

STAYING CLOSE FEASIBILITY STUDY REPORT

CONTENTS

Contents	2
Executive summary	4
Introduction	4
Theory-building.....	4
Design testing.....	5
Recommendations.....	5
Introduction	8
Methodology	10
Theory-building.....	10
Design testing.....	12
Economic analysis scoping.....	13
Synthesis and recommendations	14
PROGRAMME	15
Why: what is the purpose of the programme?	15
Who: young people.....	17
Who: professionals.....	18
What: business as usual (BAU)	19
What: programme	19
How	20
When	21
Where	22
How much, how often	23
Modifications and tailoring.....	24
How well.....	24
Evaluation Questions	25
1. Does it work?	25
2. Does it work as expected?	27
3. Does it work differently for some groups?	28
4. Does it work differently in some places?.....	28
5. Was it implemented as intended?	29
6. Is it a good use of resources?.....	30

7. What else do we want to learn?	31
Evaluation Methods	32
Counterfactuals and comparisons.....	32
Data sources and properties.....	35
Risks and strategies to mitigate	38
Overall Synthesis and Recommendations	40
Evaluation method and classification	40
Other design choices	41
Evaluation questions.....	43
References	53

EXECUTIVE SUMMARY

Introduction

Staying Close is a programme of support for young people leaving residential care and transitioning to independence, developed by the Department for Education (DfE). We conducted a feasibility study between May 2022 and March 2023 with 15 local authorities (LAs) who had been awarded funding to implement Staying Close. This aimed to understand what questions should be asked as part of the Staying Close evaluation, and how the intervention's impact might be robustly evaluated. We carried out a range of qualitative and quantitative work and considered findings from both strands of work to make design recommendations for the next stage of evaluation due to start in October 2023.

The study found that there are two outcomes which are most suitable for further evaluation: accommodation changes and education, employment and training (EET).

Theory-building

We carried out work to improve our theoretical understanding of Staying Close, and to develop a Theory of Change (ToC) for the programme. We reviewed evaluations of Staying Close pilot sites and the existing literature on support for care leavers as they transition to adulthood. We also reviewed bids from the 15 LAs that the programme was rolled out to in 2022, conducting interviews with staff. Important/expected outcomes for Staying Close included:

- An increase in the ability to maintain a tenancy in a suitable home
- An increase in the ability to form and draw on supportive relationships
- A reduction in involvement in the criminal justice system (as perpetrators)
- An increase in the ability to live independently
- An increase in the ability to appropriately care for one's own wellbeing.

Local variations in Staying Close implementation were observed, including differences in accommodation, relationships, wellbeing, independent living skills, and education, employment or training (EET) support, influenced by each LA's unique context.

We found that mechanisms through which programme activities (e.g. one-to-one support from a key worker) achieved outcomes varied depending on how an LA had chosen to implement Staying Close. Broad mechanisms that applied across contexts were care leavers' awareness of the programme, their experience of the programme and the take-up of the programme within an LA (all influencing the extent of engagement with programme activities). We also identified that stable

and suitable accommodation, having trusted relationships with staff and improved wellbeing were both mechanisms and short-term outcomes for young people.

Consistent barriers to implementation of the programme included limited housing availability, high staff turnover, unfilled specialist staff roles, poor engagement with care leavers in programme development and suboptimal matching of different young people in shared accommodation. Facilitators were suitable accommodation being available within the LA, staff training in trauma-informed care and co-production with care leavers.

Design testing

To understand how the impact of the intervention might be robustly evaluated, we first created a list of all possible designs and systematically assessed their suitability given what we knew about the programme, its recipients and practitioners and DfE's roll-out plan. We then assessed the most promising options according to a set of criteria. We concluded the most suitable methods were a randomised control trial (RCT) or a difference-in-differences design (DiD) if randomisation was not possible. An RCT involves randomly assigning LAs to implement Staying Close or not. For a DiD, young people receiving Staying Close are compared with those with similar characteristics from LAs that are expected to show the same outcomes as the treatment LAs if the programme were not implemented.

The data scoping work involved mapping all variables suitable for the evaluation and their possible sources. After shortlisting them, we collected administrative and survey data from LAs and explored the feasibility of linking individual records in national datasets.

After analysing the administrative data, we have no major concerns that it would be of sufficient quality and reliability for use as primary outcomes. We identified low staff capacity, internal miscommunications and high turnover rates within teams as the main barriers to data returns. We also found that the timelines for accessing national databases would not be compatible with reporting requirements for DfE. However, we will continue to scope the use of this data as a mitigation for risks of administrative data collection. Lastly, we concluded that relying on survey data would pose substantial risks to the evaluation, but these risks could be offset by including administrative measures as primary outcomes and including formative questions to identify effective data collection strategies.

Recommendations

We synthesised findings from these two strands of work to produce recommendations for a trial design that can most robustly evaluate the Staying Close programme.

Changes to DfE's policy objectives following increases to the programme funding (announced in March 2023) made a randomised design unfeasible. Following conclusions from our design framework, and risks associated with our preferred outcomes, we recommend that the evaluation

be considered a **pilot DiD evaluation**. This will aim both to estimate the impact of Staying Close on the outcomes of interest and to test various strategies for data collection and different measures of wellbeing and social connectedness.

We propose the pilot evaluation runs from October 2023 to October 2024. If LAs begin delivery after October 2023, we recommend allowing for 12 months of delivery up to December 2024. Participants will be included if they start receiving the intervention by April 2024, or six months before the endpoint for their LA's delivery period.

The DiD will involve identifying comparator LAs based on historical trends in the outcomes of interest (before the introduction of the programme). Differences in these trends in the two sets of LAs after the programme is delivered are then assumed to be the result of the intervention (the "parallel trends assumption"). We recommend accessing individual-level data from the comparator LAs and matching individuals with young people receiving the programme to estimate the effect of Staying Close on outcomes.

We recommend focusing on two outcomes as part of the DiD: accommodation changes and education, employment and training (EET). This recommendation is based on their centrality to the programme's theory, data access and characteristics, outcome maturation, stakeholder interest and suitability for cost analysis.

Note on subsequent evaluation

The feasibility study concluded that the evaluation of the programme should focus on NEET status and accommodation outcomes. This study found that there is not enough high-quality data on social connectedness to conduct a difference-in-difference analysis. Social connectedness data is not routinely collected; this means there is no historical data that can be used to create a matched control group. A future evaluation is, therefore, limited to measuring outcomes where there is good historical data, which largely comes from administrative datasets.

To facilitate the use of social connectedness as an outcome in future studies, we recommended further formative work to identify a measure for social connectedness and a process for collecting this data is undertaken. Findings from this formative work on social connectedness will be used to inform our strategic objective of establishing reliable indicators for the strength of care leavers' relationships.

In view of the conclusions about primary outcomes, we have agreed that the Centre for Homelessness Impact (CHI) are best placed to take forward the evaluation of Staying Close and they will commission this evaluation via their Panel of Evaluators. We have consulted with CHI throughout the feasibility study to make use of their expertise in programmes of this kind. CHI's experience and knowledge in this area and expertise measuring EET status and accommodation outcomes means they are well placed to deliver the impact evaluation of Staying Close.

When taking forward the further evaluation of Staying Close, for the formative aspect of the evaluation, CHI will explore other variables such as social connectedness, wellbeing and homelessness. Social connectedness in particular was a key outcome for the programme, but uncertainty around data collection and suitable measures made it unsuitable as a main outcome in the evaluation. Formative work will be undertaken to establish the most effective data collection strategies to use, and scales for measuring it with this cohort. We also recommend primary data collection for this outcome to allow for exploratory comparisons between young people receiving Staying Close and those in control sites.

CHI will also answer questions relating to mechanisms, subgroup and local effects using qualitative approaches such as interviews, focus groups and an exploration of contextual factors at the LA level through administrative data. An implementation process evaluation (IPE) will supplement the impact evaluation and focus on programme acceptability, sustainability and fidelity.

INTRODUCTION

Staying Close (Department for Education, 2023) is a programme of support for young people leaving residential care and transitioning to independence, developed by the Department for Education (DfE). The model (which is explained in further detail in the “Intervention” section) is designed to offer enhanced assistance in finding and maintaining accommodation and supportive relationships. This is designed to allow for a managed and gradual transition to independent living through the development of appropriate skills and improving wellbeing outcomes. The DfE intends for the programme to be comparable to Staying Put, a programme to enable care leavers in England to continue to be supported by their foster carers after the age of 18 (DfE, Department for Work and Pensions (DWP) and Her Majesty’s Revenue and Customs (HMRC) Guidance, 2013). Work on the programme is additionally informed by the government’s strategy on implementing recommendations made by the Independent Review of Children’s Social Care (2022); this includes a focus on improving wellbeing, increasing education, employment and training and reducing homelessness in care leavers.

In 2017, Staying Close launched as a pilot programme in five local authorities (LAs) and three independent providers. The programme was evaluated in each of the eight sites (Department for Education, 2022), with indicators of positive outcomes for participants across a range of health and independence-related outcomes. However, these were small-scale before-and-after comparative studies that could not establish causal relationships between the programme and these outcomes.

The DfE has since expanded the funding available for the programme, holding rounds of bidding in 2022 and 2023 to allow LAs to apply for funding to implement Staying Close for their care leavers.

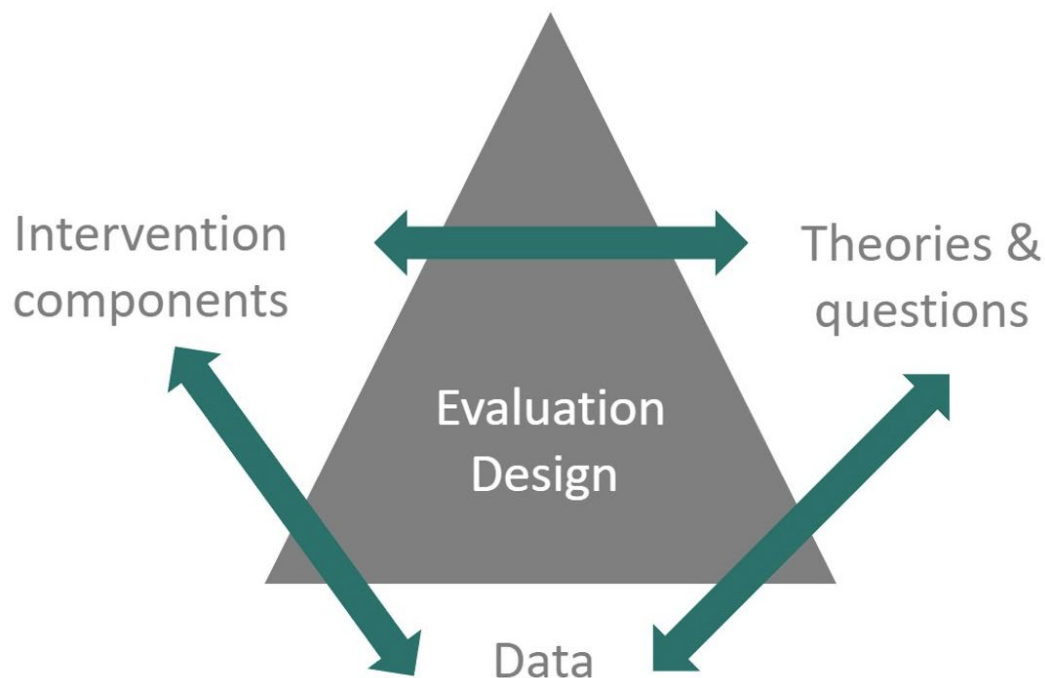
In May 2022 the Evaluation Task Force (ETF) awarded Foundations¹ funding to evaluate Staying Close as it is implemented in more LAs. The purpose of the ETF is to ensure that government spending decisions are supported by evidence, including funding robust evaluation of policies and programmes. To achieve this, we conducted this one-year feasibility study. This aimed to inform the upcoming large-scale evaluation, which will assess the impact of Staying Close on care leavers’ outcomes to establish whether the programme is effective and worth investing resources into, and where delivery can be refined ahead of a potential national roll-out. Both the DfE and ETF have supported Foundations with the development of the feasibility study. In particular, the DfE has supported communication between Foundations and the participating LAs and has presented their participation in the evaluation as an important part of implementing the programme.

¹ Funding was initially awarded to What Works for Children’s Social Care, which in December 2022 merged with the Early Intervention Foundation to become Foundations – What Works Centre for Children & Families (Foundations).

The feasibility study was conducted between July 2022 and May 2023 with the LAs that started implementing the programme in 2022. The first aim of this work has been to better understand the parameters and properties of Staying Close, as a programme that allows LAs considerable autonomy over how they use their resources for programme participants, and the implications of these for further evaluation. The feasibility study also aimed to establish the most appropriate research questions to ask in further evaluation, based on an understanding of the Theory of Change (ToC), and identification of appropriate methods and data that can be used to answer these questions. At the time of writing, some questions regarding the evaluation design and outcomes remain outstanding. Further feasibility work is going to be done over the summer of 2023 and an evaluation protocol will be published in the autumn of 2023.

This feasibility study is conducted with acknowledgement that there are constraints faced in evaluation, illustrated in Figure 1. Three competing constraints are in play for this evaluation: what is known about the programme, the questions that could be asked about the programme, and the data available about the programme. Each constraint influences what is possible for the other and has required multiple layers of research (both theory-building and methods testing) as well as multiple layers of synthesis of generated knowledge. This report aims to document both this research and synthesis, and the recommendations made about further evaluation by the evaluation team, given the constraints, and the criteria used in considerations.

Figure 1: Evaluation constraints



METHODOLOGY

This section describes the methodology used to explore each of the three constraints in Figure 1. Our aim was to solidify our understanding of the Staying Close programme in participating LAs and to build our theoretical understanding of how the programme works and understand how its impact may be robustly measured in subsequent evaluation. To build our understanding of the programme and prepare for subsequent evaluation, we identified seven specific questions that we intend to answer through evaluation, which can be informed by the feasibility study:

1. Does it work?
2. Does it work as expected?
3. Does it work differently for some groups?
4. Does it work differently in some places?
5. Was it implemented as intended?
6. Is it a good use of resources?
7. What else do we want to learn?

Our work for the feasibility study was divided into theory-building (qualitative-focused) and design testing (quantitative-focused) work; the methods used for each are described below.

Theory-building

We carried out a number of activities to generate data that would allow us to build our theoretical understanding of the programme and their implications for the mainstage evaluation.

Literature review

We conducted a literature review to identify existing research around support for young people leaving residential care, with the aim of understanding potentially important factors for the Staying Close programme. Research questions were:

- What are the experiences of children transitioning from residential care to independence?
- Which interventions show promise of improving outcomes for children leaving residential care?
- What are the key features, mechanisms, enablers and barriers of the interventions that show promise of improving outcomes for children leaving residential care?
- What are believed to be the most likely and relevant activities to improve outcomes and experiences of children transitioning to independence from residential care?

The full review is appended in Appendix 1.

Review of pilot site evaluations

We reviewed the evaluations of the programme conducted in the eight sites that started implementing Staying Close in 2017/2018. We extracted information from evaluation reports that could inform a subsequent implementation and process evaluation and an impact evaluation, and thematically analysed extracted data, using our seven evaluation questions as a framework (see the introduction to this section above). We also synthesised commonalities between any logic models that were included in evaluation reports.

A list of extracted data points and a write-up of the synthesis of pilot site evaluations is in Appendix 2.

Data collection tool

We created a data collection tool, in which we captured activities and support provided to care leavers by LAs in the feasibility study, both in business as usual (BAU) and as part of the Staying Close programme. The Care Leaver Local Offer website, Ofsted inspection reports and government guidance were used as sources for BAU data, and LAs' application forms that were submitted to the DfE when they were bidding for funding were used as sources for programme data. We additionally captured demographic characteristics about each LA using government statistics about children in care and care providers, and the Care Leaver Local Offer website.

The data collection tool is in Appendix 3.

Interviews with LAs

We conducted interviews with staff members from all 15 LAs included in the feasibility study. Twenty-eight interviews with staff were completed using semi-structured interview guides which covered programme activities, perceived impact, perceived mechanisms and barriers and facilitators to implementation.

Audio recordings of interviews were transcribed and uploaded to NVivo software for analysis. Following Braun and Clarke's six-step approach to reflexive thematic analysis (Clarke, Braun & Hayfield, 2015), we identified common themes across LAs related to the ToC. Analysts coded transcripts independently, coming to an agreed coding framework through discussion. We also used interview responses to clarify and check the accuracy of data points in the data collection tool. One focus group with staff and care leavers from an LA implementing Staying Close since 2017 was also conducted, with a focus on understanding mechanisms and impacts. This was analysed separately from one-to-one interviews (because the data generated differed considerably from one-to-one interviews), but informed thinking at the synthesis stage.

Analysts' reports from the analytic work are in Appendix 4.

Additional information sources

We also compared LA operationalisation of the programme to guidance from the programme developers (DfE) using the DfE's Staying Close Handbook and Expression of Interest templates, as well as direct discussion with DfE representatives.

We consulted with our Expert Advisory group (consisting of stakeholders with relevant professional and/or life experience in support for care leavers) throughout the process, checking understandings and interpretations with the group. We also held internal workshops with the project team to explore our theoretical understanding and expectations of the programme – for example, which subgroups or local effects we anticipated being of relevance to the evaluation. These expert opinions were weighted against data generated from the sources described above.

Once data generation had been completed, we synthesised data across all sources, and considered evidence that informed each of our seven evaluation questions (listed above). We considered findings in light of consistency of evidence, noting whether findings were present or absent across sources, and whether they were consistent or inconsistent across sources. Findings are described in the “Evaluation questions” section of this report.

Design testing

This section describes our approach to understanding how the impact of the programme might be measured robustly in the upcoming Staying Close evaluation. Below, we present the methods we used to address two key areas of the impact evaluation design:

- **Comparative methodology:** This considers the methods we can use to reliably answer our evaluation questions and estimate the impact of the programme, represented by different evaluation designs and their practical implications
- **Data sources:** This considers new or existing data we can collect to capture the impact of the programme.

Comparative methodology

In this part of the work, we considered what the most appropriate methods would be to measure the impact of Staying Close. Impact analyses involve comparing young people receiving Staying Close with another group of care leavers who are not receiving it. However, the strength of those comparisons depends on the similarity of the groups of young people whose outcomes we are comparing. The selection of young people forming the comparison group for those receiving Staying Close will depend on the research design selected for the evaluation. Young people assigned to a control group via a randomised design, for example, will differ from those identified via a quasi-experimental design (QED), and each design has associated theoretical and practical risks.

To identify the most appropriate comparisons, we first created an exhaustive list of all possible impact evaluation designs and systematically assessed their suitability given what we knew about the programme, its recipients and practitioners, and DfE’s roll-out plan (Appendix 5). We then assessed the most promising options according to their internal validity, ease of implementation and analytical simplicity to identify the most suitable designs (see Appendix 6). After holding consultations with external researchers with expertise in methods in similar evaluation contexts, we identified and ranked the most suitable evaluation designs. These are presented in the “Evaluation methods” section.

Data sources

In our data scoping work, we aimed to identify the most suitable data sources that allow us to make the desired comparisons and that capture the impact of the programme.

In the first stage of this work, we mapped all variables suitable for the evaluation, to be used either as outcomes, mechanisms or matching/control data, and their possible sources. To do this, we consulted with DfE stakeholders and partner organisations College of Policing (CoP) and Centre for Homelessness Impact (CHI), and held a workshop with Foundations’ Young Advisors Group, including care-experienced young people. The list of provisional outcomes identified can be found in Appendix 7.

To refine the list of variables for further testing in the final stage of the feasibility study, we held conversations with all 15 LAs and undertook a detailed assessment against the following criteria: alignment with ToC, internal validity, expected ease of implementation, expected missingness and timeliness of access (Appendix 7).

Following this round of shortlisting, we collected administrative and survey data from LAs to test our hypotheses on their acceptability, reliability, ease of implementation, distribution, timeliness and missingness. We also explored the feasibility of linking individual records in national datasets (see Appendices 7 and 8).

Economic analysis scoping

To understand the feasibility of an economic evaluation of Staying Close, we commissioned this scoping exercise to Bonin Analytics. This work aimed to:

- Assess the overall feasibility of an economic analysis of Staying Close alongside the main evaluation
- Develop a tool to capture intervention cost and activity data
- Present possible strategies for economic analysis and associated data requirements
- Evaluate the levels of uncertainty around each strategy and provide mitigations.

The detailed methodology and results can be found in Appendix 9 and recommendations are made in the “Overall synthesis and recommendations” section based on this work.

Synthesis and recommendations

After the theory-building and design testing work had been completed, we used project meetings to synthesise the two strands of work and to make recommendations for the optimal design of the upcoming large-scale evaluation. In deciding the optimal design, we considered the balance of risks (such as collecting incomplete or low-quality data) against the value of successfully testing a specific programme element (such as the impact of the programme on accommodation stability). In the discussion, we weighed options for evaluation design, explicitly considering benefits and drawbacks of different choices and combinations of choices. Our aim in this exercise and our reporting is to be transparent about the rationale for decision-making and risks associated with these choices to minimise possible sources of bias (particularly overconfidence bias).

To decide on our recommendations for how we would answer each of our seven evaluation questions in the mainstage evaluation, we considered multiple factors for chosen variables and design:

- Centrality to ToC
- Data access and characteristics
- Outcome maturation (whether we expect the outcome can be affected within the trial timeline)
- Suitability for cost analysis
- Stakeholder interest and policy relevance.

The full recommendations for the evaluation design, including all questions from our evaluation question framework described above, are presented in the “Overall synthesis and recommendations” section of this report.

PROGRAMME

This section describes and analyses Staying Close according to the TIDieR checklist (Hoffmann et al, 2014) and considers the risks and implications each aspect of the programme presents to the evaluation. The information presented was synthesised from DfE publications and discussions with the DfE Leaving Care team concerning current practice.

NB: In this document, Staying Close is referred to as a ‘programme’ rather than an intervention (see “What: business as usual”).

Why: what is the purpose of the programme?

What we know

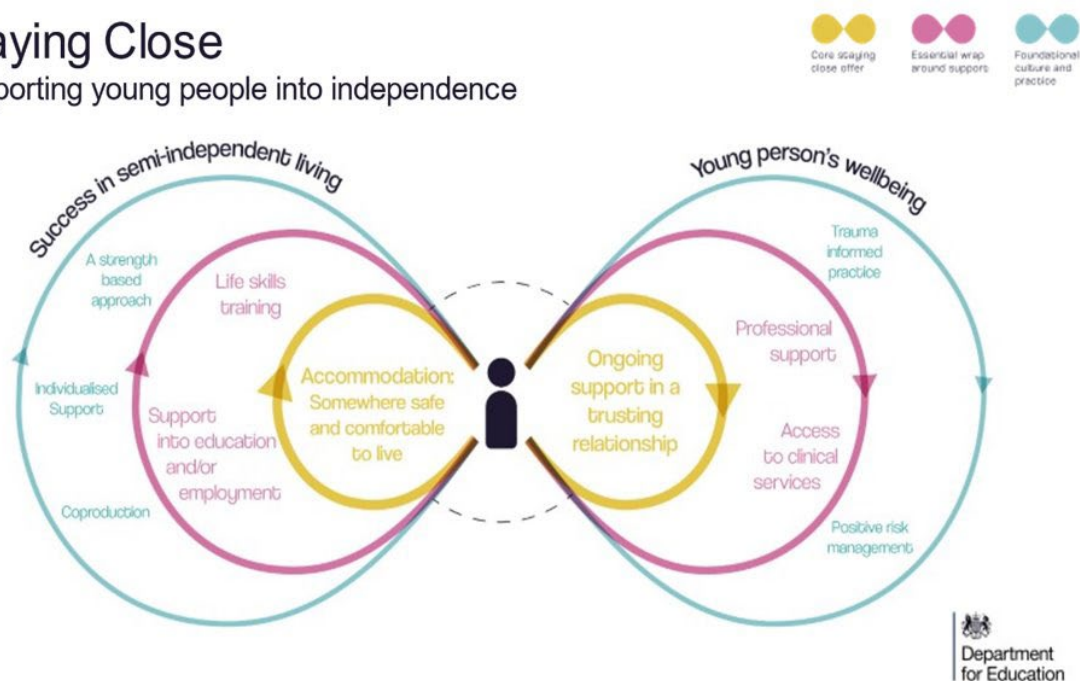
In programme documents, the objectives of Staying Close are represented in the form of concentric circles with core and peripheral objectives, as illustrated in Figure 2:

- **Core offer:** To ensure accommodation stability (reduced rates of eviction and fewer placement moves) and help young people build healthy relationships
- **Wrap-around support:** Access to clinical services, life skills training, professional support, support into education and training
- **Foundational culture and practice:** To encourage ‘positive risk management’, trauma-informed support, a strength-based approach, individualised support and co-production.

Figure 2: Staying Close objectives

Staying Close

Supporting young people into independence



Implications for the evaluation

Although Staying Close has been previously evaluated, it is unclear whether the evidence generated to date supports the programme’s current ToC. Given the governance of the programme (see “Who” below), we recommend assessing its operational alignment – i.e. the extent to which views and activities are aligned with the current programme objectives. This is discussed further in the “Does it work?” section.

In terms of data and methods, some of the programme’s objectives need to be better defined, both conceptually and operationally. Some outcomes relating to the programme objectives may be more difficult to ‘detect’ in an evaluation than others. This may be due to their position in the ToC (short-term outcome vs long-term outcome), the likelihood of getting valid data (both quantitative and qualitative), and the probability of detecting such outcomes given the sample size. See the sections on “Data sources and properties” and “Risks and strategies to mitigate”.

Who: young people

What we know

The young people taking part in the programme are those who are leaving or who have already left residential care in England after their 16th birthday and were in the care of their LA for at least 13 weeks since the age of 14. In the most recent funding round, LAs were permitted to extend their offer to other looked-after young people on a 'needs-led' basis, including for example unaccompanied asylum-seeking children (UASC) living in other forms of accommodation.

As of 2022, according to government statistics, there were a total of 82,170 looked-after children in England. Of these, 20,260 were aged 16 and above and 5,570 were UASC. The average number of children looked after across LAs in 2022 was 540.

In 2022, the number of young people aged 16 and over placed in total secure units, children's homes and semi-independent living accommodation was 13,010. The number of young people accommodated in residential care homes aged 16 and above was 380. In 2022, the total number of care leavers aged 17–25 was 45,050 in England. Of these, 20,370 young people were UASC.

According to the latest published statistics for 2022 in England, the most common reasons for care leavers to be looked after included risk of abuse or neglect, exposure to family disfunction or acute stress, child's or parent's disability or illness, low income and socially unacceptable behaviour. The population of children in care includes an overrepresentation of males, older children (39% were 10–15 years old and 25% were 16+) and children from Black, mixed or other ethnic groups. Children placed in residential care today are also more likely to be male and between age 14 and 17 when compared with earlier populations in residential care (Narey's Independent Review of Residential Care, 2016). They may also face more issues with their mental health and special education needs when compared with the entire care leaver population (Boddy, 2013). Care leavers may experience poorer health and mental health outcomes (Marmot Review, 2013) and increased risk of becoming homeless than children who were not looked after as a result of a lack of support in the years immediately after being in care (rather than the experience of being in care itself) (National Audit Office, 2015).

In this context, leaving care is understood as the period where children looked after by the LA make the transition towards independence. As part of this transition, young people who leave residential care may move towards independent or semi-independent accommodation or reunite with their families. However, it is unclear whether Staying Close will be a universal or targeted programme as the overall aim is to deliver support to as many young people as possible while, at the same time, LAs are asked to prioritise delivery based on assessment/perceived needs.

Implications for the evaluation

As our focus is on a relatively small population, there will be implications for sampling and inferences made. Young people may have little control over their engagement with LAs and may not have a strong understanding of the programme. Additionally, there are implications for identifying and constructing a comparison group in a quasi-experimental design (QED) due to differences in delivery capacity and eligibility criteria across LAs. There may also be a risk of selection effects if Staying Close is not ‘offered’ or taken up by all care leavers in the LA (see the section on “Counterfactuals and comparisons”).

Who: professionals

What we know

- **Who developed:** The programme was initiated by the DfE, based on one of the recommendations of the 2016 report on ‘residential care in England’ (Narey’s Independent Review of Children’s Residential Care, 2016). However, Staying Close is not a manualised intervention; it is a policy framework that (1) sets objectives and (2) provides resources for delivery partners (DPs) to achieve these objectives, but (3) let DPs decide how the objectives are achieved. In other words, Staying Close has been ‘co-produced’ or ‘co-developed’ with DPs
- **Who delivers (1):** LAs (with exceptions because, according to the DfE’s guidance, some LAs may choose to partner with private providers, charities or other voluntary organisations to deliver services). As explained above, the role of DPs is not limited to delivery. They have been actively involved in the development of the programme
- **Who delivers (2):** The ‘key worker/trusted person’ is responsible for supporting young people and for referrals/signposting
- **Who delivers (3):** Some LAs might work with or commission some activities from other organisations and providers (e.g. organisations running children’s homes or accommodation for care leavers or delivering training). See the section on “How well” below.

Implications for the evaluation

As the programme allows LAs to make use of external DPs, there may be less clarity and consensus on the ToC, when compared with a manualised intervention. It is possible that a DP’s understanding of the ToC is not entirely consistent with that of DfE or other DPs. This has several implications. First, it makes it more challenging for external stakeholders to understand what Staying Close ‘really is’ and to replicate the programme. It affects the evaluation of mechanisms (“does it work as expected?” in our evaluation questions) because these could be different across LAs. It is also possible that participants are ‘nested’ in LAs and that some of the variance in

outcomes can be explained by LA-level characteristics (e.g. the quality of local practices or teams). This is discussed further in the section on “Evaluation questions”.

In terms of frontline workers, from an evaluation perspective, the key worker is also a ‘gatekeeper’. This may affect survey response rates and care leavers’ engagement with the evaluation. This is discussed further in the section on “Evaluation methods”.

What: business as usual (BAU)

What we know

Local authorities have duties to care leavers. Some are legal (a personal advisor (PA), publishing a local offer for care leavers), others are discretionary (other forms of support). LAs use their judgement and core funding to decide how the policy should be implemented. Before Staying Close, some LAs offered more support than others. However, we did not detect consistent BAU patterns or models. We also observed that some Staying Close provisions interact with universal provisions (health services) or initiatives targeting broader groups (EET opportunities for disadvantaged young people).

Implications for the evaluation

In some LAs, Staying Close may be virtually indistinguishable from BAU, whereas in others it might significantly improve the local offer. It is possible that the effect of Staying Close is larger in the latter case. In the former case, it may be more accurate to say that we are evaluating the effect of receiving funding from DfE to deliver Staying Close.

Some outcomes may be wrongly attributed to Staying Close, whereas other Staying Close outcomes may be wrongly attributed to other services/initiatives. A clear understanding of BAU is essential to understand how the programme works and differs. We are also unable to say whether the mechanisms we propose in the section on “Evaluation questions” are causal. This is discussed further in that section.

What: programme

What we know

For many (but not all) LAs, Staying Close is an improvement and expansion of the care leaver local offer, rather than a set of new provisions. It is not intended to be mandatory, but rather to provide increased or better options for young people to decide where they live, who they maintain or strengthen relationships with, or what EET opportunities to pursue. The support is expected to be tapered as young people transition into adulthood. The model emphasises increasing supportive

relationships to bolster the support around young people leaving care, beyond support from personal advisors. This enhances the availability and choice in where young people can seek support. The types of relationships offered varies across LAs, often involving a Staying Close key worker or enabling residential key workers to stay in touch. Some LAs provide Lifelong Links workers as part of BAU, which has been integrated within their Staying Close offer.

Many LAs already offered provisions to aid young people to live independently and participate in EET opportunities as part of BAU. Staying Close expands these options, providing a wider range of choice in living arrangements and specialist workers such as housing, tenancy or education officers for additional support. Structured life skills programmes may be offered as part of Staying Close in addition to generic life skills support typically offered by PAs. Some LAs have chosen to expand provisions to support young people's health and wellbeing. In some areas, this may be introducing new specialist wellbeing practitioners or clinical psychologist roles, whereas others have chosen to invest in expanding wellbeing provisions already established as part of BAU.

The DfE expected both BAU and the amount and type of support provided to young people to vary significantly between LAs. Our research confirms this. It is likely that most young people are not aware of the details of the Staying Close model (because this is unlikely to be explained in detail by staff) being used in their LA but should be aware of their options for accommodation and will understand that a key worker has been assigned to them. It is unlikely that young people will be able to compare their experience of Staying Close with BAU, never having experienced BAU transitional support.

Implications for the evaluation

The programme assumes that there is an awareness of provisions, take-up and a positive experience of the process. These assumptions should be tested because experiences and outcomes may differ between LAs. This is further discussed in the section on "Evaluation questions".

How

What we know

A Staying Close handbook for LAs is available through the DfE; however, it is not clearly specified what support is available for LAs in national programme documents. It is for LAs to decide how the service/provision will be delivered and what training/qualifications are required for Staying Close staff. We expect variations between LAs, and that LAs will update their care leaver local offer. We also assume that key workers will work in co-ordination with other specialist professionals working within the model (such as EET workers) and be able to connect young people with other relevant services available through the LA's local offer for care leavers. We assume that some programme activities will require an active opt-in (e.g. they will be supported in their application to a training programme), but be something they choose to have support with.

Implications for the evaluation

The assumptions above need to be verified (see the sections on “Does it work as expected?” and “Was it implemented as intended?”). Additionally, as local authorities may already be providing services described in the Staying Close programme, some key workers will consider the programme as BAU. This could negatively impact their engagement with the evaluation. It is also unclear whether young people will be made aware of their options and the amount of choice they will be given (See “Does it work as expected?”).

When

What we know

Staying Close has been implemented in several stages between 2017 and 2023:

- **Wave 1, 2017/2018:** First Staying Close roll-out to eight pilot sites. Evaluations were commissioned and published for each site.
In April 2021, the government said it was “committed to the national roll-out of Staying Close in the future (House of Commons: Support for Care Leavers, 2016)”
- **Wave 2, August 2022:** Fifteen additional LAs receive funding for Staying Close and take part in the feasibility study from August 2022 to July 2023. In March 2023, the government confirmed its intention to roll out Staying Close and allocated an additional £8 million to the programme (HM Treasury Spring Budget, 2023)
- **Wave 3, May 2023:** Twenty-seven additional LAs are selected and expected to start implementing Staying Close by October 2023.

The upcoming large-scale evaluation will cover the years 2023–2025. However, we recommend further exploration of whether LAs in waves 1 and 2 could be included in certain analyses. These would be conducted as part of a quasi-experimental design (QED). This will depend on our selected outcomes and the existing data collected in these LAs for the population. The timeline and milestones for the evaluation are:

- October 2023: Evaluation starts; baseline data collection
- October 2024: Expected date for upcoming Spending Review. The DfE wants to have evidence ready to submit in a bid by April 2024 (assuming there is no general election at this time, which may push the Spending Review back)
- October–December 2024: Endline data collection
- May 2025: Expected report publication.

Implications for the evaluation

Accessing data via national datasets (see the section on “Evaluation methods”) could result in reporting timelines that are exceptionally delayed, due to the time it takes for LAs to feed information into these datasets and for it to become available to researchers approximately one year later. The delivery timelines for treatment LAs will stretch into 2024–2025, meaning that accessing that data would not be possible until late in 2025 or early 2026. However, national datasets may be a suitable contingency if data collection from LAs (and control LAs in particular) is challenging.

Even assuming administrative data can be collected directly from LAs, the full impact evaluation will not be completed in time for a Spending Review bid (March 2024). Other types of analyses (e.g. descriptive or qualitative data) can be conducted within these timescales so that we can generate some evidence that can be used in these bids. However, the timeline also indicates that it may be too early to measure some outcomes (e.g. homelessness).

If randomisation is not feasible, using a QED approach would involve identifying a separate control group of LAs not implementing Staying Close for comparison (see the section on “Evaluation methods”). This provides the option of including LAs from waves 1 and 2 of the roll-out, which would increase the statistical power of the analysis. The feasibility of this will depend on the existence and quality of pre-intervention data for our outcomes in these and other sites, and on whether evaluators can establish common outcome trends for a comparison group for these time periods. It is also possible that the publication of reports and other materials from wave 1 of the roll-out has increased the risk of spill-over to control LAs. As LAs in the control group may be delivering similar activities to those in Staying Close, it might be relatively straightforward for them to improve or refine their offer. This would risk invalidating the identification strategy for the mainstage evaluation, diluting the treatment effect and increasing the risk of failing to detect an effect of the programme. To mitigate this, BAU in all LAs identified as potential controls should be verified. This is further discussed in the “Evaluation methods” section.

Where

What we know

The Staying Close roll-out plan, including timelines and participating LAs, is in Appendix 10.

The following is a description of the roll-out process and selection criteria:

- The LAs selected for waves 2 and 3 submitted applications to the DfE’s call to receive funding for implementing Staying Close
- Applications were evaluated against the following criteria:
 - Quality of the proposed Staying Close model
 - Quality of their stakeholder engagement plan

- Quality of the project’s governance
- Quality of the delivery plan
- Bids that scored less than 3 (on a 5-point scale) on any of the sections were deemed unsuccessful. The DfE funded as many LAs with highest-scoring bids as they could with the available budget
- In wave 3, a number of LAs bids scored the same but there was insufficient funding to deliver Staying Close in each. The DfE therefore considered which combination of LAs would deliver Staying Close to most young people.

Implications for the evaluation

It is possible that Staying Close was first delivered in LAs where Staying Close was already similar to BAU. That would reduce the generalisability or the ecological validity of findings. Additionally, the large number of LAs already receiving Staying Close (see Appendix 10) presents a risk to a QED because it reduces the number of potential comparison sites that could be used as controls. Lastly, the way that LAs were selected to receive funding presents risks of selection effects. Specifically, LAs who can write better bids may be systematically different from LAs who wrote lower-quality ones in ways that influence the outcomes we are measuring, introducing bias into the design. Equally, selecting the combination of LAs that allowed them to deliver the programme to most young people may introduce a separate bias if LAs who deliver the programme at the lowest cost per young person differ from the wider set of LAs systematically.

How much, how often

What we know

We do not know much about the ‘intensity’ of the programme (e.g. how often young people are expected to meet/be in touch with their key worker, how long sessions will be, how long young people’s key workers are expected to be in touch for). As Staying Close is not mandatory, not all care leavers are expected to take up the offer and, when they do, they may not engage regularly or very intensively. Therefore, we expect variations. The ‘how much’ question does not seem relevant to other Staying Close components.

Implications for the evaluation

These are discussed in relation to subgroup effects in the section “Does it work differently for some groups?”

Modifications and tailoring

What we know

Since there is no manualised intervention, LAs have quite a lot of freedom to tailor their Staying Close offer to the needs of young people. These local variations were expected by the DfE, but the reasons for variations and whether they are correlated with outcomes or experiences are unclear.

LAs are also able to decide which of their care leavers they offer Staying Close to, and how many of their care leavers they offer it to; it is likely that LAs will offer the programme to young people who they believe will find the transition to adulthood most difficult.

Implications for the evaluation

The way the programme works is likely to be a 'black box' as we will be unable to determine statistically which 'version' of the programme is more/less effective. This makes the refinement of the programme, and future improvements, more difficult. However, which versions or aspects of the programme are perceived to be most effective, and where refinements can be made, should be explored via the IPE. Exploratory analysis could also be conducted to understand whether there are correlations between some model typologies and outcomes. This is discussed further in the section on "Data sources and properties".

How well

What we know

There is currently no evidence of quality assurance of the programme at the national and local level. There is a lack of clarity over how LAs will be monitored to ensure that they do not 'go rogue' and deliver a version of Staying Close not in line with the original intention. It is also possible that Staying Close could be delivered to young people who are not care leavers – i.e. those who are supported by the care system but do not meet the legal definition of a care leaver. Additionally, in some LAs Staying Close is going to be delivered by third parties (charities, private providers) who may have a different understanding of the programme and different quality criteria. There is a need for a safe mechanism or system for collecting feedback from young people to ensure quality.

Implications for the evaluation

If the programme is not implemented with high fidelity, or if a low proportion of care leavers benefit from it, the effect of the programme may be 'diluted' - i.e. evaluators may not be able to detect an impact. As a result, fidelity and compliance will be key questions in the IPE. This is further discussed in the section "Was it implemented as intended?"

EVALUATION QUESTIONS

This section aims to identify the most meaningful questions to be answered in the mainstage evaluation of Staying Close (October 2023–May 2025). To do this, it considers the most likely outcomes to be affected, the mechanisms expected to mediate the effect of Staying Close, whether certain groups are expected to benefit more/less from the programme, and the expected drivers and obstacles to implementation. Below, we present this information according to our seven evaluation questions.

The information has been drawn from four sources:

- A synthesis of the evaluations of wave 1 LAs
- The data collection tool
- Interviews with LAs who received funding to deliver Staying Close in 2022
- A literature review of interventions that improve outcomes for children leaving care.

1. Does it work?

From the programme developers

DfE guidance states that specific benefits should be delivered through Staying Close: access to improved housing options, reduced rates of evictions and fewer placement moves, positive changes in both mental and physical health, reductions in antisocial behaviour, episodes of going missing and criminal activity for young people at risk of these behaviours, increasing numbers of young people in EET (including through support gaining work experience, with application forms and advice and guidance), an increase in social connectedness and developing healthy relationships, and improvements in living skills (cooking, budgeting, time management and practical skills).

Although these are not all explicit behavioural outcomes that can be applied to a ToC, they do provide guidance on expected areas of impact. Impact areas have not been explicitly prioritised in guidance, but outcomes linked to safe and comfortable accommodation, and trusted relationships, would be expected to be prioritised because they are core parts of the offer.

From our research

Across sources, we identified six outcome areas that Staying Close may impact. By our definition, mechanisms are the participants' responses to activities within an intervention or programme; these responses then lead to behavioural change, which we define as outcomes. Sources often described mechanisms or short-term outcomes (such as young people living in a stable, suitable home) as opposed to behavioural outcomes (such as young people able to identify what they want

from their home) as a result of programme activity. We have interpreted these to state what we believe to be the key behavioural outcomes in each area.

- **Accommodation:** Consistent evidence across three sources (previous evaluations, interviews, literature review) for this as an outcome.
Sources identified care leavers' satisfaction with accommodation, suitability of accommodation and stability of accommodation as important short-term outcomes.
We believe the key outcome is a care leaver's ability to maintain a tenancy in a home that is suited to their needs
- **Relationships:** Consistent evidence across three sources (previous evaluations, interviews, literature review) for this as an outcome.
Sources identified increased connectedness and social support, and increased ability to form and maintain healthy relationships as important short-term outcomes.
We believe the key outcome is a care leaver's social network – their ability to form and draw on supportive relationships
- **Crime and safety:** Limited but consistent evidence across two sources (evidence synthesis, interviews) for this as an outcome.
We believe the key outcome is whether a care leaver becomes involved in the criminal justice system as a perpetrator
- **Independent living skills:** Mixed evidence across two sources (evidence synthesis, interviews) for this as an outcome.
Sources identified household management, managing finances and cooking skills as important short-term outcomes.
We believe the key outcome is whether a care leaver is able to live independently.
- **EET:** Limited mixed evidence across two sources (evidence synthesis, interviews) for this as an outcome.
We believe the key outcome is whether a care leaver is in employment, education or training.
- **Wellbeing:** Limited evidence across two sources (evidence synthesis, interviews) for this as an outcome.
Sources described both physical and emotional wellbeing as important outcomes. We believe the key outcome is a care leaver's ability to appropriately care for their physical and emotional wellbeing.

Outcomes derived from sources reflect the intended benefits outlined by DfE guidance. Sources did not provide any data about unintended consequences; we see this as something important to explore in subsequent evaluation.

2. Does it work as expected?

From the programme developers

DfE guidance does not describe explicit mechanisms through which they expect Staying Close to have an impact.

From our research

We were able to identify a number of activities and programme components using all four data sources:

- **Accommodation:** Offer of suitable accommodation, support/training in household management, being a good tenant, staff acting as a safety net if things go wrong
- **Relationships:** Being able to maintain supportive relationships with staff across the transition to adulthood, supportive culture within programme, more opportunity to ask for and be proactively offered support
- **Crime and safety:** Secure and safe accommodation, having a trusted person for advice
- **Independent living skills:** Personalised, flexible and intensive support/training
- **EET:** Personalised, flexible and intensive support/training, fostering self-belief
- **Wellbeing:** Personalised, flexible and intensive support, reduced isolation.

However, we were not able to use sources to effectively identify “mechanisms” in the way that we would use them for a ToC. Additionally, it is difficult to identify consistent mechanistic pathways for a programme that varies considerably from site to site and therefore includes a range of activities in a range of contexts.

Sources (interviews and the evidence review) also identified “indirect mechanisms” that acted as both outcomes and mechanisms:

- **Accommodation:** Stable and suitable accommodation improves wellbeing (basic needs being met, desired environment), reduces chances of criminality (care leavers are matched appropriately and in a safe area), improves relationships (located near family/friends, staying in one place allows community relationships to develop) and facilitates EET (having a stable address in a suitable area allows you to apply for local jobs/courses)
- **Trusted relationships:** Trusted relationships with Staying Close staff improve the chance of securing appropriate accommodation, improve wellbeing (decreased isolation, increased resilience), reduce chances of criminality (appropriate guidance), improve relationships (modelling positive supportive relationships), improve independent living

skills (one-to-one training and support) and facilitate EET (one-to-one training and support)

- **Wellbeing:** Better wellbeing improves the chance of maintaining appropriate accommodation (more likely to be able to manage a home), reduces chances of criminality (less likely to seek illegal activities such as drug use), improves relationships (more likely to seek social interaction and maintain relationships) and facilitates EET (more likely to seek opportunities and maintain a job/training).

As a team, we proposed the following broad mechanisms, which were not derived from our data sources, but which we believe are feasible based on our understanding and experience having conducted the study. These mechanisms are intended to be ones that are applicable to the broad intervention group, but also allow for some comparison against the control group receiving usual post-18 support:

- **Awareness:** Whether care leavers are aware that they can receive Staying Close
- **Experience:** The experience of care leavers taking part in Staying Close
- **Take-up:** Whether care leavers being offered are Staying Close are engaging with it.

3. Does it work differently for some groups?

From the programme developers

DfE guidance does not identify specific subgroups that may be impacted differently by Staying Close.

From our research

There are no consistent subgroups identified across sources. Interviews identified care leavers who are parents or have special educational needs and disabilities (SEND) as potential subgroups who may have different outcomes from the programme due to their specific needs. In our internal exploratory workshops, we suggested age (or life stage), UASC and SEND status, and care history could be important factors.

4. Does it work differently in some places?

From the programme developers

DfE guidance does not identify any local effects that may impact on Staying Close.

From our research

We used the data collection tool to identify variations in Staying Close across LAs. Variations are extremely common, because each LA has been able to shape the programme in a way that works best for their local context. There are several themes of variation found in a considerable minority of LA across the sample that may have an impact on outcomes:

- **Accommodation:** The type of housing on offer varies across LAs, with some solely offering Staying Close accommodation, others taking on private tenancies on behalf of care leavers or having arrangements with the council to have a stock of council flats specifically for care leavers. The additional choice some LAs have on offer may mean that care leavers may be more likely to find a home that suits them.
Urbanicity may also impact accommodation-related outcomes, with urban areas having fewer affordable options for care leavers to choose from if not taking up Staying Close accommodation, whereas rural areas may have fewer accommodation options to choose from overall.
- **Relationships:** The point at which the key relationship was introduced varied across LAs. For some, key relationships came from staff working in children's homes (and so knew young people before they left care); for others, key relationships were introduced once a care leaver was in the Staying Close programme. This difference in the length that a key relationship has been established for may affect relationship-related outcomes
- **Wellbeing:** Some LAs had bought out time with a clinician, or hired a full-time specialist trained in mental health (to varying degrees of seniority), whereas others had not
- **Independent living skills:** Some LAs had hired a specialist worker to support care leavers with their development of independent living skills, whereas others had not
- **EET:** Some LAs had hired a specialist worker to support care leavers with their employment/education/training, or had local partnerships with EET programmes, whereas others had not.

5. Was it implemented as intended?

From the programme developers

DfE guidance does not specify details about how Staying Close should be implemented, such as engagement targets or quality assurance.

From our research

Across all sources, barriers and facilitators to implementation were identified:

Barriers

- Lack of housing that LAs can use for Staying Close accommodation – consistent evidence across three sources
- High staff turnover, limiting ability to form trusted relationships with care leavers – consistent evidence across three sources
- Unfilled staff roles for both specialist work (such as mental health workers) and key workers – consistent evidence across three sources
- Poor engagement with care leavers, limiting care leavers' willingness to use Staying Close resources – consistent evidence across two sources
- Poor matching for care leavers in shared accommodation – consistent evidence across one source.

Facilitators

- Availability of suitable accommodation – consistent evidence across three sources
- Staff being trained in and using trauma-informed care – consistent evidence across two sources
- Co-production of Staying Close with care leavers – consistent evidence across two sources.

Using Proctor et al's (2011) conceptual framework for implementation outcomes, we have grouped these barriers and facilitators according to criteria that may impact on implementation:

- **Acceptability:** Co-production, poor engagement with care leavers
- **Feasibility:** Poor matching, trauma-informed care, availability of suitable accommodation
- **Fidelity:** Lack of housing, unfilled staff roles
- **Sustainability:** High staff turnover.

6. Is it a good use of resources?

From the programme developers

There is limited evidence of value for money or cost savings from any source. Pilot evaluations conducted limited cost–benefit analyses indicating that there may be a net saving, or equal existing costs. Interviews suggest that learning and new models for practice developed by an LA (e.g. a life skills development programme) as part of the Staying Close programme could be applied across the entire care leaver population within that LA, representing potential savings. Interviewees also expect that there would be a drop in numbers not in education, employment or training (NEET) and in involvement with the criminal justice system, representing a general saving to the economy.

From our research

Across sources, there was limited evidence related to cost savings. From interviews with LAs, some expected that learning and training for staff working on the programme may positively impact work with the wider cohort of care leavers in the LA, representing a potential cost saving. Staff also noted that they expected reduced criminality and NEET could represent a cost benefit to the state.

7. What else do we want to learn?

Through the data collection tool and interviews with LA staff, it was not possible to typologise Staying Close. This is likely to be more possible in further evaluation; with the addition of 27 more LAs, it is likely that the saturation point for variation in possible models of the programme will be reached. Identifying and typologising models of the Staying Close programme would have potential benefits to the DfE, LAs and further evaluation. For the DfE, specific models of the programme allow for clarity around the mechanistic pathways that result in positive change and facilitate quality assurance through easier monitoring. Should the programme be rolled out nationally, LAs setting up their own Staying Close programme would have templates to refer to. Further evaluation would also benefit by having specific models to refer to, requiring less theory-building work and potentially allowing for comparison of the effectiveness of models.

EVALUATION METHODS

This section answers the question of how the impact of the programme might be robustly evaluated. As described in the section on “Design testing”, we systematically assessed the suitability of a range of experimental/quasi-experimental methods, outcomes and data sources and strategies to access or generate them.

Below, we present the findings from this strand of work focusing on the:

- **Counterfactuals and comparisons:** This considers the comparisons which can be made in each part of our evaluation design and the most reliable methods to achieve them
- **Data sources and properties:** This considers what new or existing data we can reliably collect to make these comparisons and measure the true impact of the programme
- **Risks and mitigations:** This outlines the main risks identified to the success of the evaluation and proposes strategies to mitigate them.

Counterfactuals and comparisons

Next, we present the methods we ranked as most suitable for evaluating the programme, according to our evaluation questions. To start, we focus on the first question (“Does it work?”) to discuss the most appropriate comparisons, sample size and identification strategies for measuring the impact of Staying Close. Then, we consider the other evaluation questions regarding mechanisms and subgroup and local effects.

Comparisons

The first evaluation question, “Does it work?”, concerns our ability to measure the net effect of Staying Close. Answering it requires comparing young people receiving Staying Close and those who are not, and the strength of those comparisons depends on the similarity of the groups being compared and the number of people making up those groups. Our evaluation design scoping work (described in the section on “Design testing”) concluded that the strongest methods we could use to understand the average effects of Staying Close are:

- **Cluster RCT:** Young people from LAs randomised to receive Staying Close are compared with young people in LAs randomised to the control group in a two-armed trial
- **Difference-in-differences (DiD) design:** Young people in LAs receiving Staying Close in 2022 and 2023 are matched to and compared with young people in LAs identified as suitable comparator LAs not offering Staying Close (with parallel trends in outcomes).

A **cluster RCT** was considered the most suitable method because it provides a control group that is, in expectation, as similar as possible to the treatment group on both observed and unobserved

characteristics. This would enhance evaluators' ability to detect the true impact of the programme. Although randomising within large units, such as LAs, increases the risk of group imbalance, this can be mitigated by stratifying randomisation by variables expected to influence the effectiveness of the programme, such as Ofsted ratings and urban/rural categorisation. Additionally, an RCT offers analytical simplicity because the data would be processed and analysed internally, regressing the treatment variable against selected outcomes to estimate the impact.

If randomisation is not feasible, a **DiD** design serves as the next best option. DiDs construct a comparison group that is as similar as possible to the treatment group on observed characteristics by using existing data to identify the control individuals for the analysis. (Unlike a randomised design, however, it rests on assumptions that unobserved characteristics are also equivalent in the treatment and control groups.)

A major advantage of DiDs over other QEDs is their ability to address selection bias when comparing LAs delivering a programme with those that do not. By matching LAs based on similar outcome trends over time (parallel trends), the analysis accounts for LA fixed effects (time-invariant heterogeneity between LAs, such as local socioeconomic conditions), that may influence their selection into the programme. In the case of Staying Close, the DiD design could also include all LAs implementing the programme since 2018 (assuming suitable comparator LAs with parallel trends are identified), significantly increasing the number of young people included in the analysis and evaluators' ability to detect the programme's effects.

The choice between a DiD design and other quasi-experimental methods such as Synthetic Control Method (SCM) was influenced by the availability of data and by its analytical implications. DiDs typically rely on fewer assumptions than SCMs and are less technically and analytically complex than SCMs. However, evaluators can continue to explore the use of the SCM or other matching designs as contingencies to the DiD. If, for example, there is insufficient pre-intervention data on our outcomes of interest, we may revert to an alternative design.

Sample identification strategies

In an RCT the assignment of treatment and control individuals is conducted as follows:

1. Delivery partners who have access to the target population identify eligible individuals
2. Once the eligible participants are identified, they are randomly assigned to either the treatment or control group
3. Baseline data are collected from all participants to establish a pre-intervention benchmark
4. Following the implementation of the programme, endpoint data are collected to measure the desired outcomes.

DiDs rely on retrospective identification strategies for sample selection, enabling researchers to identify suitable comparison ('control') groups against whom to compare the treated individuals. This involves several steps:

- Evaluators identify comparator LAs based on historical data on the outcomes of interest. For an LA to be considered suitable, it must demonstrate parallel trends in the outcome variables without either being exposed to the treatment. The control group should provide a valid counterfactual for estimating the treatment effect
- Baseline data are collected from both groups before the implementation of the programme to establish a baseline against which subsequent changes can be assessed. This can be collected retrospectively if the outcome is based on administrative data and if data are collected at the right time by the data processor (when the programme started)
- Evaluators identify young people from the eligible population in comparator LAs who match the treatment sample on certain observable characteristics such as age or other relevant factors
- Following the programme, endpoint data are collected to capture the outcomes of interest.

Sample size and power

Statistical power measures the ability of a study or experiment to detect an effect or relationship when one exists. It is partially determined by sample size: the more participants an analysis includes, the more likely it will be that it can detect a programme's effects. Considering the number of LAs that applied for funding, the eligible population size and the budget allocated by the DfE to Staying Close, we estimated that approximately 1,800 young people would be included in the **cluster RCT**. The treatment group-to-control group ratio was anticipated to be roughly 2:1, due to policy requirements to spend the allocated Staying Close funding by a given deadline. Accounting for other factors such as the intra-cluster correlation coefficient (ICC – assumed to be between 0.02 and 0.05) and the predictiveness of baseline data (estimated R^2 of 10%), the minimum detectable effect size (MDES) was expected to range from 0.18 to 0.22 standard deviations for (relatively complete) administrative data. However, for survey data, the MDES was projected to be between 0.25 and 0.5 standard deviations (when using a range of assumptions for survey response rates and ICC – see Appendix 11 for more information).

The sample size for the **DiD** is currently uncertain. With the increased funding from HM Treasury (HMT), more LAs will be implementing Staying Close, resulting in a larger treatment group. It is plausible that data from LAs that initiated Staying Close delivery between 2018 and 2021 can be included in the treatment group, potentially increasing the sample size and statistical power of this design. However, the feasibility of this approach relies on the availability of historical data for the selected outcomes, which we continue to explore. This requires not only pre-intervention aggregate LA data, but also individual-level pre-intervention data in treatment and control sites (i.e. young people who would have received Staying Close if they were born a few years later in treatment sites for comparison with young people who would have been our control sample if they had been born a few years earlier in comparator sites). It is also currently unclear how many suitable comparator LAs will be identified, and therefore how large the control group will be. This will also have a large influence on power.

Other evaluation questions

Besides measuring the impact of the programme, we also recommend answering our other evaluation questions focused on the mechanisms of Staying Close and subgroup and local effects. These questions involve different types of comparisons, such as comparing different subsets of young people receiving the programme:

- **Mechanisms** (does it work as expected?): We recommend that evaluators plan to make the same comparisons as for the main average effects to test out mechanisms, comparing the treatment and control groups. These groups will either be assigned via randomisation (RCT) or control LAs and young people will be identified using historic data (QED). If appropriate data at scale cannot be collected (see “Data sources and properties” and “Risks and strategies to mitigate” below), qualitative analyses can be used to understand how practitioners and recipients felt about the hypothesised mechanism, without assessing their causal influence. In this process, there must be a focus on minimising the burden on both practitioners and young people
- **Subgroup effects** (does it work differently for some groups?): This relies on comparing different subsets from the treatment group with the same subsets in the control group, or with other young people receiving Staying Close. Although LAs are confident they can identify relevant subgroups, we anticipate that evaluators will not be sufficiently powered to estimate these effects as part of the impact evaluation. Instead, exploratory analyses can be conducted to see if there are correlations between subgroups and outcomes, and qualitative analyses
- **Local effects** (does it work differently in some places?): This involves comparing young people from different subsets on factors identified as significant by the ToC work (e.g. deprivation or housing market conditions) in the treatment group with young people in the comparison group sharing the same characteristics, or other young people in the treatment group. Although we are confident in being able to categorise LAs according to certain characteristics, we anticipate that evaluators will not be sufficiently powered to estimate interaction effects of the programme and local factors. Similarly to the subgroup effects above, exploratory and qualitative information can be provided instead.

Data sources and properties

The data framework work (described in the section on “Design testing”) aimed to identify the most suitable variables and data sources for measuring the impact of Staying Close by assessing them on their acceptability, reliability, ease of implementation, distribution, timeliness and missingness. As with the evaluation design findings described above, we describe findings as they relate to the first evaluation question (“Does it work?”) as well as variables that could be used for matching or control in the analyses.

The final recommendations for the evaluation design, which synthesise these findings with those from the “Evaluation questions” section and consider the policy context, are detailed in the “Overall synthesis and recommendations” section.

Does it work?

To answer the first evaluation question, we explored three data sources: LA administrative data, national datasets and survey data.

Local authority administrative datasets

We requested the following variables from LAs as candidate outcomes:

- Number of episodes of homelessness
- Length of each homelessness episode
- Number of address changes
- Reason for each address change
- NEET status
- NEET category.

We also collected information that could be used either to match young people to comparable controls and as control variables. The full list of variables we requested from LAs is in Appendix 12.

We received administrative data from seven out of the 15 LAs. Low response rates and delays in data sharing present a threat to evaluation timelines and sample sizes. However, the analysis of the data collected (n=83 young people) indicates high data quality and low levels of data missingness for these variables. The NEET category stands out with 74.5% missingness, with the rest ranging from 0–3% (Appendix 12). This was due to a formatting error in the data collection form that did not allow LAs to fill the respective column, despite LAs having confirmed access to this information. Considering that there were no missing values for the NEET status, we expect that to be the case for NEET category as well for future data collection, given that LAs record these details simultaneously. To remedy this shortcoming, we provided LAs with instructions on how to reformat the respective column. However, two LAs were unable to do so. We have no concerns over missingness for matching or control variables.

Based on discussions with all LA project managers we have no major concerns that LA administrative data would be of sufficient quality or reliability for use as primary outcomes. They identified low staff capacity, internal miscommunications and high turnover rates within teams as the main barriers to timely administrative data returns.

We did not entirely simulate the data collection we plan to conduct during the impact evaluation because we did not include LAs not delivering Staying Close. However, control LAs are likely to be less engaged with the evaluation. This may be particularly likely in a DiD. Therefore, we anticipate that data collection from control LAs will be a resource-intensive process with substantial risks

with respect to the timeliness and completeness of data returns. This risk and its implications for the subsequent evaluation are discussed in the section on “What else do we want to learn? (Formative research)”.

National administrative datasets

We also considered the use of national databases such as the Department for Levelling Up, Housing and Communities’ Homelessness Case Level Information Collection (H-CLIC) to access the above variables of interest.

We quickly determined that the timelines for accessing national databases would not be compatible with reporting requirements for the DfE. However, we will continue to scope use of this data as a mitigation for risks of LA administrative data collection. This data may also provide an opportunity for long-term follow-up analysis on outcomes captured in these datasets, such as homelessness.

A list of variables we would be able to access through national datasets is in Appendix 7.

Survey data

We also launched surveys for young people in all 15 LAs to test this form of data collection for other potential outcomes based on subjective experiences:

- Mental health and emotional wellbeing
- Confidence in life skills
- Sense of connection and relationships
- Feeling supported
- Financial stability
- Satisfaction with accommodation
- Readiness for independence.

Details of the survey instruments tested, our strategy to create and distribute the surveys, and their analysis are in Appendix 13.

Our findings highlight a high level of risk associated with using survey data as primary outcomes for the evaluation due to the low response rates (n=49, from ten out of 15 LAs). Discussions with LAs revealed several barriers encountered during the survey distribution process. These included overlapping timings with Ofsted surveys for young people, limited staff capacity for distribution and support, and delays in recruiting Staying Close staff and identifying eligible young people.

We ruled out contacting young people directly to collect this data. The project team felt this would be unethical without permission, and to establish this permission beforehand would be extremely resource-intensive. As a result, LA staff acted as gatekeepers for this aspect of the data collection.

If outcomes not captured by administrative data are deemed crucial to the ToC, survey usage may be considered with an enhanced plan for data collection processes and strategies. However, relying on this data would pose substantial risks to the evaluation. Mitigation measures could include

selecting alternative outcomes that do not rely on survey data and developing an enhanced data collection strategy and communication plan to boost response rates. For example, we might consider requesting LAs and key workers collect this information themselves and asking LAs to include it as part of their administrative data returns.

Notably, survey data cannot be used in DiD evaluation, because this method requires pre-intervention data for identifying LAs with parallel trends in outcomes.

Other evaluation questions

The data sources described above were also considered for answering questions about the programme's mechanisms, subgroup and local effects:

- **Mechanisms (does it work as expected?):** Survey measures were tested before the ToC work was completed, so the variables we chose to test in the survey were both candidate outcomes and mechanisms. Therefore, these findings also apply to our mechanisms question
- **Subgroup effects (does it work differently for some groups?):** This information will be obtained from LA administrative data returns. Our results showed no missingness for this information and LAs were confident in their ability to collect it. UASC status and protected categories are some examples of variables that will be used for this purpose
- **Local effects (does it work differently in some places?):** We anticipate that most variables for this analysis will be available either through publicly available data (e.g. levels of deprivation) or from LAs (e.g. the type of housing provided to young people).

Risks and strategies to mitigate

We identified five key risks to the validity of the evaluation through the methods and data scoping work described above. These are described below, followed by identified strategies for their mitigation:

- The risk that young people offered and taking up the programme may differ systematically from those who do not take it up, posing a threat to the external validity of the treatment effect for the wider population. To reduce this risk, we recommend evaluators use an **Intention-to-Treat (ITT)** analysis approach to assess the programme's impact based on eligible young people's initial assignment, regardless of whether they received Staying Close. This method helps preserve real-world conditions, where not all young people assigned to the Staying Close group will actually engage with the programme throughout the entire study period. As policy decisions should be based on the effectiveness of programmes in the general population, rather than just those who are most engaged, an ITT analysis will provide the most relevant impact estimate for policymakers.

- The risk of systematic differences between LAs in the treatment and control groups, which may have caused the former to apply for funding compared with the latter. To address this risk in an RCT, we suggest using stratified randomisation, considering important LA-level factors to mitigate the potential failure of randomisation. For the DiD design, parallel trends testing should be conducted to assess the comparability of the treatment and control LAs.
- The risk of significant differences in the use of residential care between LAs, such as in their in-house provision or out-of-authority protocols. We expect this to emerge from regional variations and uneven density of residential care homes across the country. Stratified randomisation, based on location, and LA selection based on parallel trends should also help mitigate this risk.
- The risk of inaccurate estimation of the programme’s impact on the selected outcomes due to low statistical power. This risk was made particularly acute given the constraint of the LA-level unit of delivery. As described previously, we anticipate sufficient statistical power when using administrative data for the outcomes of interest. This risk can be mitigated further by collecting baseline data to preserve power. However, the risk is higher for survey data due to high expected levels of missingness.
- In March 2023 DfE announced that more LAs will be funded to deliver Staying Close in 2023–2025. This increased the risk that we will not identify sufficient control LAs from the LAs not delivering the programme. Reducing the pool of potential comparator LAs threatens evaluators’ ability to make meaningful comparisons between treatment and control young people as part of the DiD. This risk cannot be convincingly mitigated. (The number of LAs delivering the programme at each stage of DfE’s roll-out can be found in the “When” section above).

These identified risks and mitigation strategies provide a comprehensive assessment of the challenges and potential limitations faced in the evaluation. It is crucial to carefully consider these factors when interpreting the findings and drawing conclusions from the study.

OVERALL SYNTHESIS AND RECOMMENDATIONS

The work on identifying which questions we should ask in the evaluation and the strand focused on which methods we should use to answer them were conducted independently. Different researchers worked on each strand, with conclusions shared only at the conclusion to ensure findings weren't influenced or biased by other sources of information. In this section we synthesise the findings from these two questions to produce recommendations for a trial design that can most robustly evaluate the Staying Close programme.

Evaluation method and classification

After assessing the evaluation design framework (see “Design testing” and Appendix 7), we concluded that a cluster RCT is the preferred evaluation method for Staying Close. However, this design was incompatible with the changes made to the government's plan to roll out Staying Close announced in March 2023. Specifically, the DfE was tasked with starting to deliver the programme to a larger number of LAs, meaning LAs who would have been assigned to the control group instead needed to deliver the programme from 2023.

We are therefore proposing a QED to evaluate Staying Close. Our findings (“Counterfactuals and comparisons” section) found that the most suitable method was the **DiD**. Assumptions required for this design will be tested more comprehensively in the next stage of the project and the final design decisions and their rationale will be outlined at the protocol stage (see the “QED evaluation components” section for more information). We will also consider the use of an SCM (as described in the “Counterfactuals and comparisons” section) if we consider the DiD assumptions to be implausible.

As for the scale of the project, we recommend that this should be **considered a pilot evaluation**. This means that the evaluation will be both summative (aiming to answer ‘impact’ questions about the programme's effect on outcomes) and formative (aiming to answer questions about which methods are most suitable to evaluate it – e.g. how best to collect data from hard-to-reach participants).

We believe there are compelling reasons to recommend a pilot design. Some aspects of the evaluation, such as the use of survey data for one of our three outcomes (see below), entail a high level of risk, as evidenced by findings from our data collection testing detailed in the “evaluation methods” section. Large-scale data collection from the care leaver population has never been attempted before. It is also currently unclear what the most suitable scales are for subjective measures for this cohort and what the properties of this data would be.

There is also an increased policy focus on programmes which build relationships for care leavers. The recent Independent Review of Children’s Social Care² recommended that “young people leaving care (should) have at least two loving relationships to support them” (p.156). The DfE is likely to commission more evaluations of programmes for this cohort in the future. As a result, findings from a pilot evaluation aiming not only to evaluate the programme but also to identify effective methods for future evaluations for this cohort will be extremely valuable.

Formative aspects could involve, for example, the development of measures pertaining to social connectedness and the exploration of data collection methods.

We are therefore proposing a pilot design with two primary objectives:

1. To estimate the impact of Staying Close on the accommodation changes and EET of young people
2. To test various strategies for data collection and different measures of wellbeing and social connectedness.

The decision to proceed with a DiD rather than an RCT also influenced this recommendation. Survey data cannot be reliably used for DiD impact estimates. Therefore, this pilot provides an opportunity to test different strategies and methods for generating this kind of data without presenting any new risks to the quality of our findings. Including formative elements in the evaluation will increase the chances of identifying successful mechanisms for data collection, setting precedent for future evaluations with a similar cohort or working with LAs in this way. See the “What else do we want to learn? (Formative research)” section for full recommendations for the formative aspects of the evaluation.

Other design choices

Findings from the sections on “Evaluation Questions” and “Evaluation Methods” also informed a series of other trial design and analytical decisions.

Sample eligibility

The target population consists of care leavers or future care leavers between the ages of 16 and 25. However, not all outcomes might be suitable for 16–17 year olds, as some may still be living in residential care homes. We recommend including other young people who are offered Staying Close by their LA. Specifically, LAs were permitted to extend their offer to other looked-after young people on a ‘needs-led’ basis, including for example UASC living in other forms of accommodation.

²https://webarchive.nationalarchives.gov.uk/ukgwa/20230308122820mp_/https://childrensocialcare.independent-review.uk/wp-content/uploads/2022/05/The-care-experience.pdf

To be included in the analysis, all young people should have been receiving Staying Close for at least six months.

ITT vs CACE

In the “Risks and strategies to mitigate” section, we acknowledge the potential risk of systematic differences between young people accepting or opting into Staying Close support and those who do not. Differences in important characteristics, such as their needs or the quality of their previous care experiences, might influence the effectiveness of the programme, which risk biasing results if only compliers are included in the analysis.

Therefore, we recommend evaluators adopt an intention-to-treat (ITT) approach. This approach is most useful for policy because it estimates the effect of the programme on the entire population. The alternative, complier average causal effect (CACE) analysis, includes only those young people who receive the intervention in the analysis. This would introduce a risk of bias (for the reasons described above) and would therefore reduce the generalisability of our findings.

The primary challenge lies in identifying and collecting data from an equivalent sample in control LAs, which is further complicated by the inclusion criteria based on needs introduced in the most recent round of DfE funding. To mitigate the risk of bias, the IPE should focus on establishing clear and consistent criteria for eligibility across the treatment LAs, minimising potential differences between young people in the treatment and control groups in the analysis. (See the “Counterfactuals and comparisons” section for a detailed discussion of the risks and strategies to mitigate them.)

Length of trial

The start dates for Staying Close are likely to vary across different LAs. The DfE anticipates that LAs will start delivering the programme in October 2023, though some activities may begin earlier in a small number of LAs where their BAU practices are more similar to their Staying Close model. Therefore, we recommend that baseline data collection starts in October 2023 for LAs who have started delivering the programme, and extending this for programmes that start later. The latest that LAs can start delivering the programme and be eligible for inclusion in the trial will be January 2024.

We are recommending that the trial length is 12 months in each LA, which we believe is sufficient for our selected outcomes to have an effect and is compatible with the DfE’s expectations for the project. Therefore, we propose that the trial should be conducted from October 2023 to October 2024, assuming that all LAs start delivering the programme in October 2023. If LAs begin delivery later than this (up to January 2024), we recommend allowing for 12 months of delivery. Final endpoint data collection will therefore be in December 2024.

As described in “ITT vs CACE ” above, young people will be eligible if they receive six months of the intervention. Therefore, participants will be included if they start receiving the intervention by April 2024, or six months before the endpoint for their LA’s delivery period.

As discussed previously, we recommend that evaluators conduct some analyses (e.g. qualitative or descriptive work) earlier so that they can be used in the DfE’s Spending Review bids.

We also propose that evaluators consider collecting homelessness data at a later point, as part of the formative rather than the summative aspect of the evaluation (see the section on “What else do we want to learn? (Formative research)”).

Evaluation questions

Below, we present our recommendations for the other components that complete the design, addressing our evaluation questions, including suggestions for where formative research could be incorporated into the pilot (see “What else do we want to learn? (Formative research)”).

1. Does it work?

A total of 13 possible outcomes for Staying Close were identified and tested as part of the work described in the sections above on “Evaluation questions” and “Evaluation methods”. However, we cannot estimate the effect of the programme on all of these in the impact evaluation, because this would increase the risk of bias as well as the cost of the evaluation. Programme outcomes not selected as primary outcomes in the design may be partially explored using qualitative methods.

To select the most suitable ones to focus on, we considered the following criteria:

- Centrality to the ToC
- Data access and characteristics
- Outcome maturation
- Stakeholder interest
- Suitability for cost analysis.

We recommend two outcomes for the trial:

- Accommodation changes
- EET.

The factors considered in selecting these outcomes are described below.

Accommodation changes

Positive change in young people’s accommodation represents one of the key objectives of Staying Close (section on “Why (What is the purpose of the programme?)”). This focus on accommodation outcomes is strongly supported by the ToC work conducted, which identified the ability to maintain a tenancy in suitable accommodation as the ultimate desired outcome for young people.

To capture this, we recommend using the number of accommodation changes as a proxy measure. Although an accommodation change is not always a negative outcome for young people (they may

be in transitional accommodation, for example), we believe this measure will still provide important insights into the success of the programme taken over the entire sample. Evaluators should also collect data on the reason for accommodation changes and considering whether they are scenarios where we can justifiably exclude these from analysis. Notably, this outcome may not be suitable for the entire sample, specifically 16–17 year olds still living in residential care homes. However, we are confident that is a relatively small proportion of the sample based on collected administrative data in the feasibility study and DfE discussions.

We also suggest evaluators explore the complexities around this measure qualitatively. Measures for ‘suitability of accommodation’ already exist in the LA SSDA903 reporting (annual looked-after children reporting), but the definitions used make them unacceptable for the evaluation, and we concluded that creating a new measure would not be feasible.

We are confident that LAs can provide reliable data on the number and reasons of address changes (section on “Data sources and properties”). Furthermore, we are confident that the properties of the data will allow for analysis (Appendix 13).

EET

We are confident about the data access, quality and distribution of this variable (see “Evaluation methods”), with baseline rates around 45% EET and a favourable distribution contributing to statistical power. The ToC supports the hypothesis that EET outcomes can be positively influenced by Staying Close, even if they are not the primary focus. The programme may enhance uptake and engagement in existing EET activities, due to the enhanced level of support for other life domains, providing an opportunity to test awareness and tailoring mechanisms. While the support offered under Staying Close may not differ significantly from BAU practices, observing an effect would indicate the effectiveness of these specific elements. Moreover, EET is an important outcome for policymakers and allows us also to conduct a cost analysis.

Other variables that were considered as candidates for primary outcomes in the evaluation were social connectedness, homelessness, involvement in the criminal justice system and wellbeing. Social connectedness will be explored in the formative part of the evaluation (see the section below on “What else do we want to learn? (Formative research)” for the rationale), while considerations when we excluded these other outcomes are outlined below:

- **Homelessness:** Although homelessness holds greater significance for policymakers than accommodation changes, and its improvement can be monetised, there is uncertainty regarding the baseline rate of homelessness among the target population. This would affect statistical power and evaluators’ ability to detect an effect of the programme. We are therefore recommending this outcome be explored as part of the formative aspect of the pilot.
- **Involvement in the criminal justice system:** Although there was strong support for this outcome in the qualitative work, the confidence in obtaining accurate data from LAs is

low. Alternative options, such as collecting data from local police, would be likely to impact implications for project timelines, and the quality and distribution of the data is uncertain.

- **Wellbeing:** Despite being considered a higher priority for practitioners and policymakers compared with social connectedness; this received the least support of our candidate outcomes from the ToC work. However, this may be attributed to our limited abilities to measure wellbeing quantitatively rather than its theoretical significance. This variable is believed to have potential as a mechanism supporting other outcomes, such as EET. We recommend testing ways to capture this data as part of the formative aspect of the pilot, given its policy importance and potential DfE plans for universal collection of this data from this cohort in the future.

A matrix detailing the full assessment of candidate variables is in Appendix 8.

Evaluators should conduct qualitative interviews and focus groups with staff and recipients of Staying Close to explore their perspectives on the proposed outcomes, and their broader perspectives on the impact of the programme.

2. Does it work as expected?

The qualitative work highlighted the challenges of identifying universal mechanisms for such a heterogeneous programme where LAs have different focuses and some activities or desired outcomes are prioritised over others. Below, we have included mechanisms that came from the analysis of our data sources and also suggested mechanisms that we believe are theoretically plausible.

We are proposing the following mechanisms for the evaluation.

Direct mechanisms

Awareness, take-up and experience

- The mechanism here is increased awareness of the support offer, which in turn increases take-up of these offers compared with BAU activity in LAs
- This is a universal mechanism for each of the outcomes, but it might be particularly relevant to the EET outcome
- This mechanism was not derived from the data collected during the feasibility study (for example, from practitioner interviews). However, based on our experience of conducting this feasibility study, the research team considered this plausible as a part of the theory of the intervention's effectiveness
- We are recommending that this mechanism be explored primarily through qualitative data collection via interviews and focus groups with young people
- Evaluators should consider survey measures in treatment and control sites that ask young people about their awareness/understanding of their rights and entitlements as care leavers. However, the response rate to a survey is likely to be low and it would be very

difficult to collect survey data from young people in control LAs. Therefore, a statistical analysis is unlikely to be supported and any survey data would be exploratory only, rather than evaluative. We also suggest qualitative work with young people in control LAs to explore this.

Indirect mechanisms

Trusted relationship

- This mechanism emphasises the importance of the key worker understanding young peoples' history and interests, enabling personalised and flexible support, rather than the effectiveness of specific programme activities or training
- This mechanism is theorised to support the social connectedness outcome
- This outcome would potentially allow for a statistical analysis by comparing measures with control young people. However, the response rate to a survey is likely to be low, especially in control sites. Therefore, a comparative analysis is unlikely to be supported and any survey data would be exploratory only, rather than evaluative
- As there will be no pre-intervention data on trusted relationships, we will not be able to match LAs on historic trends in the QED. Therefore, we anticipate that evaluators will not be able to make causal claims for this mechanism, but could collect and compare relationship data as an exploratory analysis. This can be supplemented with qualitative analysis.

Wellbeing

- There is theoretical evidence that this could support other outcomes, such as EET
- Based on the survey data collection in the feasibility study, we anticipate that collecting this data to be very challenging, especially from control LAs, and we would not be confident that evaluators would be able to do so successfully enough to answer evaluative questions on this mechanism
- However, there are plans to add routine collection of wellbeing data to LAs' responsibilities for their care leavers. This makes inclusion of formative research testing how this data could be collected, what scales to use and what distribution the data has more valuable. We suggest that evaluators explore whether the DfE could support and/or whether the evaluation could act as a pilot for their planned activity. This would allow evaluators to collect the data from both treatment and control sites, allowing for a statistical analysis
- As with the other subjective measures, there will be no pre-intervention data on wellbeing, and we will not be able to match LAs on historic trends for the QED. Therefore, we anticipate that evaluators will not be able to make causal claims for this mechanism, but could collect and compare wellbeing data as exploratory analysis. This could be supplemented with qualitative analysis.

Qualitative interviews and focus groups with staff and recipients of Staying Close will be used to explore their understanding of the proposed mechanisms and understand how they expect programme activities to elicit change in the recipients.

3. Does it work differently for some groups? (Subgroup effects)

Although we anticipate that evaluators will not be sufficiently powered to estimate effects of Staying Close on different subgroups, we recommend conducting descriptive analyses as well as qualitative work, such as interviews, on the following (comparing outcomes with the equivalent subgroups in the control LAs):

- **Care history:** this could involve whether young people are placed out of authority, reason for entering care and placement type
- **UASC:** binary
- **SEND:** binary.

Qualitative interviews and focus groups with staff and recipients of Staying Close can be used to explore how different groups of recipients might experience and be impacted by the programme.

4. Does it work differently in some places? (Local effects)

To understand local effects of the programme, we will focus on two areas:

- **Housing market conditions:** Interviews with practitioners and service managers to understand the influence of housing market conditions (e.g. LA access to housing stock in the form of local flats or agreements with housing associations/tenancies)
- **Urbanicity:** Comparing young people in treatment LAs with those in urban LAs within the three urbanicity categories. This will be exploratory only – we won't be powered to make causal claims within these subgroups of LAs.

5. Was it implemented as intended? (Implementation)

We recommend that an implementation process evaluation (IPE) is carried out as part of the pilot study. This comprises mostly qualitative work at the beginning and end of the evaluation period, with both LAs implementing Staying Close and control LAs to act as a comparison where appropriate. Due to the large number of LAs receiving funding in the third round of bidding, in-depth qualitative interviews and focus groups should be conducted with a sub-sample of LAs, and a survey allowing open text responses will be sent to the wider cohort.

We recommend that this work is informed by initial indicators of barriers and facilitators identified in the section on “Evaluation questions” and focuses on the following areas:

- **Programme sustainability:** interviews are conducted with staff to understand the resources required to prepare to and continue to deliver Staying Close
- **Programme feasibility:** interviews are conducted with staff and care leavers to understand how functional the programme is once it has been set up and how well it meets the needs of care leavers
- **Programme acceptability:** interviews are conducted with staff and care leavers to explore the perceived usefulness and compatibility of the programme
- **Programme fidelity:** interviews are conducted with staff and care leavers to understand how well sites were able to adhere to their initial programme plan (based on bids submitted to the DfE) and the core Staying Close objectives defined by the DfE, and what factors impacted fidelity. Evaluators should collect information on what is being delivered in control sites to ensure there is sufficient distinction from Staying Close activities.

Monitoring data can also be used to understand implementation, with administrative data capturing information related to staffing, resources and engagement with programme activities.

We also recommend that evaluators undertake work to typologise Staying Close and create programme models that apply across sites. To do so, they may update the data collection tool to include new LAs based on their submitted bids and use interviews with staff to clarify and refine our understanding of each site's programme. Evaluators can then identify common themes in variations across all sites to inform our typology.

6. Is it a good use of resources? (Value for money)

The economic feasibility study concluded that although a Value for Money analysis of Staying Close is possible, because both recommended outcomes (EEF and accommodation changes) are monetisable, social connectedness (which we recommend including as part of data collection as formative research) can be quantified but not monetised.

The suggested core economic model (cost-offset) therefore includes:

- Cost of Staying Close (for intervention LAs):
 - Accommodation
 - Relationship
 - Bespoke package
- LA expenditure associated with accommodation and (prevention of) homelessness of care leavers:
 - EET (individual level or propensity from LA) + cost from literature.

We recommend that this is compared with the cost associated with BAU/the cost incurred for a comparison or control group.

For all of these items, data can either be provided as totals for the LA so that an average cost saving can be calculated, or as part of an individual-level dataset, the latter being both more useful for

evaluation and more resource-intensive. The confidence that these data can be obtained from LAs at the level of aggregation described above is relatively high based on the available data, perhaps with the exception of the cost associated with BAU, which will need to be collected from a comparison group.

The main risk involved relates to LAs' capacity for data collection. Mitigating this risk requires ensuring sufficient capacity and funding and aligning the economic analysis with the main evaluation. The availability of data for a control group or counterfactual is also an uncertainty, as a final decision on this matter has not been made.

To facilitate economic data collection, it would be beneficial to leverage the plans for rapid cycle testing. This can involve exploring whether data on service use can be directly collected from young people and if routine data collection systems of external and collaborating services (non-LAs) can be aligned with LAs' systems. Implementing key performance indicators (KPIs) as part of service agreements could be a useful tool to achieve this. The study also found that it would be valuable to consider the development of a wellbeing measure that can support economic evaluation in children's social care.

7. What else do we want to learn? (Formative research)

The decision to classify the evaluation as a pilot was influenced by the level of risk around data collection for some of our outcomes of interest. The formative aspects of the pilot provide a valuable opportunity to test various elements involved in conducting trials focused on the care leaver cohort. More specifically, the pilot aims to test different strategies for data collection and measures of **social connectedness and wellbeing**, addressing uncertainties regarding data properties, appropriate measures and scales and effective data collection methods.

We consider social connectedness to be of high importance both to this programme theory and in the wider policy context. The DfE's interest in this outcome has grown following the Care Review, and it is likely that future evaluations of similar programmes will also focus on social connectedness. This (and the wellbeing measure described below) may also be part of the Children's Social Care National Framework and a Data Dashboard, a draft of which has recently been [published for consultation](#).³

The survey response rate in our feasibility study was very low. Therefore, research to identify suitable measures and data collection methods for this outcome will be of huge value to future evaluations.

³ https://consult.education.gov.uk/children2019s-social-care-national-framework/childrens-social-care-national-framework/supporting_documents/Childrens%20Social%20Care%20National%20Framework%20Consultation%20Document%20February%202023.pdf

It should also be noted that social connectedness cannot be included in the DiD analysis due to the lack of data for parallel trends matching. However, it can be used for comparisons with LAs showing parallel trends in other outcomes, albeit exploratorily. In including this as part of the evaluation we are capitalising on the opportunity for formative research, which may increase the value of future work.

As part of the pilot's formative phase, we suggest exploring different approaches for data collection and evaluating the suitability of various scales, including integrating LA-collected data. Below we highlight the formative research questions that we are of interest and the strategies that can be used to answer them:

- What is the most effective way to collect survey data from young people in this cohort?
 - **Convening a data collection working group:** Including representatives from the DfE such as analysts and the Leaving Care policy team, LA project leads from pilot sites, LA data protection officers (DPO), Foundations DPO
 - **Identification of data collection strategies to test:** This will likely include asking key workers to collect social connectedness data directly from young people and integrating this information into LA data returns; asking LAs to send links to young people via text messages or WhatsApp; reimbursing LAs for the data submitted
 - **Implementing these strategies** in different sites and recording response rates; following up with LA staff via interviews to understand barriers/facilitators for effective data collection.

We also recommend evaluators include a formative research question focusing on the programme/cohort:

- What is the baseline rate of homelessness for the cohort according to LA data returns?
What confidence should we have in this data?

Highly valued by policy makers, the focus on homelessness in the formative research process through descriptive analysis will allow evaluators to establish a baseline rate to better understand the extent of harm and inform future trials. However, this data may be collected beyond the timescale of the main evaluation to allow for outcome maturation.

QED evaluation components

The feasibility study did not aim to identify the precise design of a QED study. The research had focused on methods for a potential cluster RCT, before this approach became unfeasible following changes to the policy context in April 2023. As a result, the finer details of the QED are still to be decided. The team is now expediting work to test assumptions related to a DiD design and finalise the impact evaluation plan based on the recommendations provided.

While the full analysis plan for the QED will be confirmed in the protocol stage, we include below details of what we currently know and what uncertainties are still to be addressed before confirming the protocol.

Identification strategy

- **Selection of comparison LAs:** The construction of a suitable control group for the evaluation hinges on identifying enough LAs with comparable historical trends in the outcome measures, adhering to the parallel trend assumption. However, the additional funding provided to 28 LAs in 2023 by the DfE increases the risk of an insufficient number of LAs meeting this criterion. To facilitate the identification process, we require data that is aggregated at the LA level, is publicly available and encompasses at least two periods of pre-intervention data. Data with more time points will increase the reliability of the identification process. When the information is not publicly accessible, we suggest that evaluators explore the use of national government datasets, albeit with potential implications for reporting timelines. First, LAs can be assessed based on their similarity of historic trends to treatment LAs through basic tests. Next, their suitability can be assessed using individual-level data. LAs that meet the assumptions will be contacted, and qualitative work will be conducted to verify their suitability
- **Identification strategy for control young people:** In this case, the challenge lies in identifying the equivalent subset of young people from control LAs who would have received the programme if it were implemented in their LA. For this, evaluators will collect information on the eligibility criteria being used in treatment LAs and applying these to data requests from control LAs.

Outcomes and data sources

Primary outcomes include EET and accommodation changes. We believe that pre-intervention data will be available for these outcomes. However, if historical data on accommodation changes are not available, homelessness could be used as an alternative outcome with extended analysis timelines. Wellbeing and social connectedness can be analysed through pre-post analysis, supplemented by qualitative work. Survey data cannot be used for a QED, but we can include exploratory work comparing young people receiving Staying Close with those matched with them based on historical trends in other outcomes.

Static/dynamic design

Determining whether the DiD design will be static (programme implemented at one time point) or dynamic (staggered programme starts, including LAs from previous phases of the roll-out) is yet to be finalised.

Statistical power

The power of the study depends on the number of treatment LAs included from the first and second waves of the Staying Close roll-out and the identification of parallel trend LAs. This will be detailed in the evaluation protocol.

References

Boddy, J. (2013). *Understanding permanence for looked after children: A review of research for the care inquiry*. *Care Inquiry*.

Clarke, V., Braun, V., & Hayfield, N. (2015). *Thematic analysis. Qualitative psychology: A practical guide to research methods*, 3, 222–248.

Department for Education (2022). *Children's social care innovation programme: Insights and evaluation*. www.gov.uk

Department for Education (2023). *Staying Close 2023–2025: Application guide for local authorities*.

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1132086/Staying_Close_2023_to_2025_-_application_guide_for_local_authorities.pdf

Department for Education, Department for Work and Pensions and Her Majesty's Revenue and Customs (2013). "STAYING PUT" arrangements for care leavers aged 18 and above to stay on with their former foster carers. <https://www.gov.uk/government/publications/staying-put-arrangements-for-care-leavers-aged-18-years-and-above>

Foley, Niamh, R. Harker, D. Foster, W. Wilson, A. Powell, M. Sandford, T. Powell, N. Roberts, J. Lewis, and S. Kennedy. "Support for care leavers." (2022). [CBP-8429.pdf \(parliament.uk\)](#)

Her Majesty's Treasury (2023). *Spring Budget 2023*. [Spring Budget 2023 \(publishing.service.gov.uk\)](#)

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., ... & Michie, S. (2014). *Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide*. *BMJ*, 348.

Marmot, M. (2013). *Fair society, healthy lives*. *Fair society, healthy lives*, 1-74.

Narey, M. (2016). *Residential care in England: report of Sir Martin Narey's independent review of children's residential care*. Department for Education. [Residential-Care-in-England-Sir-Martin-Narey-July-2016.pdf \(publishing.service.gov.uk\)](#)

National Audit Office and Department for Education (2015). *Care leavers' transition to adulthood*. [Care leavers' transition to adulthood \(nao.org.uk\)](#)

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). *Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda*. *Administration and policy in mental health and mental health services research*, 38, 65–76.