STATISTICAL ANALYSIS PLAN

Cash for Families: A randomised controlled trial of a social care intervention

Anna Freud, University of Exeter, University of Plymouth

Principal investigator: Julian Edbrooke-Childs

YOUTH ENDOWMENT FUND

Cash for Families: A randomised controlled trial of a social care intervention

Statistical analysis plan

Evaluating institution: Anna Freud

Principal investigator(s): Julian Edbrooke-Childs

Project title	Cash for Families: A randomised controlled trial of a social care intervention
Developer (Institution)	The Policy Unit, King's College London
Evaluator (Institution)	Anna Freud, University of Exeter, University of Plymouth
Principal investigator(s)	Julian Edbrooke-Childs
SAP author(s)	Julian Edbrooke-Childs, Vashti Berry, Nick Axford, Amy Bond, David Westlake, Matthew Constantinou, Abigail Thompson, Ping San, and Jessica Deighton
Trial design	Two-armed randomised controlled trial with randomisation at the individual index-child level
Trial type	Efficacy
Evaluation setting	Local Authorities
Target group	Families with at least 1 child with a Child in Need (CiN) or Child Protection Plan (CPP) status and who are eligible to receive financial support under Section 17
Number of participants	1,291 children (aged 10-16) recruited from 8-15 Local Authorities

Primary outcome and data source	Child-reported externalising difficulties measured by the conduct + inattention difficulties subscales of the Strengths and Difficulties Questionnaire (SDQ) measured at baseline (prior randomisation), endline (35 weeks post-randomisation), and follow-up (52 weeks post-endline.
Secondary outcome and data source	Child-reported outcomes measured at baseline and 35 weeks post-randomisation: • Internalising difficulties measured by the emotional + peer difficulties subscales of the SDQ (baseline, endline, and follow-up) • Total difficulties measured by the total difficulties score of the SDQ (baseline, endline, and follow-up) • Prosocial behaviour measured by the prosocial behaviour subscale of the SDQ (baseline, endline, and follow-up) • Impact of difficulties measured by the impact supplement of the SDQ (baseline, endline, and follow-up) • Family functioning measured by the total score of the Systemic Clinical and Routine Outcome Evaluation (SCORE-15; baseline and endline) • Substance and alcohol misuse (baseline and endline) • Time spent engaging in activities with parent/carer measured by the Global Appraisal of Individual Needs - Initial (GAIN-M90; baseline and endline) Parent/carer-reported outcomes measured at baseline and 35 weeks post-randomisation: • Parent/carer depression measured by the total score of the Patient Health Questionnaire-8 (PHQ-8; baseline and endline) • Parent/carer anxiety measured by the total score of the Generalised Anxiety Disorder (GAD-7; baseline and endline)

	• Parent/carer subjective wellbeing measured by the Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS; baseline, midline, and endline) • Family functioning measured by the total score of the SCORE-15; baseline and endline) Parent/carer-reported outcomes measured at midline (18 weeks post-randomisation): • Financial stress/circumstances measured by the modified Financial Stress Questionnaire and bespoke questions on worry about finances, family conflict about money, and home/neighbourhood satisfaction • Educational resources / opportunities measured by adapted items from HOME-21 • Material living conditions / housing quality measured by items adapted from the Center for Guaranteed Income Research • Parent/carer subjective wellbeing measured by the SWEMWBS • Parenting stress measured by the Parenting Stress Scale (PSS; Berry et al., 1995) • Time spent engaging in activities with child measured by 5 items from section B2 of the Global Appraisal of Individual Needs - Initial (GAIN-M90). Local Authority reported outcomes measured at 35 weeks post-randomisation: • Welfare concerns about the child measured by number of strategy discussions, contacts/referral to the Multi-Agency Safeguarding Hub, and number of other welfare concerns • Number of unauthorised school absences and fixed-term and permanent exclusions

SAP version history

Version	Date	Changes made and reason for revision
1.0 [***original***]	23/03/26	

Table of contents

Introduction

The Cash For Families study is a two-arm individually randomised controlled trial of a local authority intervention, providing unconditional cash transfers (UCTs) to families involved with Children's Social Care (CSC). Cash transfers are unconditional, meaning that no expectation is placed upon families (e.g., to seek work or take part in community service). Cash transfers are also unrestricted, meaning that families can choose to spend the money as they wish. Each family receives payments over 45 weeks. Families in London will be given 20% more. Payments are made weekly for 45 weeks and paid directly to the primary caregiver's bank account. There are fixed dates for starting the intervention and receiving cash, and there is tapering in the last 10 weeks. The intervention aims to improve child, parent and family outcomes by reducing family and financial stress and increasing families' ability to invest in their children's wellbeing and development.

Families are eligible if they have at least one child aged 10-16 years who is under a Child in Need (CiN) plan or Child Protection Plan (CPP), and who is eligible for financial support under Section 17 of the Children Act 1989. Participating families are randomly allocated in a 1:3 ratio to either receive UCT plus Business-as-Usual (BAU) services or to continue with BAU only (e.g., access to existing financial and material support from local authorities, and financial capability support like benefits). By over-recruiting to the control group with a 1:3 allocation ratio, we could maximise on power whilst recruiting enough families to the intervention condition within our budget. Randomisation occurs at the level of the index child, stratified by Local Authority (LA) and social care status (CiN vs. CPP).

Outcomes are collected at baseline (prior randomization), midline (18 weeks post-randomization), and endline (35 weeks post-randomization, before transfers are tapered) and finally at follow-up (52 weeks post-endline). The primary outcome is child-reported externalising difficulties, measured at baseline, endline and follow-up by the combined conduct and inattention subscales of the Strengths and Difficulties Questionnaire (SDQ; Goodman et al., 1998). Secondary outcomes include child-reported internalising difficulties, family functioning, substance and alcohol use, and parent/carer engagement; parent-reported depression, anxiety, well-being, parenting stress, financial stress, child engagement, and family functioning; and Local Authority-reported indicators of child welfare and school engagement.

The main purpose of the analysis is to evaluate whether providing UCTs to families reduces externalising difficulties compared to BAU alone. Additional analyses will examine the effect of UCTs vs. BAU on secondary outcomes, explore subgroup effects (e.g., by income, social care status, amount of cash received, number of children in the family, child gender and

ethnicity), and evaluate the two main causal pathways in the literature: the Family Stress Model and Resource Investment Model. The Family Stress Model suggests that UCTs improve child outcomes through reducing caregiver distress, parental conflict, and disrupted parenting linked to economic hardship. The Resource Investment Model suggests that UCTs improve child outcomes through increasing investment in material resources (e.g., higher quality housing) and child engagement (e.g., more family time, engagement in extracurricular activities). We will test these pathways through a mediation model, whereby (1) UCT predicts financial stress at midline, (2) financial stress then predicts family stress and home and child resources at midline, and (3) family stress and home and child resources in turn predict changes in child behavioural difficulties at endline.

Design overview

Trial design, including number of arms		Two-arm efficacy randomised controlled trial: unconditional cash transfer + business as usual vs. business as usual only
Unit of randomisation		Individual index-child level
Stratification variables (if applicable)		Local authority, social care status (Child in Need vs. Child Protection Plan)
Primary outcome	variable	Behavioural difficulties
	measure (instrument, scale, source)	Child-reported externalising difficulties measured at baseline (prior randomization), endline (35 weeks post-randomization), and follow-up (52 weeks post-endline) by the combined conduct and inattention difficulties subscales of the Strengths and Difficulties Questionnaire (SDQ; Goodman et al., 1998).
Secondary outcome(s)	variable(s)	Child-reported outcomes: • Internalising difficulties (1) • Total difficulties (2) • Prosocial behaviour (3) • Impact of difficulties (4) • Family functioning (5) • Substance and alcohol misuse (6) • Time spent engaging in activities with parent/carer (7) Parent/carer-reported outcomes: • Depression (1) • Anxiety (2) • Wellbeing (3) • Family functioning (4)

		• Financial circumstances and stress (5) • Educational resources / opportunities (6) • Material living conditions / housing quality (7) • Parenting stress (8) • Time spent engaging in activities with child (9) Local Authority reported outcomes: • Child welfare concerns (1) • School attendance and fixed-term and permanent exclusions (2)
	measure(s) (instrument, scale, source)	Child-reported outcomes: • (1-4) Strengths and Difficulties Questionnaire (SDQ; Goodman et al., 1998) at baseline, endline (35 weeks post-randomization), and follow-up (52 weeks post-endline) • (5) Systemic Clinical and Routine Outcome Evaluation (SCORE-15; Stratton et al., 2010) at baseline and endline • (6) Custom substance and alcohol use items (binary “yes” or “no” reports of use within the past month) at baseline and endline • (7) Global Appraisal of Individual Needs-Initial (GAIN-M90), Section B2 (Dennis et al., 2022) at baseline and endline Parent/carer-reported outcomes: • (1) Patient Health Questionnaire-8 (PHQ-8; Kroenke, Spitzer, & Williams, 2009) baseline and endline • (2) Generalized Anxiety Disorder-7 (GAD-7; Spitzer, Kroenke, Williams, & Löwe, 2006) baseline and endline

		• (3) Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS; Tennant et al., 2007) baseline, midline, and endline • (4) Systemic Clinical and Routine Outcome Evaluation (SCORE-15; Stratton et al., 2010) at baseline and endline • (5) Custom income question; Financial Stress Questionnaire (Center for Guaranteed Income Research); bespoke questions on worry about finances, family conflict about money, and home/neighbourhood satisfaction at midline • (6) Adapted Home Observation for Measurement of the Environment (HOME-21; Caldwell & Bradley, 1984) at midline • (7) Adapted Housing quality items (Center for Guaranteed Income Research) • (8) Parenting Stress Scale (PSS; Berry et al., 1995) at midline • (9) Global Appraisal of Individual Needs-Initial (GAIN-M90), Section B2 (Dennis et al., 2022) at baseline and endline Local Authority reported outcomes: • (1) Social worker reported number of strategy discussions, referrals to the multi-agency safeguarding hub (MASH), and welfare concerns reported at endline • (2) Social worker reported number of unauthorised absences, fixed-term, and permanent exclusions at endline
Baseline for primary outcome	**variable**	Behavioural difficulties
	measure (instrument, scale, source)	Child-reported externalising difficulties measured by the combined conduct and inattention difficulties

		subscales of the Strengths and Difficulties Questionnaire (SDQ; Goodman et al., 1998).
Baseline for secondary outcome	**variable**	Child-reported outcomes: • Internalising difficulties (1) • Total difficulties (2) • Prosocial behaviour (3) • Impact of difficulties (4) • Family functioning (5) • Substance and alcohol misuse (6) • Time spent engaging in activities with parent/carer (7) Parent/carer-reported outcomes: • Depression (1) • Anxiety (2) • Wellbeing (3) • Family functioning (4)
	measure (instrument, scale, source)	Child-reported outcomes: • (1-4) Strengths and Difficulties Questionnaire (SDQ; Goodman et al., 1998) at baseline, endline (35 weeks post-randomization), and follow-up (52 weeks post-endline) • (5) Systemic Clinical and Routine Outcome Evaluation (SCORE-15; Stratton et al., 2010) at baseline and endline • (6) Custom substance and alcohol use items (binary “yes” or “no” reports of use within the past month) at baseline and endline • (7) Global Appraisal of Individual Needs-Initial (GAIN-M90), Section B2 (Dennis et al., 2022) at baseline and endline Parent/carer-reported outcomes:

		• (1) Patient Health Questionnaire-8 (PHQ-8; Kroenke, Spitzer, & Williams, 2009) baseline and endline • (2) Generalized Anxiety Disorder-7 (GAD-7; Spitzer, Kroenke, Williams, & Löwe, 2006) baseline and endline • (3) Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS; Tennant et al., 2007) baseline, midline, and endline • (4) Systemic Clinical and Routine Outcome Evaluation (SCORE-15; Stratton et al., 2010) at baseline and endline

Sample size calculations overview

		Protocol	Randomisation
Minimum Detectable Effect Size (MDES)		Standardised mean difference = 0.19	TBC*
Pre-test/ post-test correlations	level 1 (participant)	-	-
	level 2 (cluster)	-	-
Intracluster correlations (ICCs)	level 1 (participant)	-	-
	level 3 (cluster)	-	-
Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		Two-sided	Two-sided
Average cluster size		-	-
Number of clusters	intervention	-	-
	control	-	-
	total	-	-
	intervention	323	TBC*

		Protocol	Randomisation
Number of participants	control	968	TBC*
	total	1,291	TBC*

* The last participant is expected to be randomized on 17 May 2026. The actual sample size and MDES calculation are anticipated to be updated by 1 June 2026.

The primary population of interest are families currently engaged with children's services, with at least one child aged 10-16 who is classed as a Child in Need or who is subject to a Child Protection Plan. One child will be selected as the index child from each family for which primary and secondary outcomes will be completed.

The sample size was determined a priori using the PowerUp! tool (Dong & Maynard, 2013). The protocol presented different scenarios for sample size calculations based on a MDES between 0.15 and 0.20 and comparing 1:3 and 1:1 allocation ratios at 80% power and a two-sided alpha of 0.05 (see Appendix 6 in protocol). We chose an MDES of 0.19 based on Ozer et al.'s (2009) finding of a reduction in aggressive and oppositional problems following cash transfers. Furthermore, we chose a 1:3 allocation ratio between the intervention (UCT+BAU) and control (BAU-only) arms to optimise costs by providing comparable power to a 1:1 ratio while reducing the number of families receiving cash transfers. In the present SAP, we also consider attrition rates of 10%, 15% and 20% and pre-post correlations of 0 and 0.40 (Goodman, 2001; see table below).

Sample Size Estimates For Different Attrition Rates (10%, 15%, 20%) and Pre-Post Correlations (*r* = 0 or 0.40).

Attrition Rate	**Pre-post (*r* = 0)**	**Pre-post (*r* = 0.40)**
10%	1,291	1,080
15%	1,343	1,129
20%	1,401	1,178

Across these estimates, recruiting between 1,178-1,291 total families would allow us to detect an MDES of 0.19 under more and less favourable conditions of attrition and pre-post correlation. On this basis, we plan to randomize 1,291 children: 323 in the UCT + BAU arm and 968 in the BAU-only arm. Assuming 75% of referred families will consent to take part, 1,721 referrals are required to achieve the target MDES of 0.19. Our sample size estimate of 1,291 has been inflated for 10% attrition, which is based on a previous cash transfer trial with families that initially estimated a 20% attrition rate but observed rates of 6% after a one-year follow-up – 3% in the intervention arm and 8% in the control arm (Duncan et al., 2025; Magnuson et al., 2025). The recruitment and consent rate will be reassessed after recruiting the first cohort.

Note that we did not adjust the sample size for clustering of families within local authorities (LA) because randomization is at the family (individual) level. A cluster-RCT-type sample size inflation is neither necessary nor intended. Instead, nesting will be handled analytically through stratification by LA and adjusting for LA as a fixed effect in our statistical model.

Analysis

Analyses will be conducted by an analyst masked to intervention group assignment following the intention-to-treat principle. The hypotheses and general analysis framework were decided a priori; the first draft analysis plan was written during the recruitment period, before baseline data collection was completed.

Primary outcome analysis

To recap, the primary outcome is child-reported externalizing difficulties (i.e. conduct + inattention subscales of the SDQ; 0-20) at endline (35 weeks post-randomization). We have chosen externalising difficulties, rather than a broader total difficulties score, as previous studies of cash transfer interventions show that effects for young people are most pronounced for externalising difficulties (Jaffee et al., 2025), but we expect widespread changes in child mental health (see Secondary Outcomes Analysis below).

Our headline impact estimate is the intention-to-treat effect of UCT+BAU versus BAU-only on externalizing difficulties among all randomised children. We will use a generalized linear model whilst controlling for the fixed effects of baseline externalizing scores and stratification variables (LA membership and social care status). The model can be expressed as follows:

$$Y_i = \beta_0 + \beta_1 Baseline_i + \beta_2 Group_i + \beta_3 Status_i + \beta_4 LA_i + \beta_5 Wave_i + \varepsilon_i$$

Where:

- Y is the predicted SDQ externalizing score for index child i at endline
- β_0 is the intercept or predicted SDQ externalizing score when all covariates equal zero
- $\beta_1 Baseline_i$ is the baseline SDQ externalizing score for index child i
- $\beta_2 Group_i$ is the randomly assigned intervention group (0 = BAU only; 1 = UCT+BAU)
- $\beta_3 Status_i$ is the index child's social care status (0 = Child in Need; 1 = Child Protection Plan)
- $\beta_4 LA_i$is the index child's local authority membership (dummy coded)
- $\beta_5 Wave_i$ is the wave within which the family were recruited into the study
- ε_i is the residual variance for index child i

Estimates will be reported with 95% confidence intervals and exact, continuous *p* values. Alpha thresholds will be set at the 5% level unless otherwise stated, and dichotomous interpretation around specific thresholds will be avoided. Standard errors will be estimated using robust methods (e.g., Huber-White) to account for potential heteroskedasticity in residual errors. Analyses will be conducted in Stata v14.

Secondary outcome analysis

To recap, secondary outcome variables are as follows:

Child-reported outcomes:

- Internalizing difficulties [peer problems + emotional problems scales, 0-20], prosocial behaviour (0-10), total difficulties (0-40), impact of difficulties (0-10) scales of the SDQ
- Family functioning total scores (SCORE-15, 15-75)
- Time spent engaging in activities with parent/carer (GAIN-M90; 0-5)
- Substance and alcohol misuse (custom items; 0-4)[a]

Parent/carer-reported outcomes:

- Family functioning total scores (SCORE-15, 15-75)
- Total depression symptoms (PHQ-8, 0-24)[a]
- Total anxiety scores (GAD-7, 0-21) [a]
- Total wellbeing scores (SWEMWBS, 7-35)[a]

Local Authority reported outcomes:

- Child welfare concerns (number of strategy discussions, number of referrals to the Multi-Agency Safeguarding Hub, number of welfare concerns, whether the child went into care or was reunited with a primary caregiver)[a]
- School attendance and fixed-term and permanent exclusions (number of unauthorised absences; number of fixed-term and permanent exclusions)[a]

[a]Secondary outcomes treated as exploratory.

In similar projects, we consulted young people of a similar age to the target sample. They completed the full set of measures within 10–15 minutes. Based on an estimated completion rate of three to four items per minute for the target population, we anticipate that the surveys for both the young person and parent-carers will take approximately 15–20 minutes to complete. We will review participant burden during Cohort 1 and, based on completion rates, will consider prioritising secondary outcomes for Cohort 2.

For child- and parent-reported secondary outcomes, we will use a similar, single-level generalized linear model like in the primary outcome analysis:

$$Y_i = \beta_0 + \beta_1 Baseline_i + \beta_2 Group_i + \beta_3 Status_i + \beta_4 LA_i + \beta_5 Wave_i + \varepsilon_i$$

Where:

- Y is the predicted score for a given secondary outcome for index child i at endline
- β_0 is the intercept or predicted score on a given secondary outcome when all covariates equal zero
- $\beta_1 Baseline_i$ is the baseline score on a given secondary outcome for index child i
- $\beta_2 Group_i$ is the randomly assigned intervention group (0 = BAU only; 1 = UCT+BAU)
- $\beta_3 Status_i$ is the index child's social care status (0 = Child in Need; 1 = Child Protection Plan)
- $\beta_4 LA_i$is the index child's local authority membership (dummy coded)
- $\beta_5 Wave_i$ is the wave within which the family were recruited into the study
- ε_i i s the residual variance for index child i

For local authority outcomes, we will use parametric tests (e.g., poisson regression for count variables and independent t-tests for percentage variables) or non-parametric tests (e.g., Mann-Whitney U or Kolmogorov-Smirnov tests) to compare UCT+BAU and BAU-only groups, as no baseline data will be collected.

Estimates will be reported with 95% confidence intervals. Secondary outcomes related to child mental health (SDQ total scores and subscales), family functioning (SCORE-15), and family activities (GAIN-M90) are treated as confirmatory, given prior evidence that cash transfers can affect child behavioural and family functioning domains. Whilst we did not

formally estimate power for our secondary outcomes, we consider an MDES of 0.19 to be a meaningful benchmark. Analyses of confirmatory secondary outcomes will be adjusted for multiple comparisons.

Outcomes related to parental mental health and wellbeing, child welfare concerns and school attendance, and child substance/alcohol use are treated as exploratory, as there is limited prior evidence on the effects of cash transfers on these outcomes. Analyses will be conducted in Stata v14.

Subgroup analyses

We will test for moderation by including an interaction term between intervention assignment (UCT+BAU vs. BAU-only) and each moderator variable across the whole sample to the primary analysis model (see above). This approach follows YEF guidance and avoids the loss of power associated with splitting the sample into separate groups. Each moderator will be tested separately. Continuous moderators will be mean-centred.

The following subgroups have been pre-specified in the protocol and are therefore considered confirmatory:

- Total annual household income pre-tax at baseline (ordinal):
    - <£15,000
    - £15,000-£24,999
    - £25,000-£34,999
    - £35,000-£44,999
    - £45,000-£59,999
    - £60,000-£79,999
    - >=£80,000

  (Note. Participants are provided with guidance on how to respond to this item, including the types of income to report, to minimise self-reporting errors (e.g., income from all individuals living in the household, wages, benefits, pensions, business income, and investments).
  We will consider collapsing into low, mid, and high incomes relative to UK percentiles for that year.

- Amount of cash received in the intervention group (continuous; to be used as a measure of intervention dose to evaluate fidelity. Note, we will consider transforming or using as a proportion depending on the distribution)

- Number of children in the household (continuous)
- Child gender (categorical):
    - Male
    - Female
    - Non-binary
    - Gender questioning
    - Prefer to self-describe
    - Prefer not to say

We will consider collapsing into male/female/gender diverse.

- Child ethnicity (nominal):
    - Arab or Arab British
    - Asian or Asian British
    - Black, Black British, Caribbean or African
    - Mixed or Multiple Ethnic Groups
    - White
    - Other ethnic group
    - Prefer not to say

We will consider analysing ethnic subgroups (e.g., African, Caribbean, Any Other Black or British or Caribbean background) if sample sizes permit. If sample sizes in non-White categories are small, we will consider collapsing into "White" vs. "minority ethnic group".

Significant interactions ($p < 0.05$) will be explored using predicted marginal means and simple-slope plots. Results will be reported with interaction coefficients, 95% confidence intervals, and adjusted *p*-values to account for multiple testing.

Further analyses

Further analyses include sensitivity analyses to test the robustness of primary findings, and mediation analyses to test the theory of change.

Sensitivity and robustness checks

To assess the robustness of findings from the primary outcome analysis, we will run a series of pre-specified sensitivity analyses by modifying the generalized linear model specified in the primary outcome analysis above with the following:

- Model specification: inclusion of additional child-, parent-, and area-level covariates that may confound treatment effects (child age, parent/carer age, area-level deprivation, and LA characteristics[1]), recruitment wave, and any covariates that are unbalanced between the intervention and control groups.
- Missing data: comparing results from the ITT model with missing data, to the results after handling missing data (see Missing Data section below).

Mediation analysis

We will run a series of mediation analyses to evaluate two main causal pathways in the literature: the Family Stress Model and Resource Investment Model (see Introduction). Specifically, we will examine whether financial circumstances at midline predict between-family variation in family stress and home/educational resources at midline, and whether family stress and home/educational resources in turn mediate the relationship between intervention allocation (UCT+BAU vs. BAU-only) and the primary outcome (behavioural difficulties) at endline. Note, we did not specify these models in full or power for them in the protocol.

To recap, midline/mediator variables include:

Financial Stress/Circumstance

- Composite measure of FSQ and bespoke items

Family Stress

- Parenting stress (PSS, 18-90)
- Parent subjective wellbeing (SWEMWBS, 7-35)
- Family conflict about money (Bespoke measure: 'Talking about money at home')

Resource Investment

- Educational resources / opportunities (Adapted HOME-21, 0-21)
- Time parent and child spend together doing activities (GAIN-M90; 0-5)

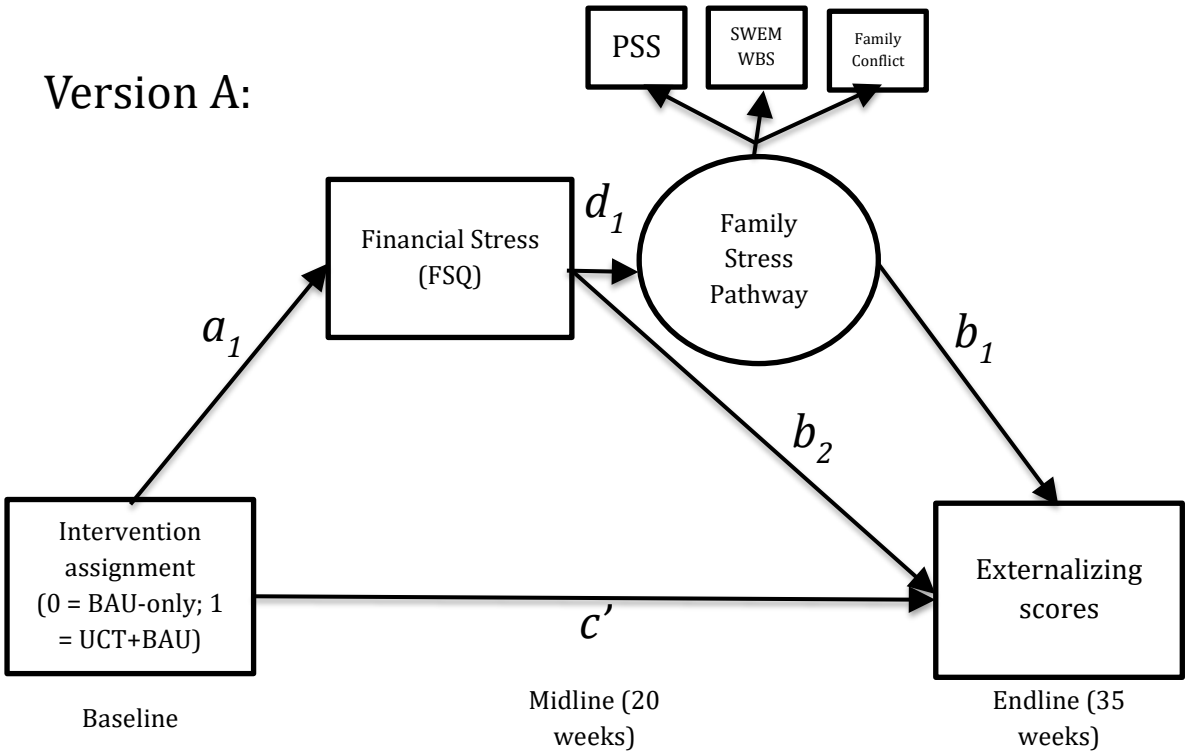
[1]We do not expect differences in the intervention and control arms by local authority following stratified randomization. However, any differences in LA characteristics that persist between the intervention and control groups will be controlled for via fixed effects to provide an extra layer of precision.

- Material living conditions / housing quality (items adapted from the Center for Guaranteed Income Research)

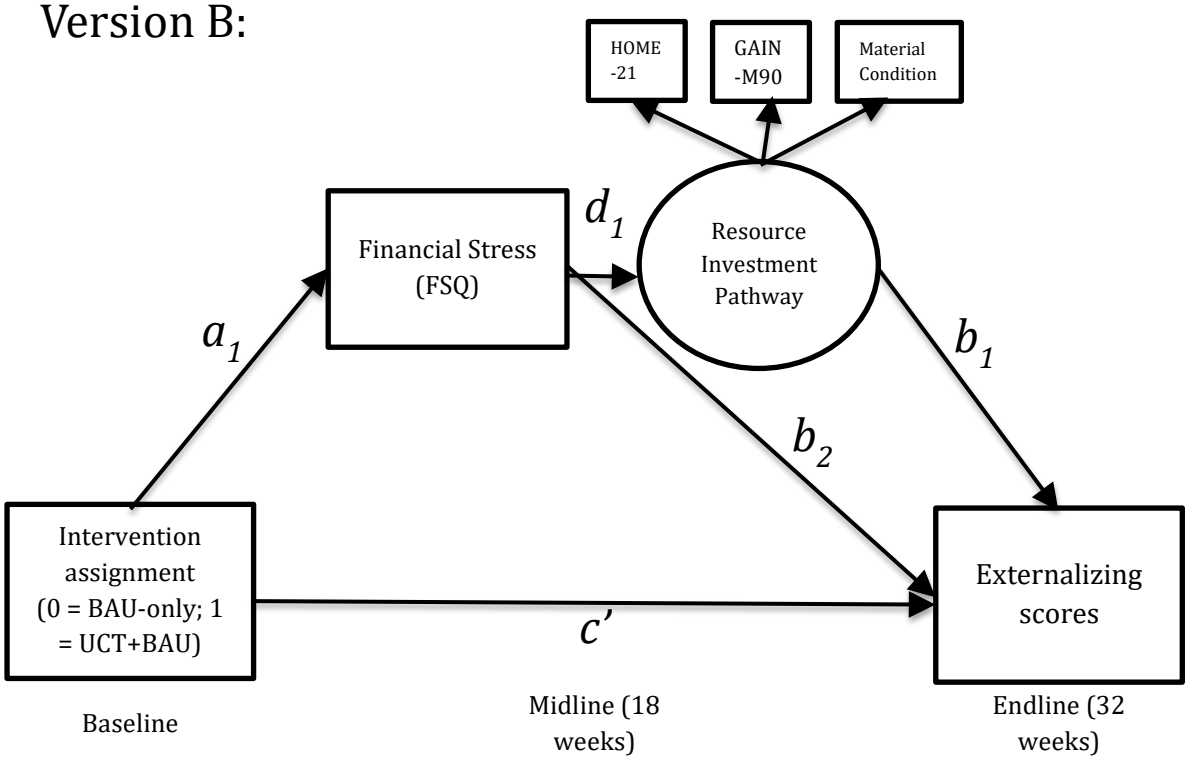
We will run the mediator analysis in stages. In the first stage, we will review the degree of overlap between variables using bivariate correlations and the variance inflation factor (VIF), where $r > 0.7$ and VIF > 5 indicate possible conceptual overlap. We hypothesise that financial stress will be strongly positively correlated with family stress and home/educational resources (indicating shared common variance), while the latter two variables will be moderately, positively correlated (indicating partially overlapping but unique causal pathways).

In the second stage, where we will test whether baseline intervention assignment predicts endline externalizing problems directly and indirectly, first via financial stress, and second via family stress or home/educational resources (see figure below). We will create composite variables related to the family stress pathway (averaging standardized versions of the parenting stress, reverse-coded parent subjective wellbeing, and family conflict about money, or creating a latent variable with both measures and testing model fit, e.g., CFI > .95, TLI > .95, and RMSEA < .06) and the resource investment pathway (education resources/opportunities, parent-child activity, and material living conditions). We will include family stress (version A) and resources investment (version B) in separate models to understand their unique associations, in addition to baseline externalizing scores and stratification variables (not included in the figure below).

Version A:

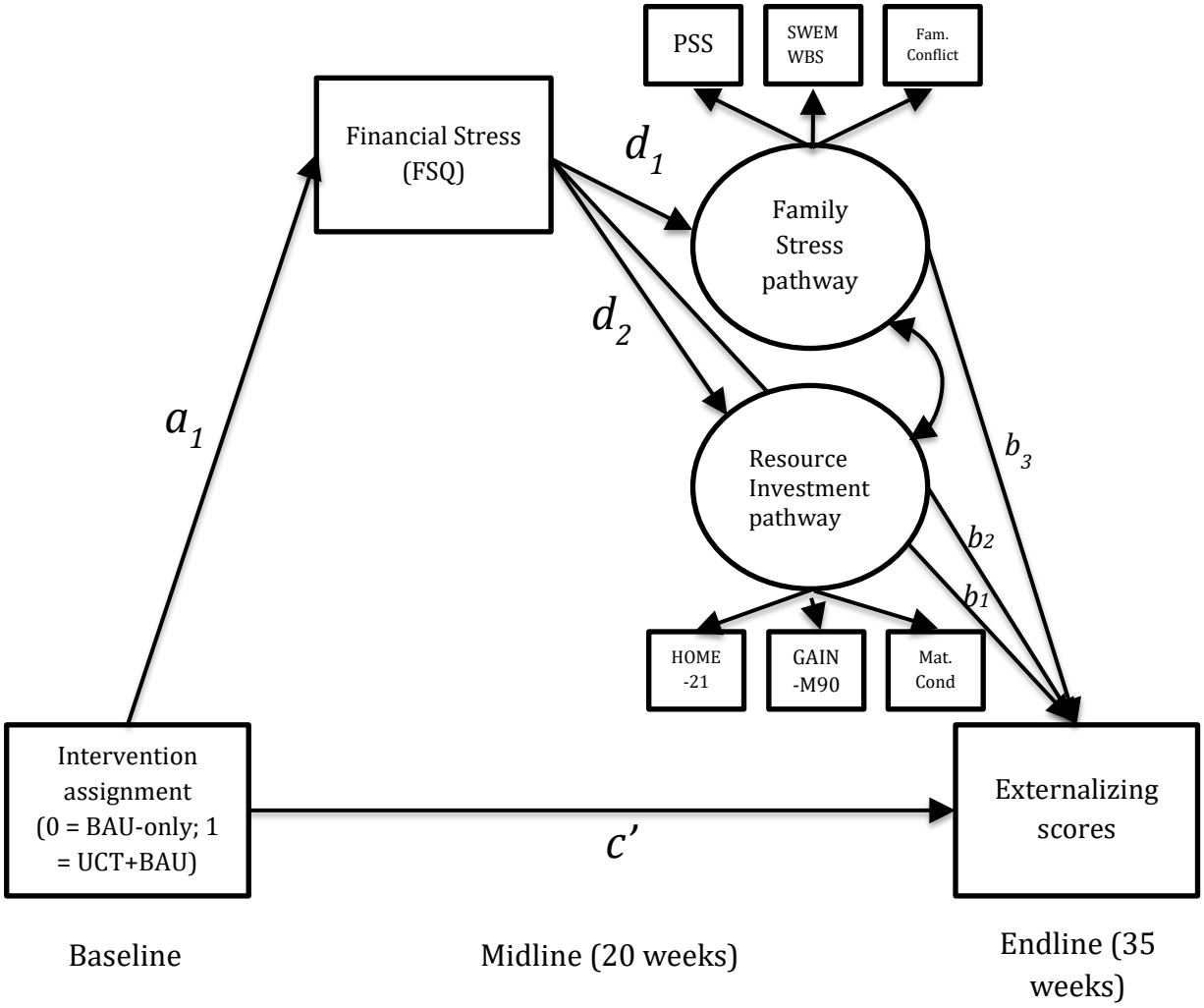


Version B:



Note. Variables are treated as latent variables in these figures, but we might instead use composites of observed variables depending on power limitations.

In the third stage, we will run a parallel mediation model including both family stress and resource investment pathway composites to assess their relative indirect associations on endline externalising problems.



Indirect effects in each model will be estimated using bias-corrected bootstrapped 95% confidence intervals (5,000 draws). Maximum likelihood estimation with robust standard errors (MLR) will be used.

In the fourth stage, if both family stress and resource investment pathways are supported and power allows, we will use sequential models to explore if one pathway's associations predict the other. Furthermore, if power allows, we will also explore moderated mediation to test

whether the indirect effects of family stress and/or resource investment differ across ethnic groups. We will run all mediation analyses using Mplus v8.

Interim analyses and stopping rules

We did not specify an interim analysis or stopping rules in the protocol. We will, however, frequently monitor trial progress and will prepare a report in [TBC] for the Data Monitoring and Ethics Committee (DMEC) covering recruitment and adverse events. We will stop the trial if the DMEC make a recommendation for the trial to terminate early based on safety concerns**.**

Longitudinal follow-up analyses

Children will complete the SDQ at 52 weeks post-endline. Scores will be analysed using the primary outcome analysis model, with SDQ externalizing scores at 52-week post-endline follow-up as the main outcome, controlling for baseline externalizing scores. We will consider running post-hoc paired samples t-tests comparing baseline to endline externalizing scores, and endline to follow-up scores.

Imbalance at baseline

We will create a table of baseline characteristics measured at enrolment, including the primary outcome (e.g., SDQ externalizing scores), secondary outcomes, stratification variables (e.g., the % of UCT+BAU vs. BAU-only in each LA and the proportions of CiN and CPP in each arm), and demographic variables (e.g., child and parent/carer sex, gender, age, ethnicity; parent/carer-reported number of jobs, hours worked per week, type of employment, highest level of education, benefits received, household income, household composition, neighbourhood deprivation).

We will include separate columns of means and standard deviations (for continuous variables) and counts and percentages (for binary and count data) for the UCT+BAU and BAU-only groups. Furthermore, we will present histograms of frequency distributions for pre-test scores on primary and secondary continuous outcomes. We will compare group means/percentages using independent t-tests (continuous variables) and Chi-Square tests (categorical/count variables) to evaluate whether the randomization was successful.

Differences between the randomised sample and the analytic sample at endline will be presented in future reports to assess whether attrition introduced systematic imbalance. A

supplementary table with LA characteristics (e.g., ethnic composition, % CiN vs. CPP) will also be provided.

Missing data

We will assess the extent and pattern of missing data by reporting the number and percentage of complete cases for each outcome variable, the percentage of missing values per variable, and cross-tabulations to describe patterns of missingness across variables.

To investigate whether missingness in the primary outcome, secondary outcome, and covariates is systematically related to observed variables, we will estimate logistic regression models predicting missingness (0 = observed, 1 = missing) separately for baseline and endline variables. Baseline models will include baseline predictors only, whereas endline models will include baseline and endline predictors to preserve temporality.

Primary analyses involving more than 5% missing data will assume a Missing At Random (MAR) mechanism and will be conducted using Multiple Imputation by Chained Equations (MICE; van Buuren & Groothuis-Oudshoorn, 2011) in R. The imputation model will include all variables from the primary or secondary analysis models, as well as auxiliary variables associated with missingness or the outcome ($|r| \geq 0.2$ or $p < .05$ in logistic regressions; van Buuren, 2018). Temporal ordering will be preserved.

At least 25 imputations will be generated, with the number of imputations matching the percentage of missing data (e.g., 40% missing → 40 imputations). Parameter estimates will be pooled using Rubin's rules. Model convergence and plausibility of imputed values will be assessed using diagnostic plots and summary statistics.

Results will be compared to a complete-case analysis, Full Information Maximum Likelihood (FIML) analysis, and delta-adjusted pattern-mixture models.

We will conduct sensitivity analyses under Missing Not At Random (MNAR) mechanisms using delta-adjusted pattern-mixture models (Carpenter, Kenward, & White, 2007). Imputed values for missing endline outcomes will be shifted by δ values of 0.25, 0.5, and 1.0 standard deviations of the observed outcome distribution to represent plausible departures from MAR. Treatment effects will be re-estimated under each scenario. A tipping-point analysis will identify the magnitude of δ required to alter the substantive conclusions of the trial.

Compliance

As specified in the protocol, compliance with the intervention will be assessed using routinely collected data from the payment agent on the amount of cash successfully transferred to a parent/carer. A threshold of 80% of cash successfully transferred to each family (i.e., without delay/disruption) will be used to indicate compliance.

Complier average causal effect (CACE) analysis will be used to estimate the intervention effect among compliers in the intervention group, and control participants who would comply if offered the intervention. A two-stage least squares (2SLS) approach will be used with intervention assignment as a predictor for compliance status (0 = < 80% compliance, 1 = ≥ 80% compliance), followed by expected compliance (the outcome variable from the first stage) as a predictor of the primary outcome (externalizing difficulties). The 2SLS approach can be expressed as follows:

Stage 1:

$$\text{Compliance}_i = \pi_0 + \pi_1 Group + \pi_2 \text{X}_i + u_i$$

Stage 2:

$$Y_i = \beta_0 + \beta_1 Compliance + \beta_2 \text{X}_i + \varepsilon_i$$

Where:

- Compliance_i = predicted compliance predicted by intervention assignment and covariates.
- Y_i= endline externalising difficulties (primary outcome).
- X_i= baseline covariates (baseline externalizing scores, social care status, LA membership, etc).

Intra-cluster correlations (ICCs)

We will estimate intra-cluster correlations (ICCs) for the primary outcome (externalizing problems) to describe the proportion of variance attributable to between- vs. within-LA differences.

For baseline and endline externalizing scores, we will fit two-level random-intercept models with children (level 1) nested within LAs (level 2):

$$Y_{ij} = \gamma_0 + u_{0j} + \varepsilon_{\text{ij}}$$

Where:

- $Y_{ij}$is the outcome for child i in LA j
- $u_{0j} \sim \mathrm{N}(0, \tau^2)$ is the LA-level random effect
- $\varepsilon_i \sim \mathrm{N}(0, \sigma^2)$ is the child-level residual.

The LA-level ICC will be calculated as:

$$\mathrm{ICC}_{\mathrm{LA}} = \frac{\tau^2}{\tau^2 + \sigma^2}.$$

ICCs will first be calculated from an empty model (no covariates) and then from a model including covariates from the primary outcome analysis (e.g., baseline externalizing scores, intervention assignment, LA membership, social care status, and recruitment wave). ICC estimates will be reported with 95% confidence intervals.

Presentation of outcomes

We will report Cohen's *d* for continuous outcomes and risk ratios (RR) with 95% confidence intervals for binary outcomes.

Cohen's *d* will be calculated as:

$$d = \frac{\hat{\delta}}{\hat{\sigma}_{\mathrm{uncond}}}$$

where:

- $\hat{\delta}$ reflects the adjusted mean difference between UCT+BAU and BAU-only groups on a given outcome at endline
- $\hat{\sigma}_{\mathrm{uncond}}$ is the unconditional pooled standard deviation for a given outcome at endline (i.e., the square root of the pooled variance across both groups from a model without covariates).

RRs will be calculated as:

$$RR = \frac{P(Y = 1 \mid \text{Intervention})}{P(Y = 1 \mid \text{Control})}$$

which is the ratio between the probability that the outcome is a 1 vs. 0 for the intervention group (UCT+BAU) relative to the probability that the outcome is a 1 vs. 0 for the control group (BAU-only).

95% bootstrapped confidence intervals (≥1,000 resamples) will be reported to reflect statistical uncertainty.

References

Berry, J. O., & Jones, W. H. (1995). The Parental Stress Scale: Initial psychometric evidence. *Journal of Social and Personal Relationships, 12*(3), 463–472. https://doi.org/10.1177/0265407595123009

Caldwell, B. M., & Bradley, R. H. (1984). *Home observation for measurement of the environment*. Little Rock, AR: University of Arkansas at Little Rock.

Carpenter, J. R., Kenward, M. G., & White, I. R. (2007). Sensitivity analysis after multiple imputation under missing at random: a weighting approach. *Statistical Methods in Medical Research,* 16(3), 259-275. https://doi.org/10.1177/0962280206075303

Dennis, M. L., & Chestnut Health System J. (2022). *The Global Appraisal of Individual Needs (GAIN): Administration and scoring guide.* Lexington, KY: Lighthouse Institute.

Dong, N., & Maynard, R. (2013). PowerUp!: A tool for calculating minimum detectable effect sizes and sample size requirements for experimental and quasi-experimental designs. *Journal of Research on Educational Effectiveness, 6*(1), 24–67. https://doi.org/10.1080/19345747.2012.673143

Duncan, G. J., Magnuson, K., Kunin-Batson, A. S., et al. (2025). Cash Transfers and Their Effect on Maternal and Young Children's Health: A Randomized Clinical Trial. *JAMA Pediatrics.* 179(8), 867–875. https://doi.org/10.1001/jamapediatrics.2025.1612

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(11), 1337–1345. https://doi.org/10.1097/00004583-200111000-00015

Goodman, R., Meltzer, H., & Bailey, V. (1998). The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *European Child & Adolescent Psychiatry, 7*(3), 125–130. https://doi.org/10.1007/s007870050057

Jaffee, S. R., Lin, G., Fowle, M. Z., & Reina, V .J. (2025), Annual Research Review: Cash transfer programs and young people's mental health – a review of studies in the United States. *J Child Psychol Psychiatr, 66*, 498-515. https://doi.org/10.1111/jcpp.14101

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2009). The Patient Health Questionnaire-8: A concise measure of depressive symptoms. *Journal of Affective Disorders, 114*(1–3), 163–173. https://doi.org/10.1016/j.jad.2008.06.026

Magnuson, K. A., Duncan, G. J., Yoshikawa, H. et al. (2025). Effects of unconditional cash transfers on family processes and wellbeing among mothers with low incomes. *Nature Communications,* 16, 7517. https://doi.org/10.1038/s41467-025-62438-x

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Stratton, P., Bland, J., Janes, E., & Lask, J. (2010). Developing an indicator of family function and a practicable outcome measure for systemic family and couple therapy: The SCORE. *Journal of Family Therapy, 32*(3), 232–258. https://doi.org/10.1111/j.1467-6427.2010.00507.x

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., … Stewart-Brown, S. (2007). The Warwick–Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes, 5*(1), 63. https://doi.org/10.1186/1477-7525-5-63

van Buuren, S. (2018). *Flexible imputation of missing data* (2nd ed.). Boca Raton, FL: CRC Press. https://doi.org/10.1201/9780429492259

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

YOUTH ENDOWMENT FUND

Impetus

Home Office

youthendowmentfund.org.uk

hello@youthendowmentfund.org.uk

@YouthEndowFund

The Youth Endowment Fund Charitable Trust

Registered Charity Number: 1185413