoring pairs and choose the best (visual) matches from this pool of comparator LAs.

- For local authorities which start after 31st March 2013 but before 31st March 2015 (Brent, Bristol, Norfolk, Suffolk, Wokingham), we use the “median duration of (core) assessments” as a proxy for continuous assessments. There are three pre-treatment years of data with this proxy because the earliest publically available data is 31/03/2011 - 31/03/2012.
- For local authorities which start before 31st March 2013 (Leicestershire, Lincolnshire and West Sussex)¹¹, we use the percentage of core assessments completed within 35 working days as the median duration of assessments is not available. This data is publicly available as early as 31/03/2005 - 31/03/2006.
- In the transition year from core assessments to continuous assessments (31/03/2013 - 31/03/2014), local authorities implemented the transition at different times during the year. As a consequence, local authorities may have reported continuous and/or core assessments for this year. If data is available for continuous assessments, we use this data. If it is not available, we use core assessment data as a proxy. This might lead to decreased precision in our 2013 data since authorities might use different types of assessments. Only a small number of pilot LAs is affected by this since the period we consider for the matching ends before the 2013/2014 year for most pilot local authorities.¹²

The earliest publicly available data on the rate of ICPCs is 31/03/2011 - 31/03/2012. However, there is publically available data on the *number* of ICPCs as early as 31/03/2005 - 31/03/2006 which is early enough for our purposes. For years prior to 31/03/2011 - 31/03/2012, we calculate the rate by dividing by local authority mid-year population¹³ for children aged under 18 years¹⁴ and multiplying by 10,000.

The secondary outcomes are not collected and reported publicly, so that we have to rely on proxies for the purposes of finding suitable comparator LAs based on parallel trends before the intervention. We use referrals within 12 months of a previous referral as a proxy for secondary outcomes one and two. Data on this outcome is publicly available since the year ending 31/03/2006, which is early enough for our purposes. We inspect trends for both the number of re-referrals and the share of re-referrals as a percentage of all referrals in the respective year.

For secondary outcome three, namely the share of children that are cared for through kinship rather than non-kinship care, we use the rate of Special Guardianship Orders as a proxy for matching. We inspect trends for the rate of Special Guardianship Orders per 10,000 children and Special Guardianship Orders as a share of children who ceased to be looked after that year. Data on this outcome is publicly available since the year ending March 31, 2008. However, there are several local authorities for which data is not reported in select years to protect confidentiality due to the low number of Special Guardianship Orders in the respective years.

¹¹ and for the four local authorities mentioned above for the time before March 2011,

¹² The local authorities affected are Wokingham, Bristol and Bexley.

¹³ Office for National Statistics. *Estimates of the population for the UK, England and Wales, Scotland and Northern Ireland: Mid-2001 to mid-2018 detailed time-series.*

<https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland>

¹⁴ For years where both the rate and number are available, we compare the published rate and the rate calculated to check the validity of the methodology.

Matching robustness checks

The robustness of the match will be tested using a Mann-Whitney U test. We compare the Euclidean distances of the 10 best matches identified using the majority of the pre-treatment data to the Euclidean distances of the same 10 matches computed using a holdout dataset of pre-treatment data. A rejection of the null hypothesis indicates that the rankings of the comparators created using the training data is different from the ranking created using the holdout data which would suggest that the matches are not robust. We exclude LAs where the matches are statistically different ($p \leq 0.05$). We will also visually inspect the trends by plotting the local authority level outcome variables over time for the pilot LAs and the possible comparator LAs.

Individual-level Matching

In addition to the matching at local authority level, we also match at an individual level using coarsened exact matching (CEM)¹⁵. We do so in order to decrease the imbalance on covariates between the treated and control individuals, allowing the identification of a better causal estimate.

We chose CEM to match at an individual level because it allows analysts to specify ex ante the maximum acceptable imbalance. It also has a number of other desirable properties, for example, it removes the need for an additional process to restrict data to an area of common support, meets the congruence principle, is robust to measurement error, works well with multiple imputation (the way of handling missing data we have specified), and is computationally fast (important given that the dataset will be large relative to computing power available). CEM works by first temporarily coarsening the control variables so that the continuous variables are cut into categories (e.g. age as an integer coarsened to 0-5, 5-10, 10-15, 15-18 years) and categorical variables are combined (e.g. school year coarsened to primary school, senior school). All individuals are then assigned to strata with the same coarsened control variables. Strata which do not have at least one treatment and control individual are dropped.

We match on pre-treatment (i.e. one time period) individual-level control variables. For the purpose of CEM, we specify the coarseness of the variables as:

- Gender (included as a binary indicators: 0=male, 1= female, 2=other/unspecified)
- Age of children at the time of referral (0-4, 5-12, 13+ years)
- Ethnicity (major group)
- Academic year (primary school years, secondary school years)
- Disabled¹⁶ (included as a binary indicator: 0=No, 1= Yes)
- Free school meals (included as a binary indicator: 0=No, 1= Yes), Pupil Premium eligibility (for Reception, Year 1 and Year 2) as all infant school children in government-funded schools are FSM eligible¹⁷

¹⁵ Iacus, S., M., King, G. & Porro, G. (2018, April 12). CEM: Software for Coarsened Exact Matching. CRAN. <https://cran.r-project.org/web/packages/cem/vignettes/cem.pdf>

¹⁶ Hughes K, Bellis MA, Jones L, Wood S, Bates G, Eckley L, McCoy E, Mikton C, Shakespeare T, Officer A. Prevalence and risk of violence against adults with disabilities: a systematic review and meta-analysis of observational studies. *Lancet* 2012.

¹⁷ HM Government. Apply for Free School Meals, <https://www.gov.uk/apply-free-school-meals>

- Is child an Unaccompanied Asylum Seeker¹⁸ (included as a binary indicators, 0=No, 1= Yes)
- Number of prior times subject to a CIN or CP plan (0, > 0)

We make sure that individuals from each pilot local authority are only matched to individuals from their comparator local authorities (that have been identified as having parallel trends) in the CEM procedure. Note that the coarseness is only for matching purposes and we describe our operationalisation of covariates for inclusion in the regression below. We report the proportion matched and the multivariate imbalance score which measures imbalance with respect to the joint distribution, including all interactions, of the covariates (Iacus, King and Porro, 2011)¹⁹. We then estimate the difference-in-difference regression weighted by the weights that equalise the number of treated and control individuals within each CEM stratum.

Participants

The piloting of SofS was part of the Department for Education (DfE) Innovation Programme. Being part of the pilot involved local authorities submitting an application demonstrating that the project could be sustainable long-term, that they would secure all the buy-in necessary to make the programme a success and that they would collaborate with policymakers and national partners as well as providing data on their success. It is our understanding that the Innovation Programme accepted all applications to be part of the SofS evaluation.

Local authorities who applied to the programme had a range of levels of experience with SofS before starting. They applied to be part of the programme for a variety of reasons from empowering families, addressing low morale across the social work workforce, achieving greater consistency in social work practice, simplifying the system, engaging more effectively with multi-agency partners to using it as a platform for broader change.

The Innovation Programme (IP) Wave 1 lasted for eighteen months from 2014 to 2016. The IP Wave 2 lasted two years from September 2017 until September 2019. It is important to note that London Borough of Tower Hamlets dropped out after the IP Wave 1 finished, and Bexley joined in the IP Wave 2. Wakefield was part of the IP Wave 2 but dropped out in January 2019. We exclude local authorities which dropped out from the analysis (to follow the approach of HSCWRU).

Sample size / MDES calculations

We have not included power calculations for this analysis because there are too many assumptions that we would need to make for the power calculations to be useful. Furthermore, because of the retrospective nature of the DiD analysis, sample size is constrained more by the practicalities of data access than for other projects.

¹⁸ UN High Commissioner for Refugees (UNHCR), *UNHCR's Engagement with Displaced Youth*, March 2013, available at: <https://www.refworld.org/docid/5142d52d2.html> [accessed 14 June 2019] p28

¹⁹ Iacus, Stefano M., Gary King and Giuseppe Porro. 2011. "Multivariate Matching Methods that are Monotonic Imbalance Bounding." *Journal of the American Statistical Association* 106:345–361. <http://gking.harvard.edu/files/abs/cem-math-abs.shtml>.

Outcome measures

To answer research question 1, the outcome measure is: duration of assessment (days). We only consider assessments that have started and finished outside of the embedding period (we anticipate SofS will need a period of time after implementation to be fully functional and to allow new practices to be fully embedded - we thus exclude this embedding period from our analysis).²⁰

To answer research question 2, the outcome measure is: the likelihood of an initial child protection conference (ICPC) (0-1). We only consider the likelihood of an ICPC for children that have already been assessed and whose case is designated as open.

To answer research question 3, the outcome measure is: the rate of cases that received an NFA to be re-referred within a 6-month window (0-1).

To answer research question 4, the outcome measure is: the rate of cases that are re-referred within a 6-month window and lead to a CP plan or the child becoming looked after (0-1).

To answer research question 5, the outcome measure is: the share of looked after children and children who ceased to be looked after that are cared for through kinship care²¹ (0-1).

Analysis plan

Primary Analysis:

The quantitative work being undertaken by HSCWRU as part of their evaluation already involves identifying a group of statistical nearest neighbour (SNN) sites that do not use SofS. The SNNs are identified within the LAIT and are identified on the basis of similarity (the inverse of the Euclidean distance) between background characteristics²². This SNN group is then being compared with the ten pilot sites across a range of nationally available outcome measures to test for significant differences. As our approach to identifying causal impacts is a DiD approach, we match local authorities based on the extent to which they meet the assumptions for that identification strategy. In particular, this will involve matching on the *trends* of the outcome measures over the last few years, rather than the *levels* of those outcome measures. It will not take into consideration other background characteristics used to determine HSCWRU's SNNs.

Prior to analysis, each local authority will be matched with two candidate control local authorities based on trends in our outcomes of interest. This matching having been completed, we will estimate the effect of the intervention, β_3 on the outcomes of interest, Y_{ist}

$$Y_{ist} = \alpha + \beta_1 SS_s + \beta_2 T_t + \beta_3 (SS_s \cdot T_t) + \beta_4 X_{it} + \beta_5 \Gamma_{st} + \theta Year_t + \epsilon_s(1)$$

²⁰ If a child had several assessments during the pre- or post intervention period we take the average duration of assessment.

²¹ We consider children under reg 24 or under a Special Guardianship order to be in kinship care.

²² See Appendix A of Department of Education (2017, September). Local Authority Interactive Tool: User Guide.

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/643132/LAIT_User_Guide_2017.pdf

Where;

- α is a regression constant
- SS_s is a binary indicator of whether the local authority received SofS
- T_t is a binary indicator of whether the local authority is receiving SofS at time t
- X_{it} is a vector of participant level characteristics (see “Individual level covariates”)
- Γ_{st} is a vector of local authority level characteristics (see “Local authority level covariates”)
- $Year_t$ are year dummy variables to capture time trends common to all authorities
- ϵ_s is an error term, denoting standard errors clustered at the level of the local authority (the level at which assignment takes place)

We include data for $t - 1$ and from $t + e$ until the current year where e is the time taken to reach a steady state to allow the evaluation of the programme to start when teams understand and have become accustomed to new ways of working (the embedding period). The time taken to reach a steady state comes from triangulation of qualitative work conducted by Mary Baginsky and a survey of the pilot LAs at a SofS leadership day. For some local authorities this is a gap of up to 36 months from when SofS started until it reached a steady state.²³ We investigate whether the estimated timing of the steady state affects the treatment effect in the sensitivity analysis.

DiD involves multiple observations of the control and treatment groups over time. This can create serial correlation where errors in one period transfer to another. Serial correlation can be a particularly important issue in the DiD context because of the inclusion of fairly long time series, dependent variables which are commonly used are typically highly positively serially correlated, and the treatment variable changes little within a local authority over time. Serial correlation underestimates the standard errors, meaning that we have too much confidence in the result, and we get a higher rate of false positives. We check for serial correlation using the Durbin-Watson test. We consider there to be serial correlation if the test statistic is outside of the range of 1.5 - 2.5. Serial correlation is partially taken care of when clustering at the local authority level.²⁴ To counteract individual-level serial correlation, we also try a multilevel model with standard errors clustered at both the individual and the local authority level and consider whether the multilevel model produces a statistically different and practically different estimate of the treatment effect to the treatment effect estimated from equation (2) (where we take practically different to be 20% above or below).

Secondary Analysis:

Additional subgroup analyses

We further investigate whether the effect of SofS is more pronounced within certain subgroups (subject to availability of the data). We investigate whether local authorities which self-report to have SofS embedded well have different treatment effects. We categorise local authorities based on a composite of their self-reported profiling scores. The score came from the “Profiling survey” conducted by HSCWRU²⁵ as of the beginning of Wave 2: 0-4, 5-6, 7-8, 9-10 where 0 is “not at all embedded” and 10 is “fully embedded”.

²³ We will re-confirm exact embedding periods with the local authorities when commencing the data collection process.

²⁴ Clustering at a group level (not group-year level) seems to account well for serial correlation when there are a sufficient number of groups (p.17). Bertrand, M., Duflo, E. & Mullainathan, S. 2004. "How Much Should We Trust Differences-in-Differences Estimates?," *The Quarterly Journal of Economics*, MIT Press, vol. 119(1), pages 249-275, February <https://www.nber.org/papers/w8841.pdf>

²⁵ A tool designed in Innovation Programme Wave 1 with experienced Signs of Safety practitioners to assess the strategic leads' perceptions of their local authorities' fidelity to the practice model. Local authorities were asked their perception of their fidelity at the end of Wave 1 fidelity at the end of Wave 1 and retrospectively for fidelity at the beginning of Wave 1, and also at Wave 2 to estimate the fidelity

We also categorise local authorities based on their Ofsted ratings and MTM's scoring of the quality of delivery and conduct subgroup analyses for each of these categories. We run separate regressions for each of these categories and plot the coefficient on the DiD variable against the profiling score to visually inspect whether the subgroups have different treatment effects.

Channels

We test whether SofS influences the duration of assessments by increasing the number of same day assessments by including a dummy for same day assessment in the regression and reporting whether it is significant.

Practice models

We recognise that there may be overlap between the SofS model and two other popular practice models, Restorative Justice and Reclaiming Social Work. If the mechanisms through which the three practice models are expected to affect outcomes are the same, comparing the pilot SofS local authorities to local authorities using Restorative Justice or Reclaiming Social Work would bias downwards any potential effect. We ask the comparator local authorities whether they use either practice model. In this analysis, the SS_s in equation (1) above becomes a three-level indicator of whether the local authority received SofS (the base category), a similar practice model (Restorative Justice and Reclaiming Social Work), or neither.

Covariates

In order to increase the precision of our estimates, we include the following individual level and local authority covariates (where they are available). We do not include individual-level fixed effects because the same child is unlikely to, for example, both be assessed and have an initial child protection conference multiple times in the time period of interest.

Individual level covariates

- Gender (included as a discrete indicators: 0=male, 1= female, 2=other/undefined/unborn)
- Age of children at the time of referral
- Academic year
- Disabled²⁶ (included as a binary indicator: 0=No, 1= Yes)
- Free school meals (included as a binary indicator: 0=No, 1=Yes), Pupil Premium eligibility (for Reception, Year 1 and Year 2)
- Is child Unaccompanied Asylum Seeker²⁷ (included as a binary indicators, 0=No, 1= Yes)

at the start of Wave 2. The profiling tool involved 50 items under 6 headings and local authorities could rate themselves on a scale from 1-10 with higher numbers indicating higher fidelity. The embeddedness measure is an average of indicators identified by HSCWRU in conjunction with some pilot local authorities which best represented whether SofS was embedded. These are: "Embedding an organisational commitment to Signs of Safety", "Mapping cases by individual social workers", "Using safety plans across initial and review child protection conferences and in all related groups", "Providing advanced 5-day training for all social workers", and "Establishing practice leadership and supervision processes to support Signs of Safety".

²⁶ Hughes K, Bellis MA, Jones L, Wood S, Bates G, Eckley L, McCoy E, Mikton C, Shakespeare T, Officer A. Prevalence and risk of violence against adults with disabilities: a systematic review and meta-analysis of observational studies. *Lancet* 2012.

²⁷ UN High Commissioner for Refugees (UNHCR). (2013, March). *UNHCR's Engagement with Displaced Youth*. <https://www.refworld.org/docid/5142d52d2.html>, p28.

- The date when the referral was made (included as binary indicators: 0=Usual, 1= the day after a bank holiday or the 5 days before school holidays)²⁸
- How many previous plans a child has had (integer)
- The main need which the episode is taking care of. (included as a categorical variable: 0 = Not stated, 1 = Abuse or neglect, 2 = Child's disability/illness, 3 = Parental Disability/illness, 4 = Family in acute stress, 5 = Family dysfunction, 6 = Socially unacceptable, 7 = Low income, 8 = Absent parenting, 9 = Cases other than Children in Need) for children in need
- Latest category of abuse. (included as a categorical variable: 1= Neglect, 2= Physical abuse, 3= Sexual abuse, 4= Emotional abuse, 0=Multiple/ Not recommended)

Local authority level covariates

- Local authority fixed effect
- Numbers of assessments to Children's Social Care
- Average number of cases per children and family social worker (based on FTE counts) as defined in the LAIT (integer)
- Most recent Ofsted rating (included as a categorical variable on a 4-point scale: 'outstanding'=1, 'good' =2, 'requires improvement' = 3, 'inadequate' = 4)
- Proportion of children / young people not seen in accordance with the timescales specified in the plan (continuous variable)
- Proportion of children / young people Income Deprivation Affecting Children Indices. IDACI Score or Free School Meals eligibility (continuous variable)
- Proportion of children / young people white British (continuous variable)
- Population density (for the duration of assessments variable as this may affect how quickly assessments can be done)
- Number of other Innovation Programme Projects the LA is participating in

The data will be sourced from a variety of data sources (Characteristics of Children in Need Tables, LAIT, Ofsted reports, aggregate measures of individual-level data requested from LAs). Variables which do not change over time e.g. urban / rural are excluded because the fixed effects absorb the effects of time constant variables.

Handling missing data

In cases of missing data, we will consider the possible reasons for its missingness and undertake statistical analyses to determine whether there are any patterns relating to other recorded covariates or to the intervention variable. We will drop observations with missing outcome variables, and will consider dropping a variable which has > 30% of the data missing. We conduct multiple imputation where data is missing experimentally at random.

²⁸ Local authorities often experience a spike in referrals in these two time periods. Referrals from schools tend to increase before holidays since teachers are worried about the child staying at home all day over an extended period of time. During bank holidays, children are spending more time at home and need additional attention/supervision. Practice experience suggests that bank holiday weekends also correlate with higher alcohol consumption and violence.

Sensitivity

Placebo test using alternative control group

There are a few placebo tests to attempt to falsify DiD results. Our preferred method would have been to test for a DiD in previous periods (when no intervention took place and therefore there is no expectation of any treatment effect). However, we will not have access to multiple previous period data because we expect that requesting such data will not be seen as proportionate by local authorities. One can also test outcomes that are not expected to be affected by the treatment. However, we do not feel we understand the mechanisms of SofS fully enough to choose an alternative outcome for testing. To attempt to falsify the DiD result, we recode one comparator local authority in each match as a pilot local authority at random and re-estimate the DiD comparing this recoded LA with the other LAs in the match group. We only include match groups where there is more than one LA.

Time varying treatment effect

For the primary analysis, we have been generous in the length of time allowed for SofS to reach a “steady state” (not necessarily full implementation). We think it probable that evaluating the intervention too early is more likely to risk missing the treatment effect than evaluating the intervention too late given that a whole system change likely takes some time to reach a steady state. Furthermore, the intervention is ongoing and so we would not expect the effect to tail off after its start (except perhaps due to reduced enthusiasm after the initial novelty - we wish to capture SofS in its “business as usual” phase so do not see the risk of missing the effect of any initial enthusiasm as problematic).

Given that the treatment effect may be sensitive to the timing chosen, we treat t as the first time period treated (instead of $t + e$) and introduce lagged treatment variables from $t + 1$ to $t + M$ ²⁹.

$$Y_{ist} = \alpha + \beta_1 SS_s + \beta_2 T_t + \beta_3 (SS_s \cdot T_t) + \beta_4 X_{it} + \beta_5 \Gamma_{st} + \sum_{m=0}^M (SS_s \cdot T_{t+m}) \delta_m + \epsilon_s(2)$$

Where;

- α is a regression constant
- SS_s is a binary indicator of whether the local authority received SofS
- T_t binary indicator of whether the local authority is receiving SofS at time t
- T_{t+m} is a binary indicator that equals one if the observation is from a local authority receiving SofS at the time $(t+m)$ the observation is made, and 0 otherwise.
- X_i is a vector of participant level characteristics (see “Individual covariates”)
- Γ_s is a vector of local authority level characteristics (see “Local authority level covariates”)
- ϵ_s is an error term, denoting standard errors clustered at the level of the local authority (the level at which assignment takes place)

δ_m with $m \geq 0$ measures any lagged effects of the treatment over time. The time period in which $\delta_m \geq 0$ gives an indication of when the treatment is starting to take effect.

²⁹ M is the number of years we evaluate after the implementation of SofS. Angrist, J. D., & Pischke, J.-S. (2009). Mostly harmless econometrics: an empiricist's companion. Princeton, Princeton University Press. p. 237

Ethics & Participation

We maximise the benefit of the evaluation by providing an additional lens to understand the impact, in particular getting closer to a causal estimate of the impact of SofS, which is informative for local authority decision-making as to whether or not to invest in SofS. We believe the risk of harm is very low. The data used is administrative data which is collected / created in the course of day to day children's social work, and no further collection of data is required. The analysis does not involve innovative technology, denial of service, large-scale profiling, biometric data, genetic data, data matching, invisible processing, tracking or targeting of individuals for marketing purposes. The outputs will be presented as summary statistics and will be checked for statistical disclosure.

The low risk of harm mostly comes from the possibility of harm if the individual were identified (very unlikely) following a data breach (also very unlikely). We mitigate the risk of a data breach by using the ONS' secure research service (SRS). Data will be stored on the ONS' systems. Access to the data will be limited to the research team at WWCS; all researchers have undergone rigorous data protection training. It is very unlikely that the data requested will enable re-identification because we only ask for the data necessary to undertake the analysis and this contains no "instant identifiers" (e.g. name, address etc) or "meaningful identifiers" (which would allow matching to other datasets with more information).

This extension of the SofS evaluation Wave 2 has been approved by King's College London's GSSHM Research Ethics Panel (REP).

Registration

The trial will be pre-registered on OSF (Open Science Framework, <https://osf.io/>) run by the Centre for Open Science (<https://cos.io/>).

Data protection

The underlying data used to conduct this analysis consists of administrative data from local authorities involved in the Innovation Programme Wave 2 SofS project and comparator local authorities not conducting SofS. The data about individuals requested via the ONS secure research service will be pseudo-anonymised. We will not request any 'instant identifiers' (that would allow us to point to an individual in the dataset) or 'meaningful identifiers' (which would allow identifying someone through linking the data to another dataset). We will require 'meaningless identifiers' ('data variables used within a dataset but have no meaning beyond the dataset's boundaries') to track individuals over time.

This section is structured according to the guidance given by the Information Commissioner's Office, which "covers the General Data Protection Regulation (GDPR) as it applies in the UK, tailored by the Data Protection Act 2018"³⁰.

Principles of the GDPR

Principle (a): Lawfulness, fairness and transparency

³⁰ Information Commissioner's Office, Guide to the General Data Protection Regulation (GDPR). <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/>

1. Lawfulness:

WWCSC is the data controller. WWCSC decided to process the data and decided the purpose of the processing, what data should be collected and which individuals to collect data about. The data is collected by the DfE. The legal basis for What Works for Children's Social Care processing the data is legitimate interest.

Legitimate interest is a three part test:

1) Purpose test: are you pursuing a legitimate interest?

We are a research group within an innovation charity, whose purpose is to improve the evidence base in children's social care. We consider the processing of the data to be in our legitimate interests because it will enable us to produce research in this area, which will benefit local authorities, in particular senior leaders who make decisions about practice models.

2) Necessity test: is the processing necessary for that purpose?

The processing is necessary for the purpose because processing individual-level data allows us to conduct analysis which is better powered to detect the impact of SofS, and which allows us to better control for the circumstances of the individual which may affect the outcome. Both of these factors mean that we are more likely to be able to provide meaningful research which can be used to inform practice, with downstream effects for children involved in statutory social care.

3) Balancing test: do the individual's interests override the legitimate interest

We have published a privacy notice on our website to give general notice of this processing. While the data is quite sensitive and on a population which includes vulnerable children, we will not be using identifiable IDs and the data will be stored securely. We believe this processing falls within generally socially acceptable uses of this kind of data - it is scientific research in the public interest by a charity and for the benefit of a vulnerable group. Alongside the privacy notice, we include a form which individuals can fill in to uphold their individual data rights.

We therefore believe that the individuals' interests do not override our legitimate interest in this processing.

The legal basis for processing special category data is that it is necessary for archiving, scientific, historical research or statistical purposes (point (e) of section 10 of the DPA which refers to (j) (archiving, research and statistics) of Article 9(2) of the GDPR). The project meets condition (4) in Part 1 of Schedule 1:

(a) is necessary for archiving purposes, scientific or historical research purposes or statistical purposes,

This processing constitutes scientific research as it will be used to create evidence on pre-defined, specific hypotheses around what works to improve outcomes for children who have undergone a statutory intervention, in order to increase the knowledge base in this area. The special category data we are using is data concerning gender, and health, specifically disability. Not being able to assign gender or disability status to our data limits the scientific value of this research because they are likely moderators of social care outcomes. The likelihood of children to enter care also varies significantly by ethnic group and is thus

important to control for when trying to gauge the impact of Signs of Safety on children's services.

(b) is carried out in accordance with Article 89(1) of the GDPR (as supplemented by section 19)

Organisational and Technical Arrangements

"Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation. Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner."

The data will be pseudo-anonymised i.e. it can no longer be attributed to a specific data subject without the use of additional information. We are not requesting any 'instant identifiers' (e.g. name or address) or 'meaningful identifiers' (identifiers that allow linking to other datasets).

See "Organisational and technical arrangements".

Safeguards (DPA 2018 Section 19)

In the UK, the requirements of Article 89(1) GDPR will not be met unless the provisions of Section 19 DPA 2018 are also complied with. We have no reason to believe that the research will cause damage or distress (and certainly not substantial damage or distress) to the children or young people - the analysis requires no extra involvement of the children or young people. The data has already been collected in the course of day-to-day work with the child/young person and their family. The processing and presentation of evidence is unlikely to have distressing effects because we protect against identification of the individual and also against statistical disclosure (following the ONS standard rules outlined in the Approved Researcher training). The research is not being carried out for the purposes of measures or decisions with respect to a particular data subject but looks at the effect of SofS on the cohort as a whole.

(c) is in the public interest.

The work is intended to support work towards high standards of quality of social work practice which affects a substantial section of the public.

2. Fairness:

ICO's guidance says fairness means "you should only handle personal data in ways that people would reasonably expect and not use it in ways that have unjustified adverse effects on them"³¹. This data is being used for statistical research to understand whether a practice model is working and contribute towards improvements in public services. We believe that "the reason person" would find the use of data in this way acceptable.

3. Transparency:

This will be covered below in the section on the right to be informed. We will ensure that privacy notices are written in clear and plain language. We will also ensure that notices have a Flesch-Kincaid grade level of 7 to ensure that either older children who are able to object by themselves can do so and that the notices are accessible to all parents.

³¹ Information Commissioner's Office. Principle (a): Lawfulness, fairness and transparency. <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/lawfulness-fairness-and-transparency/>

Principle b): Purpose Limitation

This data will only be used to answer the research questions in this document, as part of a general purpose to increase the evidence base about how SofS affects the outcomes of children / young people and their families involved in social care. They will not be used for any other purpose, other than usual statistical checks to ensure the accuracy of the data.

Principle c): Data Minimisation

We have only requested data that is adequate, relevant and limited to what is necessary to fulfil the purpose of this project i.e. to build the evidence base on practice models. Broadly speaking, we can classify the data requested into two groups, broadly individual-level and local authority level variables. The individual-level variables are sourced from local authority administrative datasets, and local authority level variables are sourced from public data e.g. the LAIT and a survey of local authority representatives knowledgeable about the local authority's implementation of SofS.

Individual-level variables

- Outcome measures which are necessary to measure whether SofS was successful;
- Other individual-level variables which we expect to influence the outcomes. Not being able to include these variables limits the scientific value of this research because they are likely moderators of social care outcomes.

Local authority level variables

- Local authority level variables which we expect to influence the outcomes.

Principle d): Accuracy

The local authorities spend considerable time cleaning the administrative data so that it is suitable for data returns to the Department, and we are requesting only data that is in such returns (for example, the LAIT³², CIN Census³³). To validate data quality, we will conduct checks on the following: data-type constraints (words instead of numbers where we expect them), range constraints for numeric data; set-membership constraints for categorical data (are the categories limited to what we expect?); regular expression / formatting patterns (e.g. dates) and cross-field validation (e.g. before dates start before after dates). Please see the "Handling missing data" for our approach to missing data in the administrative datasets.

Principle e): Storage limitation

Data will be deleted 6 months after the analysis is published.

Principle f): Integrity and confidentiality (Security)

See "Data security arrangements".

Principle g): Accountability principle

³² HM Government. Local authority interactive tool (LAIT), <https://www.gov.uk/government/publications/local-authority-interactive-tool-lait>

³³ HM Government. *Statistics: children in need and child protection*. <https://www.gov.uk/government/collections/statistics-children-in-need>

The Executive Director of the What Works Centre and Principal Investigator for this research (Dr. Michael Sanders) will be ultimately responsible for the conduct of the research. Other details are below in the accountability and governance section.

Individuals' rights under the GDPR

The right to be informed

The Centre has published a privacy notice on its website detailing how the processing will be done. As this data is indirectly collected and for “scientific or historical research purposes” as well as “statistical purposes” the Centre is relying on an exemption to the requirement to individually inform participants as it would “*prevent or seriously impair* the achievement of the purposes for processing”.

This is the case because:

- It would require the Centre, which is part of a not-for-profit organisation to expend considerable resources to mail a large number of individuals thus leaving less resources to undertake the processing;
- It would require re-identifying the individuals via their addresses, which is data the Centre does not have access to.

The right to access, rectification, erasure, restriction of processing and to object

Individuals have the right to access their individual data and supplementary information. The right of access allows individuals to be aware of and verify the lawfulness of the processing. Individuals are entitled to obtain:

- confirmation that their data is being processed;
- access to their individual data; and
- other supplementary information

If an individual wishes to access this information, we cannot comply directly because we do not have identifiers in the dataset. We would point the individual towards the trial protocols to indicate the type of information that we hold on them for the purpose of this analysis. We would then collect the information necessary for the DfE to identify them via the online form, and refer the case to the DfE where the request can be handled using the DfE’s own subject access request procedures. For individuals invoking their rights to rectification, erasure, restriction of processing and to object, we would then require the DfE to inform us of which rows of data to rectify or delete.

The right to data portability

The right to data portability allows individuals to obtain and reuse their individual data for their own purposes across different services. It allows them to move, copy or transfer individual data easily from one IT environment to another in a safe and secure way, without hindrance to usability. This is not particularly relevant in the context of statistical analysis as the value of processing the data is to the public and comes from the aggregation of the data, rather than from the processing of the individual’s data, and so it is difficult to imagine the purpose of porting the data to an alternative system.

Individual’s rights in relation to automated decision-making and profiling

Nothing in this analysis is related to either automated decision-making or profiling of any individuals.

Accountability and Governance

Nesta (which is incubating WWCS) has a Data Protection Team which has the responsibility for the management of Data Protection on behalf of the Organisation. The Data Protection Team includes the Deputy CEO, ensuring compliance with GDPR at the highest level of management. Each team has a Data Protection Lead which has day-to-day responsibility for data protection. There is no formal Data Protection Officer at Nesta because Nesta does not routinely undertake large-scale processing of personal data.

The Centre takes and documents the appropriate technical and organisational measures in place to comply with GDPR.

The approach of Nesta (which is incubating WWCS) to information security is outlined in its Information Security and IT Usage Policy.

Checks on staff

The data will only be accessed by Centre research team members. Research staff at WWCS have undergone data protection training and have substantial experience in handling data. The research team continues to review the training needs of the team to ensure the Centre's approach remains up-to-date.

Data security arrangements

Data will not be stored on WWCS systems but on the ONS systems. Access will only be granted to research team members named as protocol authors.

Personnel

Evaluation team

Michael Sanders, Executive Director, WWCS, is the principal researcher. Vicky Clayton, Data Science Manager, WWCS will lead on analysis and manage the project with Eva Schoenwald and Alix Leroy, Researchers, WWCS.

For the team from HSCWRU, please see:

<https://www.kcl.ac.uk/scwru/res/knowledge/signsofsafety2>

Appendix

Estimations of when each pilot local authority started SofS started and when it reached a steady state is based on triangulation between Section 1.2 of the *Evaluation of Signs of Safety in 10 pilots*³⁴ from Wave 1, conversations with Mary Baginsky incorporating her qualitative work and a survey of each pilot local authority at the Signs of Safety leadership day in June 2019. Where there has been uncertainty, we have allowed for a longer time period for SofS to reach a steady state. This is because we anticipate that we are least at risk of missing a treatment effect by allowing a longer time period for embedding.

Pilot local authority	Innovation Programme evaluation history (from Wave 1 report)	Start of implementation (according to survey)	Number of months to reach steady state (according to survey and qualitative work by Mary Baginsky)	Date from which to evaluate
Bexley	Joined the pilot after Wave 1	July 2015	18	January 2017
Brent	2 year' grouping in Wave 1 - Up to two years' experience	June 2013	36	June 2016
Bristol	New grouping in Wave 1- beginners with either no previous experience or up to one year's experience	April 2014	36	April 2017
Leicestershire	'2+ years' grouping - more than 2 years' experience	September 2012	36	September 2015

³⁴ Baginsky, M., Moriarty, J., Manthorpe, J., Beecham, J. & Hickman, B. (2017, July). Evaluation of Signs of Safety in 10 pilots: Research report. *Social Care Workforce Research Unit, King's College London / Department for Education*.
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/625376/Evaluation_of_Signs_of_Safety_in_10_pilots.pdf

Lincolnshire	'2 year' grouping in Wave 1 - Up to two years' experience	2012	36	October 2015
Norfolk	New grouping in Wave 1- beginners with either no previous experience or up to one year's experience	October 2014	24	October 2016
Suffolk	'2 year' grouping in Wave 1 - Up to two years' experience	February 2014	24	February 2016
West Sussex	'2+ years' grouping - more than 2 years' experience	2012	36	2015
Wokingham	New grouping in Wave 1 - beginners with either no previous experience or up to one year's experience	June 2014	36	June 2017

Additional details on matching procedure

This section details additional assumptions and procedures that were required for the local authority matching due to inconsistencies in the datasets.

Several local authorities have been united under one local authority in 2019. We only analyse councils in their pre-2019 divisions for our purposes. In addition, Bedfordshire and Cheshireshire were both split into several local authorities on 1st April 2009. We include a buffer of three years until we can include these local authorities in our analysis, recognising that the separation constitutes a large change which could explain changes in outcomes. Consequently, these local authorities will not be included in our assessments.

We do not have sufficient data to match on for several potential comparator local authorities. This is either because local authorities did not report the relevant figures in their annual reports or because the data was removed due to small sample size considerations.