Dadansoddi ar gyfer Polisi

YMCHWIL GYMDEITHASOL Y LLYWODRAETH

GSR

GOVERNMENT SOCIAL RESEARCH

Analysis for Policy

Llywodraeth Cymru
Welsh Government

Social research number: 58/2025

Publication date: 10/07/2025

Basic Income for Care Leavers in Wales pilot evaluation: statistical analysis plan

Mae'r ddogfen yma hefyd ar gael yn Gymraeg.

This document is also available in Welsh.

OGL © Crown Copyright Digital ISBN **978-1-80633-022-5**

Basic Income for Care Leavers in Wales pilot evaluation: statistical analysis plan

Authors: Michael Sanders[2], Elizabeth Schroeder[3], Patrick Fahr[3], Dimitris Vallis[2], Vibhor Mathur[1], Louise Roberts[1], Zoe Bezeczky[1], Kate E Pickett[4], Matthew Johnson[5], Rod Hick[6], Stavros Petrou[3], Sally Holland[1] and David Westlake[1]

1. Children's Social Care Research and Development Centre (CASCADE), School of Social Sciences, Cardiff University
2. The Policy Institute, Kings College London
3. Nuffield Department of Primary Care Health Sciences, University of Oxford
4. Department of Health Sciences, University of York
5. Department of Social Work, Education and Community Wellbeing, Northumbria University
6. School of Social Sciences, Cardiff University

Full Research Report: Sanders, M., Schroeder, E., Fahr, P., Vallis, D., Mathur, V., Roberts, L., Bezeczky, Z., Lloyd, H., Pickett, K.E., Johnson, M., Hick, R., Petrou, S., Holland, S. and Westlake, D. (2025). Basic Income for Care Leavers in Wales pilot evaluation: statistical analysis plan. Cardiff: Welsh Government, GSR report number 58/2025.

Available at: https://www.gov.wales/basic-income-care-leavers-wales-pilot-evaluation-statistical-analysis-plan

Views expressed in this report are those of the researchers and not necessarily those of the Welsh Government

For further information please contact:

Jennie Mack

Equality, Poverty and Children's Evidence and Support Division

Welsh Government

Cathays Park

Cardiff

CF10 3NQ

Email: SocialJusticeResearch@gov.wales

Table of contents

List of tables 3

List of figures 4

Glossary 5

Introduction 11

- Overview of the pilot 11
- Overview of the evaluation 13
- Co-production 13
- Theory enhancement 13
- Impact evaluation 14
- Implementation and process evaluation 14
- Economic evaluation 14
- Aims and scope of this report 15

Statistical analysis plan for impact analysis 16

- Quasi-experimental comparisons 16
- Data sources 17
- Survey data 17
- Administrative data 18
- Outcome Measures 20
- Matching methodology: Coarsened Exact Matching 22
- Analytical strategy: Difference in Differences 22
- Five assumptions need to be met to ensure validity: 23
- No Policy Endogeneity 23
- Common Trends 23
- No Spillovers 23
- Uniqueness 23
- Stable Group Composition 24
- Sample selection 24
- Local authority level matching 24
- Post Match Balance 25
- Analytical specification: Survey Data 26
- Primary Analysis 26
- T2 only data 26

Secondary Analysis 27
Imputation 27
Robustness Checks 27
Analytical specification: Administrative Data 27
Imputation 28
Additional cohorts 28
Difference in difference counterfactual 28
Secondary Analysis 29
Robustness checks 29
Statistical analysis plan for economic analysis 30
Introduction 30
Rationale for the economic evaluation 30
Research Objectives 30
Perspective 31
Time horizon and follow-up 31
Measurement of resource use 31
Valuation of resource use data 32
Data Analysis 32
Cost Consequences Analysis 33
Uncertainty and Sensitivity Analyses 33
Conclusions 35
References 36

List of tables

Table 1: Comparisons and data sources 19
Table 2: Outcomes and data sources 20

List of figures

Figure 1: Difference In Differences Example (reproduced from Sanders and others, 2023) 23

Glossary

Basic Income

A basic income is a periodic cash payment unconditionally delivered to all on an individual basis, without means-test or work requirement.

'Better Off' Calculation

Offered pre-pilot and throughout the pilot to eligible recipients to understand whether that individual would be 'better off' being on the pilot or not. This should consider not just financial matters but also wider support that they may/may not be entitled to receive.

Bootstrap

Bootstrap is a statistical resampling method that estimates the sampling distribution of a statistic by repeatedly sampling with replacement from the original dataset.

Citizens Advice Cymru (CA)

A network of independent, locality-based charities that offer free, independent and impartial advice to anyone. The service was initially linked to the development of the social welfare service, and areas of support include: benefits, work, debt and money, consumer rights, housing, family, law and courts, immigration and health. The organisation provided additional financial advice and support for recipients of the Basic Income for Care Leavers in Wales Pilot via the Single Advice Fund. The organisation was formerly called Citizens Advice Bureau and is often referred to as CAB by participants in this study.

Coarsened Exact Matching (CEM)

Coarsened Exact Matching (CEM) is a statistical method that improves causal inference by creating balanced intervention and control groups. This approach ensures matched units are comparable on key characteristics, reducing bias and increasing statistical power compared to other matching methods.

Cost-Consequences Analysis (CCA)

Cost-Consequences Analysis (CCA) is an economic evaluation method that presents costs and outcomes separately without combining them into a single measure of cost-effectiveness. Unlike other economic evaluation approaches that synthesise results into ratios such as cost per quality-adjusted life year, CCA displays multiple outcome measures alongside the costs in a disaggregated format, allowing decision-makers to weigh the different consequences according to their own values and priorities.

Department for Work and Pensions (DWP)

The Department for Work and Pensions (DWP) is the UK government department responsible for welfare, pensions, and employment support services. DWP maintains administrative data on benefit claimants, employment outcomes etc.

Difference-in-Differences Analysis

Difference-in-Differences (DiD) analysis is a quasi-experimental research design that estimates causal effects by comparing changes in outcomes over time between intervention and control groups. The method relies on the parallel trends assumption, requiring that intervention and control groups would have followed similar trajectories in the absence of the intervention. DiD is useful for evaluating policy interventions when randomized trials are not feasible.

GAD-2 (Generalized Anxiety Disorder-2)

The GAD-2 is a brief anxiety screening instrument comprising the first two items from the GAD-7 scale, measuring feelings of nervousness/anxiety and uncontrollable worry over the past two weeks. Items are scored on the same 4-point scale as the PHQ-2, with scores of 3 or above suggesting possible anxiety disorders warranting further evaluation.

Heads of Children's Services

Heads of Children's Services are senior leaders responsible for children's services in local authorities. They typically set the strategic direction of the service, provide leadership, oversight, guidance and support for other senior managers.

His Majesty’s Revenue and Customs (HMRC)

HMRC is the UK’s tax, payments and customs authority. It is the UK government department responsible for handling taxes, regulating National Insurance, among other financial functions.

Income Deprivation Affecting Children Index (IDACI) Score

The Income Deprivation Affecting Children Index (IDACI) score measures the proportion of children aged 0-15 living in income-deprived families within a specific area, forming part of the English Indices of Deprivation. Higher IDACI scores indicate greater levels of income deprivation among children in that area, making it a key indicator for targeting resources and support services.

Income Deprivation Affecting Older People Index (IDAOPI)

The Income Deprivation Affecting Older People Index (IDAOPI) measures the proportion of older people (aged 60 and over) experiencing income deprivation within a specific area, also forming part of the English Indices of Deprivation.

Intention-to-Treat (ITT) Principle

The intention-to-treat (ITT) principle is an analytical approach in research where participants are analysed according to their original group assignment, regardless of whether they actually received the intended intervention or completed the study.

Lifelong Learning Wales Record (LLWR)

The Lifelong Learning Wales Record (LLWR) is the comprehensive administrative database that collects information on all post-16 learning in Wales, covering further education colleges, work-based learning providers, adult community learning, and higher education institutions.

Local Authority

There are 22 local authorities (councils) in Wales which make up the elected local government tier for Wales. They are typically responsible for delivering children's social care services.

Longitudinal Educational Outcomes (LEO)

The Longitudinal Educational Outcomes (LEO) dataset is a comprehensive administrative database created by the UK Department for Education that links individuals' education records with their employment, benefits, and earnings data. LEO connects de-identified person-level information from various government databases to create a longitudinal picture of outcomes over time.

Member of Senedd (MS)

A Member of Senedd (MS) is an elected representative who serves in the Senedd Cymru (Welsh Parliament). The term 'Member of Senedd' replaced 'Assembly Member' (AM) following the Senedd and Elections (Wales) Act 2020, which renamed the National Assembly for Wales to Senedd Cymru.

Lower Layer Super Output Area (LSOA) Level

Lower Layer Super Output Areas (LSOAs) are geographical units used for statistical reporting in England and Wales, designed to have relatively similar population sizes of approximately 1,500 residents on average. LSOAs are built from clusters of Output Areas and provide a consistent framework for analysing and comparing socio-economic data across different areas.

Missing At Random (MAR)

Missing At Random (MAR) is a statistical assumption about missing data where the probability that a data point is missing depends only on observed variables, not on the unobserved missing values themselves.

Missing Completely At Random (MCAR)

Missing Completely At Random (MCAR) is a statistical assumption about missing data where the probability that a data point is missing is completely independent of both observed and unobserved data values.

Multiple Imputation by Chained Equations (MICE)

Multiple Imputation by Chained Equations (MICE) is a statistical technique for handling missing data that creates multiple complete datasets by iteratively imputing missing values based on observed data patterns.

National Health Service (NHS)

Government funded medical and healthcare services that everyone in the UK can use without being asked to pay for the full cost of the service. The NHS was established as part of major social reforms following the Second World War. The founding principles were that services should be comprehensive, universal, and free at the point of delivery. There is more than one NHS in the UK, comprising of NHS England, NHS Wales and NHS Scotland.

National Pupil Database (NPD) for England

The National Pupil Database (NPD) for England is a comprehensive administrative dataset maintained by the Department for Education that contains longitudinal information on all pupils in state-funded schools in England.

OSF (Open Science Framework)

The Open Science Framework (OSF) is a web platform designed to support research collaboration and promote transparency and reproducibility in scientific research. It is developed and maintained by the Center for Open Science.

Personal Advisor / Young Person's Advisor (often referred to as a PA / YPA)

Practitioner working directly to support care-experienced young people. A Personal Advisor is responsible for working with young people who have left the care of the local authority (and is often in contact with them before this in order to transition as smoothly as possible from the provision of a social worker). Personal Advisors also contribute to pathway plans and to making sure they are carried out correctly.

Personal Social Services Resource Unit (PSSRU)

The Personal Social Services Resource Unit (PSSRU) is a leading research centre that produces annual estimates of the unit costs of health and social care services in the UK.

PHQ-2 (Patient Health Questionnaire-2)

The PHQ-2 is a brief depression screening tool consisting of the first two questions from the longer PHQ-9 questionnaire, asking about depressed mood and anhedonia (loss of interest or pleasure) over the past two weeks. The PHQ-2 is widely used in primary care and research settings as an efficient first-step screening tool for major depressive disorder.

Post-16 Pupil Collection

The Post-16 Pupil Collection is an administrative dataset that captures information about young people aged 16-18 in further education colleges, sixth forms, and other post-16 education providers in England.

Pupil Level Annual School Census (PLASC)

The Pupil Level Annual School Census (PLASC) includes individual pupil-level data from all state-funded schools in England and Wales on an annual basis.

Quasi-experimental design

A research design that aims to establish causal relationships by comparing groups or conditions but lacks the random assignment of participants that characterizes true experimental designs. In quasi-experimental studies, researchers work with naturally occurring groups or use non-random assignment methods to allocate participants to different conditions.

Real Living Wage

An independently calculated rate of income based on what people need for a decent standard of life. Annually calculated by the Resolution Foundation and overseen by the Living Wage Commission, the pilot's payment amount is based on the rate as of November 2021 for the equivalent of a 37-hour working week.

Realist Evaluation

An approach to evaluation which is underpinned by the premise that any intervention designed to solve a social problem relies on human decision-making as well other factors to make them work. This means that the same intervention implemented in a different context may work (or not work) through different mechanisms and therefore produces different outcomes. Rather than determining the 'average effect' of a policy, realist evaluations try to explain what works, for whom and in what circumstances.

Social Cost-Benefit Analysis (SCBA)

Social Cost-Benefit Analysis (SCBA) is an economic evaluation method used to assess the overall impact of a policy, program, or intervention on society as a whole. Unlike traditional cost-benefit analysis that focuses primarily on financial costs and benefits to the implementing organization, SCBA takes a broader societal perspective by considering all costs and benefits that accrue to different groups in society, including individuals, families, communities, and government.

St David's Day Fund

A fund set up by the Welsh Government in 2017 to support young people who are or have been in local authority care. It is part of the Children and Communities Grant. It is administered by local authorities and they are encouraged to be flexible and creative in deciding how to use it to best meet the needs of young people in the same way that birth parents might financially support their children.

Type I Errors

Type I errors occur when a statistical test incorrectly rejects a true null hypothesis, essentially concluding that an effect or difference exists when it actually does not. Also known as a "false positive" result.

Welsh Index of Multiple Deprivation (WIMD)

The Welsh Government's official measure of relative deprivation for small areas in Wales. It identifies areas with the highest concentrations of several different types of deprivation by ranking all 1,909 Lower-layer Super Output Areas (LSOAs) in Wales from 1 (most deprived) to 1,909 (least deprived).

Introduction

The Basic Income for Care Leavers in Wales pilot (herein referred to as 'the pilot') was launched in July 2022 with the first payments issued to recipients in August 2022 (Welsh Government, 2022a). The evaluation of the pilot began in November 2022 and two annual reports and a study protocol have been published to date (Holland and others, 2024; Mathur and others, 2025, and Westlake and others, 2024).

This report is a technical document that details the specific statistical approaches we intend to use to analyse the quantitative data sources. This analysis will test whether the pilot is effective in improving outcomes for participants and cost effective from the perspective of government and wider society. These quantitative estimates of effects will be complemented by a wide range of qualitative insights from interviews and focus groups with key stakeholders, and fieldnotes from the implementation and process component of the study. Some of this has already been published in the annual evaluation reports (2023 to 2024 and 2024 to 2025). This report is designed to complement the study protocol (Westlake and others, 2024) as a guide to the study and the methods used. As we describe in more detail below, this is intended to increase the transparency and reproducibility of the evaluation and conform with best practice guidelines for evaluation as set out in the Magenta Book (HM Treasury, 2020).

Overview of the pilot

The pilot is unique in several ways, and as such it has garnered a great deal of interest from both within the UK and around the world. No other basic income scheme has given regular payments as large as those received by young people involved in the pilot, nor have other basic income schemes been open to a national cohort of care leavers from the age of 18 for two years of payments. Welsh Government has further information about the nature of the pilot but we briefly describe the key aspects here.

Before we discuss the practicalities of who is eligible and how the basic income payments are delivered, it is worth noting that the pilot is based on four key principles. These were described by the Minister responsible for the pilot, Jane Hutt MS in her 2022 Written Statement (Welsh Government, 2022b):

- Taking part in the pilot should make no recipient worse off.
- There should be no conditionality on income received.
- The same payment should be paid to everyone.
- The payment will not be altered midway through the pilot.

To be eligible, a young person must be a category 3 care leaver. This means they must have been looked after by one of the 22 local authorities in Wales for a period of at least 13 weeks (or multiple periods amounting to at least 13 weeks), beginning after their 14th birthday and ending after they reached the age of 16 years. At enrolment, they must have been resident in Wales or being supported as a care leaver by a Welsh local authority's social services department and living elsewhere.

For those meeting these criteria, the pilot had an enrolment period lasting 12 months, from 1 July 2022 to 30 June 2023 for all those having an 18th birthday in that period. For those who enrolled, the payment started the month after their 18th birthday and continued for 24 months, with the exception of some who enrolled late for reasons outside their control, and who therefore started receiving payments shortly after this point. Those that spent time in custody during the pilot have not received 24 months' worth of payments, as in these circumstances payments are paused until they are released from custody. The total amount of the basic income payment is £1,600 gross per month, which is taxed at source to leave recipients receiving a net amount of £1,280 each month. Some recipients will be eligible for a tax rebate depending on other taxable income, such as earnings from employment. Recipients of the basic income can choose whether this is paid in one monthly transfer of the full amount, or twice-monthly transfers of half the amount. They can also choose to have housing costs paid from the basic income directly to their landlords. When the policy was developed, in 2021 to 2022, the levels of payment were set to be roughly in line with the Real Living Wage (Living Wage Wales, 2023) for a full-time employee at that time. These payments are unconditional, in keeping with the second principle mentioned above, and with the notion that recipients should be free to make their own life choices without being compelled to take a certain route.

In addition to the basic income payments each month, the young people involved are also eligible for support around budgeting and managing finances as part of the pilot. In most local authorities this is provided by Citizens Advice Cymru, but some local authorities are providing financial advice via other means. This component of the pilot is designed to offer tailored financial advice and support for individual young people, including 'signposting and additional support relating to wellbeing, education, work, as well as broader financial advice' (Welsh Government, 2022a). As part of the decision to enrol onto the pilot, an optional 'Better Off' calculation was available to inform this decision and estimate whether or not the basic income would constitute a net increase in a person's income.

No aspects of the pilot are intended to affect a care leaver's eligibility for their usual services. This is in keeping with the first principle mentioned above and reflects a desire to add to the support available for care leavers rather than replace existing provision. As care leavers, they are eligible for a range of services, including support from an allocated 'Young Person's Advisor', who provides one-to-one advice and support in various aspects of life. They are also exempt from council tax liability, eligible for higher education bursaries and cost of living payments and can apply for financial grants through the St David's Day Fund, which is administered by local authorities alongside localised forms of support.

The pilot has proved popular, with 97% of young people who were eligible taking part (n= 644; Welsh Government, 2024).

The pilot has a governance structure which includes a Steering Group, an Operational Group, a Research and Evaluation Advisory Group and an independent Technical Advisory Group. These groups include civil servants and independent experts, and their function is to give advice to Welsh Government on the design and delivery of the pilot. In addition, the evaluation has a separate advisory group, which consists of experts from various fields of academic study and professional practice.

Overview of the evaluation

The commissioned evaluation began in November 2022 and is due to end in 2027. Several aspects of the pilot are within the scope of the evaluation, and the overall evaluation uses a range of quantitative and qualitative methods. We will measure its impact in the different areas of recipients' lives that we detail below. We will also explore how the pilot is implemented, the attitudes and experiences of key stakeholders involved, and the cost effectiveness of the pilot from public sector and societal perspectives. The following research questions cover these disparate aspects of the scheme:

- Research question 1: What is the impact of the pilot?
- Research question 2: Is the pilot implemented as intended?
- Research question 3: How is the pilot experienced?
- Research question 4: How does the pilot fit into the overall offer for care leavers in Wales?
- Research question 5: How cost effective is the pilot?

The published study protocol (Westlake and others, 2024) describes our approach to answering these questions in greater detail. It also includes a more extensive summary of the background to the pilot, a discussion of previous research on the topic, and the challenges and limitations we face in conducting this study.

The evaluation is designed around five core areas called 'work packages'. These are briefly summarised as follows:

Co-production

Co-production underpins the study and participatory methods feed into the design and data collection. A group of care-experienced young adults, living in a range of educational, employment and housing situations, meet regularly to provide advice. Their role is to co-create research questions, data collection instruments, consider ethical and analytical questions and advise on policy and practice implications. Seven such meetings have taken place to date, involving between six and ten young people in each meeting, and the advice of the group has materially informed key decisions. Examples of changes and decisions made in conjunction with this group include which survey measures to choose, focus group questions for young people's advisors and a change to procedure to allow young people to take part in qualitative interviews in pairs or small groups, rather than just individually, if they choose to do so.

Theory enhancement

The study is a theory-based evaluation, and it employs techniques informed by realist evaluation (Chen, 2014; Pawson, 2013; Pawson and Tilley, 1997). The advantage of this approach is that it promises an insight into how and why the pilot may or may not have the intended effects for different people involved.

Impact evaluation

The impact of the pilot is being measured on several outcomes, so that we can see in what ways and how much of an effect it has for young people involved. The outcomes of interest for the pilot, which were specified at the outset by Welsh Government, fall into the following categories:

(1) wellbeing

(2) financial literacy / security

(3) community cohesion / engagement

(4) the effects of poverty

(5) access to labour and education markets

(6) volunteering and life skills

(7) physical and mental health.

Implementation and process evaluation

This strand focuses on how the pilot is implemented, its ongoing delivery, and how the pilot is experienced and perceived by those involved. Also explored, within this analysis, is how the pilot fits into the overall offer for care leavers in Wales, including intersections with existing services. The qualitative elements of this work package were enhanced and extended in 2023, when further funding became available. This increases the scope of the study to understand the lived experience of young people involved.

Economic evaluation

The economic evaluation will consider whether the pilot represents value for money in terms of the outcomes achieved. It aims to synthesise the costs and consequences of the pilot, to inform a social cost-benefit analysis. Social cost-benefit analysis (SCBA) is an extension of cost-benefit analysis, adjusted to consider the full spectrum of costs and benefits (including social and environmental effects) borne by society as a whole because of an intervention or programme. A further cost-consequences analysis (CCA) will examine a range of key outcomes to explore wellbeing where the data allow, and aspects of educational attainment, engagement in the labour market and financial security.

The evaluation is informed by survey research. Coram Voice, a children's rights charity, has been commissioned separately by Welsh Government to gather survey data from participants of the pilot and a comparator group (more details in Section 2).

All research undertaken by Cardiff University must undergo ethical review. This study was considered by Cardiff University's School of Social Sciences Research Ethics Committee and approved in January 2023 under the reference SREC/323. An amendment, to include additional qualitative methods, was approved on 20th June 2023. Prior to this, ethical approval for Coram Voice's administration of the evaluation surveys was obtained from the

University of Oxford under the reference CIA-22TT-149. The study is also being conducted in line with Government Social Research ethical principles.

Aims and scope of this report

Modern evaluations are increasingly publishing their analysis plans in detail (Ofosu and Posner, 2023). This typically happens prior to data collection being completed and analysis beginning. The main advantages of publishing detailed analysis plans are transparency, replicability and credibility.

Three concepts underpin what are now considered to be cornerstones of transparency in social science: (1) data access (2) 'production' transparency, and (3) analytic transparency. Elman and others (2018) explain these as follows:

> 'Data access refers to making available to others the data on which empirical claims in published research rest; production transparency implies clearly explicating the most relevant aspects of the data generation process; and analytic transparency entails conveying the processes through which data were analyzed [sic] to produce claims and conclusions' (page 32).

Welsh Government intend to make data available (where possible) after the evaluation has been completed. The study protocol (Westlake and others, 2024) gives an account of the overall study design and an overview of the analysis plans to enhance production transparency. This technical report furthers the analytic transparency of the evaluation by providing further detail on how analysis of the quantitative data sources will be conducted.

This is important because the profile of the Basic Income for Care Leavers in Wales pilot makes replication efforts likely. The pilot has generated a large amount of interest to date, in the UK and internationally, and this is likely to increase as it nears its conclusion. Being the most generous basic income in the world to date means other researchers are likely to want to replicate methods and ratify findings. In psychology and behavioural science there is a well-known replication crisis, whereby researchers have been unable to replicate research findings to check they are valid (Open Science Collaboration, 2015; Ioannidis, 2005). This is driven (in part) by not pre-specifying analysis plans for evaluations and making those public. Organisations such as the Open Science Framework (OSF) have been set up to tackle this and give guidelines for researchers. For example, OSF says:

> 'Reporting research design and analysis should maximize transparency about the research process and minimize potential for vague or incomplete reporting of the methodology.' (Center for Open Science, 2015)

They go on to suggest that the public should have access to full analysis plans as well as study protocols (i.e. analytic as well as production transparency).

Similar recommendations exist in specific disciplines. For example, within Health Economics, researchers have developed an economic evaluation best practice checklist which suggests that researchers should "indicate whether a health economic analysis plan was developed and where available" (Husereau and others, 2022).

Statistical analysis plan for impact analysis

Quasi-experimental comparisons

As described in the study protocol (Westlake and others, 2024), we will use a quasi-experimental design to analyse the impact of the pilot. Quasi-experimental evaluations attempt, in the absence of randomisation, to achieve identification of the causal impact of one or more interventions, primarily through a mix of sample selection and statistical approaches.

These approaches attempt to solve what is known as the “missing counterfactual problem”. That is, we can observe the outcomes of individuals who receive an intervention (intervention group), and we can observe the outcomes of individuals who did not receive an intervention (comparator group), but not both. As a result, any outcome which we observe for someone who receives the intervention cannot be attributed to the effect of the intervention - it might have occurred anyway. To be able to identify a causal impact, we must therefore solve this missing counterfactual problem. In the case of a randomised controlled trial, this process is relatively straightforward. Randomly assigning participants to a treatment or a control group ensures that in the absence of the intervention, both groups would have the same expected average values for their outcomes. Therefore, any difference between the two groups can be attributed to the causal effect of the intervention.

In this instance, and in many other evaluations of public policies, randomisation is not an option. Instead, we must find an alternative in which we are able to make the same claim; that, conditional on our sample selection and analytical strategy, we do not expect to see any uncontrolled for differences between the intervention and comparator groups, and so any differences between the two groups can be attributed to the intervention.

This is not so straightforward. We cannot, for example, simply compare young people’s outcomes at the start of the intervention to those at the end - their experiences of these two years are likely to shape their outcomes in many ways that do not relate to the pilot. Moreover, we cannot compare outcomes for this group to those of young people who left care in the past - their most immediate prior cohort has experienced different labour market conditions due to both the effects of the Covid-19 pandemic and recent changes in the cost of living.

Our solution is to make a number of different comparisons, each one comparing the cohort of care leavers who have experienced the pilot with a cohort of care leavers who have not. These comparisons will use two main types of data, which cover different types of outcomes:

- Self-reported data from surveys
- Administrative data from routinely compiled databases

The principal comparison (‘A1’ in Table 1 below) will be of self-reported survey outcomes of the group of young people who received the basic income (intervention group) with a group of care leavers in Wales who turned 18 the following year (the comparator group), making

them ineligible for the pilot. To increase sample size, we intend to also supplement the comparator group self-reported survey outcomes with additional survey data supplied by Coram Voice, which includes care leavers living in England ('A2' in Table 1). The second comparison ('B' in Table 1) will be of routinely collected outcome data that forms part of the Longitudinal Educational Outcomes (LEO) datasets, available for both England and Wales. In this comparison, the intervention group will be compared with care leavers in England of the same age and similar characteristics, who were ineligible for the pilot because it was only available in Wales. This will involve two quasi-experimental approaches: coarsened exact matching (CEM) and difference in differences (DID), as described below.

We detail these comparisons in Table 1 below, after specifying the data sources we intend to use.

Data sources

We will use two main datasets; a survey administered by Coram Voice, and administrative data extracted from the Longitudinal Educational Outcomes Datasets (LEO).

Survey data

We are contracted to use data collected by Coram Voice, a children's rights charity commissioned by Welsh Government to gather survey data from participants of the basic income pilot. Surveys will be administered by Coram Voice for the intervention group and comparator group at two time points (Time 1 (T1); around the participants' 18th birthday, to coincide with entry, Time 2 (T2) around the participants' 20th birthday, to coincide with exit). It should be noted that both of these time periods differ for each participant in the study, because they turn 18 at different points during the enrolment window. Therefore, data collection will, in practice, take place continuously over the study period.

The survey is based on an established Coram Voice survey used with care leavers in across the UK, called 'Your Life Beyond Care' (Briheim-Crookall and others, 2020). Coram Voice started collecting data in Wales in October 2022, four months after the pilot had started, and initially included only the original questions used in the 'Your Life Beyond Care' survey. In January 2023 an updated survey was released with additional questions designed to cover the broader range of outcomes in Table 1. The new questions added were the result of discussions between the evaluation team, Welsh Government, and Coram Voice.

After several changes to the mechanism for collecting survey data were agreed, and questions added, response rates increased, and the final response rate was 64% for the intervention group at Time 1. Welsh Government and Coram Voice take informed consent for survey data.

It should be noted that in the absence of enough quality matches from the survey data, we may also include data from other relevant studies like the Additional Financial Assistance for Care Leavers trial (Sanders and Vallis, 2023) to maximise comparability of the control group.

Administrative data

We intend to use the Longitudinal Educational Outcomes (LEO) datasets which are held by the Office for National Statistics (ONS) and Welsh Government. This resource links: educational data from the National Pupil Database (NPD) for England, the Pupil Level Annual School Census (PLASC), the post-16 pupil collection and the Lifelong Learning Wales Record for Wales, employment and earnings data from HM Revenue and Customs (HMRC) and the Department for Work and Pensions (DWP), progression and success in further education from the individualised Learner Record (Department for Education) and progression to higher education from the Higher Education Statistics Agency. English LEO also contains markers for young people's social care experience, which will allow us to identify care leavers in England, while the intervention group will be identified within Welsh LEO using matched demographic information of participants. As detailed below, we will initially consider all available data on care leavers from England and Wales, and will use a two stage matching process, to select a group of English care leavers who are (a) in English local authorities that are comparable to the 22 Welsh local authorities and (b) who are comparable to the Welsh care leavers within those local authorities.

Use of administrative data will allow us to link the intervention group with a large enough comparable group of young people experienced in care during the same period in which the intervention took place from England. This will provide the analysis with higher power, increasing precision while allowing for a broader range of outcomes to be explored. It will also enhance causal inference with respect to unobserved time-dependent covariates that may have been correlated with the outcomes of interest at the time of the intervention, an aspect of analysis which is not possible through survey-based analysis.

Outcomes within administrative data will be aligned with various outcomes already explored within the survey (see Table 2), thus allowing for accurate inference on the effect of the intervention. For outcomes not covered by the survey, a binary interaction indicator of whether the participant was in a local authority that would make them eligible for the basic income payments at the time of the intervention will be used to capture the intervention effect.

Other administrative data is also available from Welsh Government. They managed the enrolment of eligible young people in collaboration with the 22 local authorities. Each recipient completed an enrolment form at the start of the pilot. Enrolment forms contain a range of monitoring data, including some self-reported data about the individual's health circumstances. An exit form is also used by Welsh Government to gather data at the end of the pilot. Welsh Government developed a database to record all this information, and this management data has formed the basis of Welsh Government's data releases during the pilot (Welsh Government, 2024).

Some data items gathered at the enrolment stage for participants in the intervention group are added to surveys completed by participants in the comparison group (as these data would otherwise be missing for this group). Welsh Government take informed consent for monitoring data to be shared with the evaluation team.

Table 1: Comparisons and data sources

Comparison label	Intervention	Comparator	Data	Matching
A1	Care leavers in Wales who enrolled on the pilot and completed surveys	Care leavers in Wales who turned 18 after the pilot enrolment period ended (July 2023 – September 2024*) and completed surveys	Self-report survey data	None
A2	Care leavers in Wales who enrolled on the pilot and completed surveys	Care leavers included above in comparison A1, plus care leavers in England who completed a similar survey administered by Coram Voice between 2021-2024.	Self-report survey data	CEM
B	Care leavers in Wales who enrolled on the pilot for whom administrative (LEO) data is available	Care leavers in England who are the same age as the treatment cohort, for whom administrative (LEO) data is available	Administrative data routinely collected and compiled within LEO	CEM

*The window for this group is intentionally longer than 12 months to increase sample size.

Outcome Measures

As detailed in the study protocol (Westlake and others, 2024), Welsh Government identified six outcome domains of interest, and the literature suggests that it is also important to include physical and mental health outcomes more broadly. These outcome domains of interest, and the means of collection through surveys and administrative records are outlined in Table 2.

Table 2: Outcomes and data sources

Outcome domain	**Collected through**	**Specific measures**
Wellbeing/ psychological wellbeing	Survey data	Categorical indicators of frequency of emotional states, level of anxiety, happiness, feelings of positivity, including an open-ended question
Financial literacy/ security	Survey data	Questions on levels of financial coping, allocation of income, spending habits with an open-ended question on money management.
Community cohesion/ engagement	Survey data	Binary indicators of friendship, partnership, owning a pet, having a person of trust, and categorical indicators of emotional support. Binary indicators of community cohesion in follow up (T2) survey
Ameliorating the effects of poverty	Survey and administrative data	Categorical indicator of current housing, binary indicator of housing satisfaction Leisure and access to luxury items/internet

Access to labour market/ education/ lifelong learning	Administrative data	Binary indicator of employment; continuous measure of earnings; binary indicator of enrolment in education or training; binary indicator of participation in higher education; binary indicator of participation in further education. These outcomes will be extended using linked longer-term data from the Longitudinal Education Outcomes (LEO) data.
Volunteering and life skills	Survey data	Education, employment and training.
Physical and mental health	Survey and administrative data	Self-rated general health (ONS question) Limiting long-term illness (ONS question) Common mental disorders: depression (PHQ-2) (Kroenke and others, 2003); anxiety (GAD-2) (Spitzer and others, 2006).

Welsh Government has not specified the exact effect they expect the pilot to have on these outcomes, although they expect it to have positive effects across all these outcomes. They aim for the policy to empower participants, give them more agency and control, and improve their lives. The literature on basic income suggests that we should not see any detrimental effects in any of these areas, and that in many areas improvements would be hypothesised. Some of these improvements may take longer than others to materialise, meaning that some benefits may not be detectable during the timescale of the study. In our final report we will publish analysis of the observed effects on all outcomes in Table 2. The measures cited were selected by the evaluation team and approved by Welsh Government.

Now that we have specified the overall approach including the specific comparisons, the data sources, and the outcomes of interest, we will further detail the approaches to matching and the quasi-experimental analysis.

Matching methodology: Coarsened Exact Matching

For comparisons A2 and B a suitable comparator group needs to be identified from the larger group of care leavers in England for whom data is available. This requires a process called matching, where individuals are matched based on certain characteristics in order to create a comparator group which is as similar as possible to the intervention group, except that they did not receive the intervention.

The methodology we intend to use for this is called Coarsened Exact Matching (CEM) (Iacus and others, 2012). This approach lies somewhere between the two extreme forms of matching - the completely Uncoarsened Exact Matching, or the logical extreme of coarsening to a single figure - the Propensity Score (Rosenbaum and Rubin, 1983). In CEM, matching variables are preserved but are coarsened. Coarsening means redefining variables into ranges. For example, instead of a participant's height being an exact number of centimetres, which might be difficult to find a match for in small samples, this could be coarsened to heights in ten centimetres intervals. By doing this, participants continue to be matched on the values of their observable characteristics, but the likelihood of matching on any variable or set of variables is increased. This matching approach has the advantage of yielding more matches than exact matching, while also ensuring that units are matched on measures that are relevant to the outcomes of interest.

For the administrative data, we will use matching at two levels - first to match Welsh local authorities with their English counterparts, and second to match care leavers within those local authorities with each other. For the survey data, we will use matching at the level of care leavers.

Analytical strategy: Difference in Differences

Difference in differences approaches are quasi-experimental approaches which compare the differences between individuals in the intervention and comparator groups, at the start and end of a new policy or intervention (Figure 1 below). This comparison allows for time invariant differences (differences that stay the same over time), whether observed or unobserved, to be controlled for analytically. Although the comparison of two groups that have not been measured contemporaneously (which is the case with comparison A1) is a non-standard implementation of the difference in differences methodology, the underlying assumptions remain the same.

Figure 1: Difference In Differences Example (reproduced from Sanders and others, 2023)

Five assumptions need to be met to ensure validity:

1. No Policy Endogeneity
2. Common Trends
3. No Spillovers
4. Uniqueness
5. Stable Group Composition

No Policy Endogeneity

The decision to implement a particular policy in a particular place at a particular time cannot have been based on anything that is endogenous to the intervention effect. Put differently, policymakers who chose the intervention to roll out, need not to have done so specifically because the treatment effect of the intervention was higher in that local authority than in others. Given that we will match on prior levels and trends in relevant outcomes and characteristics, and the requirement for being treated is “being in Wales”, we argue that this condition is met.

Common Trends

Difference in difference analyses allow for individuals to be different in their baseline levels of the outcome measure but require that outcomes within the intervention group and comparator group have a tendency to move in a way which is correlated. In our analysis, we will seek to both verify and ensure that this assumption is met through inspection of trends of the outcomes prior to intervention roll-out, and by matching on prior trends in the outcome measure.

No Spillovers

This assumption requires that, for example, an increase in employment or wellbeing in one local authority which received the intervention does not lead to a similar increase, in another comparator local authority which did not receive the intervention. Given the temporal gap between the comparator local authorities and the intervention authorities in Wales, and the spatial distinction between local authorities in Wales and England, we think that this assumption is likely to be met.

Uniqueness

This assumption requires that the only major change in the local authorities involved in the pilot that took place at the time of the pilot’s rollout was the pilot itself (i.e. no similar

interventions were also implemented). Given the cost intensiveness of the basic income intervention, that it was rolled out through the nationally administered programme, this assumption is justifiable.

Stable Group Composition

This assumption requires that participating individuals are unable to move from a treated unit to an untreated unit, or vice versa, over the time period covered by the evaluation. Given that comparator group outcomes for self-report survey data will be estimated the following year, and that no additional major financial assistance policies are expected to be implemented at that time, our analysis will meet the assumption of stable group composition. The same is true for the administrative data, because no major financial assistance policies were implemented in England at the time of the pilot.

Sample selection

For comparison B, using administrative (LEO) data, we will use a matched difference in difference analysis, in which first local authorities, and subsequently where possible individual care leavers are matched statistically on their observable characteristics, and then comparisons over time of different care leavers within the same authorities are used to control statistically for time invariant characteristics of the local authority.

As the pilot is universal for care leavers meeting eligibility criteria across Wales' 22 local authorities, for the survey analysis we must make use of matching only at individual level between Welsh care leavers at T1 and their following year's peers at T2. We do not offer matching details at this point for the survey, as for this to be achieved, we must wait to obtain comparator data at T2.

In the evaluation specification, Welsh Government advised that around 550 young people were expected to become eligible for the intervention during the enrolment period (based on local authority estimates), so our calculations are based on this estimate (even though more young people were actually eligible and enrolled). Although ex ante power calculations of complex quasi-experimental designs are difficult to accurately compute, we anticipate being able to detect effects on self-report survey data outcomes of no more than 0.2 standard deviations (calculated via Glass's Delta), and for effects of no more than 0.12 standard deviations for the administrative data, based on our experience with other similar projects. These effect sizes are comfortably within the range of small effects, allowing us to build a clear picture of the impacts of basic income. However, it should be noted that the small sample size makes subgroup analysis, particularly for any group which is in a minority among eligible participants, difficult to conduct reliably.

Local authority level matching

As described above, we use CEM and at this stage we have done an initial exercise in matching local authorities for comparison B. This required publicly available information on the age, legal basis, gender, and numbers of care leavers in local authorities in Wales from 2018 to 2022, with a panel at local authority/year level. This panel was then used to identify

variables capturing the rate of change in the children in care within each local authority in Wales. We further accessed data on indices of multiple deprivation for Wales (WIMD – Welsh Government, 2019), particularly focusing on the indices for employment, income and childhood deprivation at Lower Layer Super Output Area level, which is subsequently collapsed to give a local authority level average for each score for all top tier local authorities (these local authorities serve as top-tier units responsible for children's services in England, equivalent to local authorities in Wales). This is in turn matched into the panel dataset created previously.

Using this data, we implemented CEM iteratively to arrive at a comparator sample. This involves using the standard CEM command in Stata version 16 (StataCorp, 2019), and coarsening follows the Scott Algorithm. This algorithm is more conservative in its coarsening than some alternatives, but has the benefit of reducing researcher degrees of freedom, reducing the risk of false positives and making replication easier.

For comparison B, our initial attempt at matching included variables across Welsh and English local authorities for; annual rates of change in care numbers in the local authority for each of the last four years; income deprivation scores, employment deprivation scores, and deprivation experienced by children scores (Ministry of Housing, Communities & Local Government, 2019). However, this set of variables, given the conservatism of the matching approach, yields relatively few matches (six Welsh local authorities find matches). Since this is too few, we iteratively reduced the number of variables used in matching to determine which combination of variables, coarsened in this way, finds a match for larger proportions of the Welsh local authorities without reducing the quality of the match.

When matching using Income Scores, Income Deprivation Affecting Children Index (IDACI) score and Income Deprivation Affecting Older People Index (IDAOPI) score, we identify 41 matches for 21 Welsh local authorities. Using Employment Score instead of IDAOPI score produces 29 matches for 18 Welsh local authorities. Reducing this to only Income scores and IDACI Scores, yields 71 matches for all 22 Welsh local authorities. Given the need for some specificity of matches (more than half the local authorities available for matching are matched in the second model) the first or second approaches, which identify matches for 21/18 Welsh local authorities, are preferred at this stage.

Post Match Balance

In order to achieve matches we have had to reduce the number of variables included in the match. This is in part a consequence of several factors:

- there are few local authorities in both groups
- the degree of aggregation makes matches more challenging to achieve
- the variables used in matching are naturally continuous and take many, fine grained, values
- we are using a conservative coarsening approach

Having identified our matches, we therefore need to additionally conduct balance checks to determine whether the difference in other variables, on which matching was intended, varies between intervention and comparator local authorities post matching.

For both of our potential models, we test balance between groups on trends in care numbers, and any omitted scores. We find that our second model, which includes Employment rather than IDAOPI, creates a more balanced sample overall, except for with respect to IDAOPI, which is significantly imbalanced.

Nonetheless, our preferred model at present is the second model, which matches 18 Welsh local authorities with 29 English ones (Appendix A). We will work to expand the match to 22 local authorities by separating the remaining four Welsh authorities and taking a different approach to matching these.

Analytical specification: Survey Data

Our analytical strategy for the survey data will follow a matched difference in differences approach, with matching taking place prior to analysis in order to select the most appropriate sample. The difference in difference strategy for this data will be to take the first time period (T1) as baseline survey data for both the intervention and comparator groups, and to make use of the endline survey data (T2) as the second time period.

Primary Analysis

Our primary analysis specification will be conducted using ordinary least squares/linear prediction model regressions, specified as:

$$Y_{ilT2} = a + b_1 Y_{ilT1} + b_2 B_l + \Gamma_1 \mathrm{X}_i + \Gamma_2 L_l + e_l$$

Where Y_{ilT2} is the outcome measure for individual i in local authority l in time T2, a is a regression constant, Y_{ilT1} is the lagged value of the participant's outcome measure (their baseline score) at T1 and b_1 its coefficient estimate, B_l is a binary indicator of whether the participant was eligible for the basic income payments, set to 1 if they are eligible and 0 else (equivalent to a binary indicator for being 18 in Wales at time T1) with b_2 its coefficient estimate (treatment effect). X_i is a vector of participant level characteristics with Γ_1 representing the related vector of coefficients, similarly L_l is a vector of local authority level characteristics including those used in matching and Γ_2 its vector of coefficients, and e_l is an error term clustered at the level of the local authority. Other figures in the equation are coefficients.

T2 only data

Some variables in the survey are only collected at T2, and not at T1. For these variables, we will adopt a less typical approach and replace Y_{ilT2} with a vector of baseline variables that are the strongest control group predictors of the outcome at T2.

Secondary Analysis

Secondary analysis will follow the same regression specification as our primary analysis but replacing the variable Y with the relevant secondary outcomes.

Imputation

Inspection of the missing data pattern will provide some initial insight into the type of missingness, and statistical testing will further help assess whether the missing data mechanism is Missing Completely At Random (MCAR) or Missing At Random (MAR). We will utilise Little's MCAR test (Little, 1988) which determines whether the missingness is related to the observed or unobserved data. We will also use a logistic regression model with an indicator of missingness as the outcome, which will show whether relevant covariates are predictive of missingness, pointing towards the plausibility of a MAR assumption.

Due to the nature of survey data collection, and the use of some optional questions, we anticipate missing data for some participants. We will not make use of imputation for outcome measures as this carries substantial risks in terms of bias. For missing data at baseline (T1) we will make use of a mixture of Multiple Imputation through Chained Equations (White and others, 2011) in which available baseline or demographic data for the participant are used in regression analyses to calculate the likely values of the missing variable. Multiple imputation avoids the risks of overfitting or spurious precision by estimating the imputed values multiple times, by reintroducing into the model estimates the variability stemming from the predictive nature of imputed values. If instead only one dataset was used, we would be effectively treating imputed values as observed data, not properly accounting for the variances of the estimations in our model stemming from our imputed values being predictions, which would lead to type I errors. Where Multiple Imputation through Chained Equations is not possible due to a total lack of baseline or demographic data, we will make use of Null Imputation whereby missing values are imputed as 0 and a binary indicator for missingness is created.

Robustness Checks

We will robustness check our analyses by:

- using Null imputation across the board, where Multiple Imputation by Chained Equations (MICE) is not possible
- conducting complete case analysis
- sing logistic and probit regression for binary outcomes

Analytical specification: Administrative Data

Our analytical strategy for administrative data will differ from that used for survey data in a few important ways. Administrative records will contain richer baseline data on participants' characteristics and backgrounds. Records will not contain missing data with the same

frequency as survey data. Furthermore, data from administrative sources are available for additional cohorts. As a result of these factors, our analytical strategy for administrative data is different to that used in the survey data in the following ways.

Imputation

In this analysis we take a different approach to imputation. For outcomes, we will make use of complete case data only - that is, where a participant has complete information for their outcomes. Where participants are lacking data on baseline characteristics or demographics, this will be much more limited in the case of the administrative data. As a result, we will use null imputation for these small number of observations, where the values are imputed as 0 and a binary indicator is created which is set to 1 if that variable's value is missing and 0 if it is present. This is less sophisticated than the Multiple Imputation through Chained Equations approach used for the survey data but also requires fewer assumptions to be met.

Additional cohorts

We will make use of two additional cohorts in our final data analysis:

- young people who left care in the 12 months prior to the launch of the pilot - from July 2021 to June 2022
- young people who leave care in the 12 months after the pilot cohort - from July 2023 to June 2024 (the same group used as the comparator in Comparison A1)

Difference in difference counterfactual

In the administrative data analysis, the counterfactual we will use will not be the baseline data for the intervention group, but time period matching data for the comparator groups, with baseline data used as control variables.

We will analyse outcome measures from administrative data (identified in the section above on outcome measures). We will use Ordinary Least Squares/Linear Prediction Models for our analyses in order to maximise interpretability, while conducting robustness checks as identified in section 2.0. Our primary regression specification will be;

$$Y_{ilT2} = a + b_1 W_l + b_2 T_{T1} + b_3 (W_l \cdot T_{T1}) + \Gamma_1 X_i + e_{lT1}$$

Where Y_{ilT2} is the outcome measure for individual i in local authority I in time T2, a is a regression constant, W_l is a binary indicator that a local authority is in Wales, T_{T1} is a binary indicator of the treatment time period, set to 1 for the treatment time and 0 else, $(W_l \cdot T_{T1})$ is an interaction term between being in Wales and being in the treatment time period, taking a value of 1 for the treated cohort of young people and 0 else. The coefficient on this variable is our coefficient of interest. X_i is a vector of individual characteristics and e_{lT1} is an error term clustered at the local authority/time level.

Secondary Analysis

Our analytical strategy for our secondary outcomes will follow the same specification as our primary outcomes, replacing Y_{ilT2} with the relevant variables.

Robustness checks

Finally, we will conduct a number of checks to ensure the robustness of our findings. These will include:

- using MICE instead of Null imputation
- using triple differences including non-care leavers in the same local authority/time period pairs to further control for local variation and trend deviation
- using logistic and probit regressions for binary outcomes.
- placebo analyses

Statistical analysis plan for economic analysis

Introduction

In addition to estimating the effects on a range of outcomes, the study intends to calculate the economic benefits or costs of the pilot. This economic analysis plan sets out the planned approach to the cost-benefit analysis and is aligned to current guidance for economic analysis plans (Thorn and others, 2021).

Rationale for the economic evaluation

Economic evaluations consider resource use, costs and benefits of a policy intervention simultaneously. They are valuable for informing resource allocation decisions for health and social care interventions. The purpose of this economic evaluation, as determined by the research team, will be to consider whether the pilot represents value for money in terms of the outcomes achieved. It aims to identify, measure and value the costs and consequences of the pilot, and to synthesise the evidence to inform a social cost-benefit analysis.

Additionally, cost-consequences analyses (CCA) are recommended for complex interventions that may have multiple implications (Drummond and others, 2005), and for public health interventions which may have an array of benefits that are difficult to synthesise in a common unit. They present disaggregated costs and disaggregated consequences (primary and secondary outcomes), together with the estimates of the mean costs associated with the comparator interventions with appropriate measures of dispersion. The CCA will examine a range of key outcomes to explore wellbeing where the data allow, and aspects of educational attainment, engagement in the labour market and financial security.

Research Objectives

The primary research question for the economic evaluation is: how cost effective is the pilot? Supplementary research questions include:

- what are the costs of the pilot?
- from a public sector perspective, where do the costs and benefits of the pilot accrue?
- from a societal perspective, where do the costs and benefits of the pilot accrue?
- does the pilot generate economic gains in the longer-term?
- what are the cost-consequences of the pilot in terms of wellbeing, physical health, mental health, financial literacy and security, and do they help with both poverty reduction and improved access to labour the market, education and lifelong learning?

Perspective

Following recent recommendations by National Institute for Health and Care Excellence, and HM Treasury Guidance for appraisal and evaluation in government (the Magenta Book) (Drummond and others, 2005, NICE, 2022, HM Treasury, 2022), the primary analysis will adopt a public sector perspective. This helps detect cost shifting between sectors and considers the impact on society as a whole, rather than focusing only on a single department or agency's budget. Principles of opportunity cost will underpin all analyses including deadweight, leakage, displacement. Secondary analyses will extend wider costs and consequences to society.

Time horizon and follow-up

The time horizon will match the length of follow-up, that is, for 24 months following baseline data capture (T1) for the intervention and comparator cohorts.

Measurement of resource use

Data will be captured to estimate the resources and costs for the pilot and for support and/or wrap around services. Data will initially be collected from primary sources, most notably through a survey of care-leaving teams in local authorities, and from meetings with key stakeholders. A scoping exercise to inform a structured approach to data collection from stakeholders will be conducted, informed by the implementation research. Interviews and meetings will be used to understand the precise nature of the services and support provided, and thus the items to be costed, and the costs associated with these will be identified by written request to the relevant organisations and stakeholders.

To identify and measure the impact of the intervention, a broader configuration of service supports will be reviewed that may include programme management, training of frontline staff and education materials. Key managerial and administrative staff may include pilot funders and related stakeholders (Senior Civil Servants) within Welsh Government, managers (Heads of Children’s Services), social workers and administrators within local authorities, Personal Advisors and other key stakeholders. Throughput (the offer and uptake of services) and other associated costs data will be collected from relevant support services including Citizens Advice. There may be variation in the configuration of support, management and administrative services between and within settings (local authorities), and an audit of planned versus actual services delivered will identify these.

The identification of direct service costs may include data for all the costs involved in providing a service, specifically operational and overhead costs of running the service. These costs will be collected directly from local authorities and other services, via a survey proforma. Direct costs will account for the cost-of-service delivery, monitoring, follow-up, and associated administrative activities. Thus, data will be collected for total costs to deliver services, including professionals providing direct contact, management, supervision, total salary on-costs, training and travel costs associated with the delivery of the service, total

overheads, number of staff involved, total number of clients served and mean number of contacts from the commencement of support to completion in the pilot.

Intended and unintended cost impacts will be evaluated through survey-based data collection and informal co-production meetings, and in reference to the monitoring data. Changes in the circumstances of pilot recipients such as education, employment or training status will be explored.

The societal perspective will explore temporal changes in the policy environment, such as issues of demand and supply in the care-leaver housing market; for example, increases in supply (availability) of housing options; and whether housing and supported-accommodation costs rise to new thresholds.

Valuation of resource use data

Data items will be captured in disaggregated units where possible, and micro-costing will be performed to capture variance in costing patterns. Unit costs for each resource input will largely be derived from national secondary sources, for example the Department of Health & Social Care's NHS Reference Costs, the Personal Social Services Resource Unit (PSSRU) unit cost compendia, or through Spinal Column Points Salary Scales for each local authority (Jones and Burns, 2021). All resource use will be valued in monetary terms using the latest and most appropriate UK unit costs or participant valuations estimated at the time of analysis. The currency used will be expressed in British Pound Sterling (£), for a base cost year. Adjustments will be made for inflation using the PSSRU hospital & community health services index, and the Green Book discount rate, known as the Social Time Preference Rate, for use in UK government appraisal. All costs accrued beyond 12 months' follow-up will be discounted to present values using nationally recommended discount rates (Jones and Burns, 2021).

Data Analysis

The economic evaluation will be conducted alongside the implementation and impact evaluations, utilising the analytical designs and frameworks of those parts of the study. Specification of comparators and approaches for accounting for selection biases will mirror those planned for the impact evaluation and supplemented by guidance (Kreif, Grieve and Sadique, 2013; Deidda and others, 2019; Stinnett and Mallahy, 1998). The analysis will be informed by a comprehensive review of the broader literature regarding universal basic income and other similar cash transfer interventions, of which there are a growing number. The full analysis set will include all cost and outcomes variables, in accordance with the "intention to treat" principle, with all participants who are allocated to an intervention or control group should be analyzed in the group to which they were originally assigned, regardless of whether they actually received the intervention or not.

Utilisation of resource use items will be summarised by comparator group and differences between groups will be analysed using t-tests for continuous variables and Pearson chi-squared (x^2) tests for categorical variables. Mean differences in costs and outcomes between the intervention and comparator arms will be estimated and the bootstrap 95%

confidence interval will be computed based on 1,000 (or more) replications. Measures of uncertainty (standard errors and confidence intervals) will also be reported for the mean estimates.

Costs and benefits will aim to capture the direct impact of a policy on wellbeing and broader social impacts such as engagement in education, financial literacy, psychological well-being and other outcomes of interest. Principles of opportunity cost will underpin all calculations. Missing data from either self-report, linked data or participant surveys will be imputed where appropriate to reduce the impact of missing data on regression results. Sensitivity checks will be conducted against non-imputed data to assess the best/worst-case scenarios of missing outcomes, which could be particularly relevant if there are indications of Missing Not At Random (MNAR) patterns. This refers to missing data where the reason a value is missing is related to the missing value itself, or to unobserved variables patterns. Finally, narrative techniques will be used for outcomes that cannot be monetised, or where further exploration will be important, such as financial levers and incentives, mechanisms of change and unintended consequences. This will link to the theory enhancement and qualitative work packages of the evaluation.

Value for money will initially be expressed in terms of social cost-benefit at 24 months post-intervention, converting outcomes to monetary values. We will follow accepted guidelines outlined in the Green Book, which are constructed to explore the stated objectives of economy, efficiency and effectiveness (HM Treasury, 2022). All analyses will be carried out using appropriate analytical software such as STATA, R or Microsoft Excel, with relevant package and version numbers recorded at reporting stage.

Cost Consequences Analysis

The cost-consequences analysis will present resource use, costs and secondary outcomes. Each table will present the mean aggregate costs with their standard errors for the intervention and comparator, and the bootstrap mean difference (risk ratio secondary outcome) with their 95% confidence intervals. The costs and consequences will be presented in a disaggregated way, i.e. summary values and measures of dispersion for each cost category (and total costs) and each consequence or outcome associated with the comparators. Decision makers then place their own relative weights on the disparate consequences/outcomes.

Uncertainty and Sensitivity Analyses

We intend to conduct a range of sensitivity analyses to explore the impact of uncertainty surrounding key components of the economic evaluation on economic outcomes. These will be carried out for key costs and outcomes, specifically where they are highly sensitive to certain values or input variables such as direct service costs, support, management and administrative services between and within settings.

Sub-group analyses will mirror those undertaken for the main analysis. Summary statistics and cluster analysis may be used to determine data characteristics.

The economic evaluation will be reported in accordance with the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) 2022 reporting checklist (Husereau and others, 2022).

Conclusions

This report has detailed the analysis plans for two major aspects of the evaluation, and as such it complements and deepens the plans outlined in the study protocol (Westlake and others, 2024). The results of this analysis will enable us to estimate the average treatment effects of the pilot for the cohort as a whole, with the potential to consider subgroup treatment effects if sufficient data is available. Combined with the qualitative analysis detailed elsewhere (see Westlake and others, 2024) this will give a rounded picture of how effective the pilot has been on the outcomes of interest, and the ways in which it has made a difference to the lives of the young people involved.

The economic analysis will provide policy makers with further information that should inform decisions about the most cost-effective ways of supporting young people leaving care in Wales.

References

Briheim-Crookall L, Baker C and Selwyn J (2020) 'What Makes Life Good: Care leavers' views on their well-being' Coram Voice and The Rees Centre, University of Oxford

Center for Open Science (2015) 'Guidelines for Transparency and Openness Promotion (TOP) in Journal Policies and Practices "The TOP Guidelines" Version 1.0.1'

Chen H T (2014) 'Practical program evaluation: Theory-driven evaluation and the integrated evaluation perspective' Sage Publications

Department for Education (2024) 'Longitudinal Education Outcomes (LEO): collection' UK Government

Deidda M, Geue C, Kreif N, Dundas R and McIntosh E (2019) 'A framework for conducting economic evaluations alongside natural experiments' Social Science and Medicine volume 220, pages 353 to 361

Drummond M, Sculpher M, Torrance G, O'Brien B and Stoddart G (2005) 'Methods for the Economic Evaluation of Health Care Programmes' Oxford, Oxford University Press

Elman C, Kapiszewski D and Lupia A (2018) 'Transparent Social Inquiry: Implications for Political Science' Annual Review of Political Science volume 21, pages 29 to 47

HM Treasury (2020) 'The Magenta Book: Central Government Guidance on Evaluation' UK Government

HM Treasury (2022) 'The Green Book: Central Government Guidance on Appraisal and Evaluation Practice Guide' Crown copyright

Holland, S; Westlake, D; Roberts, L; Vallis, D; Lloyd, H; Sanders, M; Mathur, V; Bezeczky, Z;, Pickett, K; Johnson, M; Hick, R; Schroeder, E; Fahr, P (2024) Basic income for care leavers in Wales pilot evaluation report. Cardiff: Welsh Government, GSR report number 12/2024

Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs A H, Carswell C and others (2022) 'Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations' BMJ volume 376, e067975

Iacus S M, King G and Porro G (2012) 'Causal Inference without Balance Checking: Coarsened Exact Matching' Political Analysis volume 20, number 1, pages 1 to 24

Ioannidis, J.P.A. (2005) 'Why most published research findings are false', PLoS Medicine, 2(8), e124. doi: 10.1371/journal.pmed.0020124

Jones K and Burns A (2021) 'Unit Costs of Health and Social Care 2021' Personal Social Services Research Unit, University of Kent, Canterbury

Kreif N, Grieve R and Sadique M Z (2013) 'Statistical methods for cost-effectiveness analyses that use observational data: a critical appraisal tool and review of current practice' Health Economics volume 22, number 4, pages 486 to 500

Kroenke K, Spitzer R L and Williams J B W (2003) 'The Patient Health Questionnaire-2: Validity of a Two-Item Depression Screener' Medical Care volume 41, number 11, pages 1284 to 1292

Little R J A (1988) 'A Test of Missing Completely at Random for Multivariate Data with Missing Values' Journal of the American Statistical Association volume 83, number 404, pages 1198 to 1202

Living Wage Wales (2023) What is it? - Living Wage Wales

Mathur, V. Roberts, L., Bezeczky, Z, Lloyd, H., Vallis, D., Sanders, M., Pickett, KE, Johnson, M., Hick, R. Schroeder, E. Fahr, P. Petrou, S., Lee, H., Holland, S., and Westlake, D. (2025). Basic Income for Care Leavers in Wales Pilot Evaluation: Second Annual Report 2024 to 2025. Cardiff: Welsh Government, GSR report number 20/2025

Ministry of Housing, Communities & Local Government. (2019). English indices of deprivation 2019: domains of deprivation (File 2). GOV.UK

(2022) 'NICE health technology evaluations: the manual - Process and methods [PMG36]'

Ofosu G K and Posner D N (2023) 'Pre-Analysis Plans: An Early Stocktaking' Perspectives on Politics volume 21, number 1, pages 174 to 190

Open Science Collaboration (2015) 'Estimating the reproducibility of psychological science', *Science*, 349(6251), aac4716. doi: 10.1126/science.aac4716

Pawson R and Tilley N (1997) 'Realistic Evaluation' Sage Publications

Pawson R (2013) 'The Science of Evaluation: A Realist Manifesto' Sage Publications

Rosenbaum P R and Rubin D B (1983) 'The Central Role of the Propensity Score in Observational Studies for Causal Effects' Biometrika volume 70, number 1, pages 41 to 55

Sanders, M. and Vallis, D. (2023) Trial Protocol – Additional Financial Assistance for Care Leavers Trial, Centre for Homelessness Impact & King’s College London

Spitzer R L, Kroenke K, Williams J B W and Löwe B (2006) 'A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7' Archives of Internal Medicine volume 166, number 10, pages 1092 to 1097

StataCorp LLC (2019) 'Stata Statistical Software: Release 16' College Station, TX: StataCorp LLC

Sterne J and others (2009) 'Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls' BMJ volume 338, 2393

Stinnett A and Mallahy J (1998) 'Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis' Medical Decision Making volume 18, pages 68 to 80

Thorn J, Davies C, Brookes S, Noble S, Dritsaki M, Gray E and others (2021) 'Content of Health Economics Analysis Plans (HEAPs) for trial-based economic evaluations: expert Delphi consensus survey' Value in Health volume 24, pages 539 to 547

Welsh Government (2019) Welsh Index of Multiple Deprivation (full Index update with ranks): 2019. Cardiff: Welsh Government

Welsh Government (2022a) 'Basic income pilot for care leavers: overview of the scheme' Welsh Government

Welsh Government (2022b) Written Statement: Basic Income Pilot for Care Leavers in Wales (16 February 2022)

Welsh Government (2024) 'Basic Income for Care Leavers in Wales pilot statistics: August 2022 to July 2024' Welsh Government

Westlake D, Holland S, Sanders M, Schroeder E, Pickett K E, Johnson M and others (2024) 'The basic income for care leavers in Wales pilot evaluation: Protocol of a quasi-experimental evaluation' PLOS ONE volume 19, number 10, article number e0303837

White I R, Royston P and Wood A M (2011) 'Multiple imputation using chained equations: Issues and guidance for practice' Statistics in Medicine volume 30, number 4, pages 377 to 399