

STATISTICAL ANALYSIS PLAN

# Remedi Restorative Mentors: A Randomised Controlled Trial

University of Birmingham

Principal investigators: Professor Siddhartha  
Bandyopadhyay

**Remedi Restorative Mentors: A  
Randomised  
Control Study**



**Statistical analysis plan**

**Evaluating institution: University of Birmingham**

**Principal investigator(s): Siddhartha Bandyopadhyay**

---

<b>Project title<sup>1</sup></b>	<b>Remedi Restorative Mentors: A Randomised Controlled Trial</b>
<b>Developer (Institution)</b>	Remedi
<b>Evaluator (Institution)</b>	University of Birmingham
<b>Principal investigator(s)</b>	Professor Siddhartha Bandyopadhyay
<b>SAP author(s)</b>	Siddhartha Bandyopadhyay, Emily Evans, Yiannis Karavias, Livia Menezes
<b>Trial design</b>	Two-armed randomised controlled trial with random allocation at the individual level
<b>Trial type</b>	Efficacy
<b>Evaluation setting</b>	Community and family homes
<b>Target group</b>	Children and Young People (CYP) aged 10-17 who have displayed violent behaviours

---

	and/or have committed a violent offence referred to Remediation via the police and youth justice services.
<b>Number of participants</b>	352 Treatment group – 176 Control Group – 176
<b>Primary outcome and data source</b>	The primary outcome of interest in this study will be contact with the police measured by the number of contacts of the CYP with Greater Manchester Police (GMP), as perpetrators, victims or missing person episodes. These data will be collected one month before the trial ends. Data will be taken for a period 1 year prior to the recruitment date and 3 months after delivery ends for all CYP.
<b>Secondary outcome and data source</b>	<p>The secondary outcomes will be:</p> <ol style="list-style-type: none"> <li>1. Self-reported offending/delinquency measured through the Self-Report Delinquency Scale (SRDS)</li> <li>2. Emotional and behavioural difficulties measured through Strengths and Difficulties Questionnaire (SDQ)</li> </ol> <p>Self-reported delinquency is measured by the variables:</p> <ul style="list-style-type: none"> <li>• Variety of delinquency</li> <li>• Volume of delinquency</li> </ul> <p>Emotional and behavioural difficulties are measured by the following three variables:</p> <ul style="list-style-type: none"> <li>• Internalising score</li> <li>• Externalising score</li> <li>• Total difficulties score</li> </ul> <p>The SDQ and SRDS questionnaires are completed by CYP at baseline, the end of the RM and RC interventions, and 6 months after the end of the intervention as a follow-up.</p> <p>For the SDQ CYP are asked to recall over the past 6 months, a period defined in the measure. For the SRDS, CYP are asked</p>

	to recall over the past 3 months, this was a period defined for this study, in consultation with Remedi.
--	--

## SAP version history

Version	Date	Changes made and reason for revision
<b>1.3 [latest]</b>	6/2/2024	Feedback from YEF
<b>1.2</b>	26/7/2023	Feedback from YEF
<b>1.1</b>	03/07/2023	Feedback from YEF
<b>1.0 [original]</b>	15/5/2023	<i>[leave blank for the original version]</i>

*Any changes to the design or methods need to be discussed with the YEF Evaluation Manager and the developer team prior to any change(s) being finalised. Describe in the table above any agreed changes made to the evaluation design. Please ensure that these changes are also reflected in the SAP (CONSORT 3b, 6b).*

## Table of contents

SAP version history .....	3
Table of contents .....	3
Introduction .....	4
Design overview .....	5
Sample size calculations overview .....	7
Analysis .....	8

## Introduction

The aim of this study is to evaluate whether a restorative mentoring programme for children and young people (CYP) is useful as a means of diversion from the criminal justice system. The new programme, termed Restorative Mentoring (RM), is compared to an alternative, light touch mentoring scheme termed Restorative Choices (RC). The RC scheme can be seen as a “basic” option, as it is similar to other typical programmes offered in the region. Both RM and RC are delivered by Remedi. Both interventions are described in the study protocol and in the pilot study report. CYP are referred to Remedi by the Greater Manchester Police and by Youth Justice Services.

The trial will be a two-armed individually randomised controlled trial. It will include an impact and a process and implementation evaluation. Upon referral to Remedi, eligible CYP will be randomly assigned to receive restorative mentoring (RM, the treatment group) or restorative choices training (RC, the control group) on a 1:1 basis. The primary outcome of interest in this study will be contact with police, where a CYP is a suspect including arrests and charges but also being linked to offences on which no further action is taken. The secondary outcomes will be the CYP's emotional and behavioural difficulties (as assessed by the Strengths and Difficulties Questionnaire) and their self-reported delinquency (as reported in the Self-Report Delinquency Scale (SRDS)). The process and implementation evaluation will consider the delivery of the Remedi interventions and the experience of them by all relevant groups. This design is considered by the research team to be the most appropriate to understand how the RM intervention operates and to assess its impact.

This design was used during a one-year pilot trial. No amendments have been made to the intervention or the study design as a result of this trial however the study did provide learning regarding the implementation of the study protocol which will support the current study.

The RM intervention began operation in Greater Manchester in April 2022. The RM intervention consists of 12 weeks of support comprised of three components, intensive one to one mentoring support for CYP based on 3-4 sessions per week, support for their family to address conflicts/improve communication and relationships, and restorative justice with relevant and appropriate victims of offences or incidents. All CYP in the treatment group will receive the mentoring component but the use of either of the other components will be determined by an initial needs assessment. The CYP in the control group will receive RC, a short mentoring scheme focused on the CYP understanding the causes and effects of their actions. This consists of four sessions usually lasting 1-2 hours (depending on the attention abilities of participants). The sessions will take place over a period dictated by the availability of the CYP - they can all take place during a week or at most over four weeks. Both RC and RM will be delivered by Remedi. The mentors are given extensive training to ensure that CYP are appropriately supported in both programmes and the evaluation teams will regularly

carry out fidelity checks that look at whether the two interventions are being delivered as designed by looking through the number of sessions delivered and random sampling of case worker sessions. This will pick up any hidden bias in how delivery occurs across the two groups.

## Design overview

Please ensure all details are in line with the latest version of the protocol.

<b>Trial design, including number of arms</b>		Two-arm Randomised Control Study
<b>Unit of randomisation</b>		Individual CYP
<b>Primary outcome</b>	variable	The primary outcome of interest in this study will be contact with the police, as perpetrators, victims or missing person episodes.
	measure (instrument, scale, source)	Police contact data will be taken from Greater Manchester Police administrative records. Police contact data is a count variable starting at 0. These data will be collected one month before the trial ends. Data will be taken for a period 1 year prior to the recruitment date and 3 months after delivery ends for all CYP.
<b>Secondary outcome(s)</b>	variable(s)	<p>The secondary outcomes will be the CYP's emotional and behavioural difficulties and self-reported delinquency. Emotional and behavioural difficulties are measured by the following three variables:</p> <ol style="list-style-type: none"> <li>1) Internalizing score</li> <li>2) Externalizing score</li> <li>3) Total difficulties</li> </ol> <p>Self-reported delinquency is measured by:</p> <ol style="list-style-type: none"> <li>4) Variety of delinquency</li> <li>5) Volume of delinquency</li> </ol>

	<p>measure(s) (instrument, scale, source)</p>	<p>The first three scores are measured with the Strengths and Difficulties Questionnaire (SDQ). The remaining two scores (4 &amp; 5) are measured with the Self-Reported Delinquency Scale (SRDS).</p> <p>The SDQ and SRDS are completed by CYP at the point of consent, end of the intervention and 6 months after the end of the intervention as a follow-up (for CYP who start the interventions in the first 18 months of the trial).</p> <p>For the SDQ CYP are asked to recall over the past 6 months, a period defined in the measure. For the SRDS, CYP are asked to recall over the past 3 months, this was a period defined for this study, in consultation with Remedi.</p> <p>Scales are 0-20 for the internalizing and externalizing scores and 0-40 for the total difficulties score. Variety of delinquency ranges from 0-19 and volume of delinquency ranges from 0-198 (this excludes the scoring for Question 16, which is different from the other questions).</p> <p>In the pilot study 73% of CYP consented into the interventions completed the initial questionnaires, and between one third and one half completed the end point questionnaires. We would expect these rates to be similar for this efficacy study and so could expect to receive around 257 initial questionnaires and 148 end point questionnaires. Six-month follow up questionnaires were not completed as planned during the pilot study and so we cannot use these completion rates to estimate the expected number of completed questionnaires during the efficacy study. However, we would expect completion rates to be lower than those at the end of the interventions, due to the break in contact with the CYP. As such we estimate between 10-20% of CYP will complete these questionnaires, which would equate to between 35-70 questionnaires. The expected attrition rate will affect the power of the analysis for these secondary outcomes.</p>
<p>Baseline for primary outcome</p>	<p>variable</p>	<p>The primary outcome of interest is the number of contacts of the CYP with the police, as perpetrators, victims or missing episodes.</p>
	<p>measure (instrument, scale, source)</p>	<p>Police contact data will be taken from Greater Manchester Police administrative records for events 1 year prior to the intervention. Police contact data is a discrete count variable starting at 0.</p>

Baseline for secondary outcome	variable	<p>The secondary outcomes will be the CYP's emotional and behavioural difficulties and self-reported delinquency.</p> <p>Emotional and behavioural difficulties are measured by the following three variables:</p> <ol style="list-style-type: none"> <li>1) Internalizing score</li> <li>2) Externalizing score</li> <li>3) Total difficulties</li> </ol> <p>Self-reported delinquency is measured by:</p> <ol style="list-style-type: none"> <li>4) Variety of delinquency</li> <li>5) Volume of delinquency</li> </ol>
	measure (instrument, scale, source)	<p>The first three scores are measured with the Strengths and Difficulties Questionnaire (SDQ) and the last two (4 &amp; 5) are measured with the Self-Reported Delinquency Scale (SRDS) questionnaire.</p> <p>The SDQ and SRDS are completed by CYP at the start of the intervention.</p> <p>Scales are 0-20 for the internalizing and externalizing scores and 0-40 for the total difficulties score.</p> <p>Variety of delinquency ranges from 0-19 and volume of delinquency ranges from 0-198 (this excludes the scoring for Question 16, which is different from the other questions).</p>

### Sample size calculations overview

	Protocol	Randomisation
Minimum Detectable Effect Size (MDES)	0.3	0.3
Alpha	0.05	0.05
Power	0.8	0.8



		Protocol	Randomisation
One-sided or two-sided?		Two-sided	Two-sided
Number of participants	intervention	176	176
	control	176	176
	total	352	352

The planned number of the trial participants is 352 CYP in its one year of implementation, 176 in the RM group and 176 in the RC group. The sample size is such that the trial is sufficiently powered to detect a Cohen’s d of at least 0.3 with probability 80%, which is in line with the funder guidelines and the previous literature.

The “simple” randomisation method will be used, which is a robust method against selection and accidental biases. We will use the statistical software package Matlab to implement the randomisation. Automated randomisation will ensure that the process is transparent and reproducible.

The above rest on the assumption that the alpha is equal to 0.05 and that the power is equal to 0.8. The minimum detectable effect size is equal to 0.3 as measured by Cohen’s d.

## Analysis

Descriptive statistics will be presented for the overall sample, as well as for the treatment and control groups separately. The random assignment of cases will be formally tested by comparing means of observable characteristics (such as age, sex, ethnicity among others) between treatment and control groups to check that they are similar across treatment and control groups (see discussion below on imbalance at baseline).

### Primary outcome analysis

The aim of the analysis is to answer the research question: did the Restorative Mentoring intervention (treatment) reduce post-treatment contact with the police compared to the Restorative Choices intervention (control)?

The primary outcome for this study is contact with the police, defined as the number of times the CYP appear in the police data as perpetrators, victims or for missing person episodes. These data will be taken from Greater Manchester Police administrative records.

Contact with the police is a discrete count variable starting at 0. These data will be collected one month before the trial ends, for a period 1 year prior to the recruitment date and for 3 months after completion of the intervention for all CYP.

It should be noted that the primary outcome, which is police contact with the CYP, is susceptible to biases in reporting and recording of crimes. These biases may stem from a number of factors, including characteristics of both the victim and the offender, such as ethnicity and gender. Nevertheless, we anticipate that the RCT design remains robust to these biases, as randomisation into the treatment and control groups is conducted independently of the police's reporting and recording of crimes.

There are two ways of analysing the data, depending on the distribution assumed for the dependent variable, which is total police contacts. As a discrete count random variable, it may be assumed to follow a Poisson or Negative Binomial distribution. However, if the mean of that distribution is far from 0, the Poisson and Negative Binomial distributions may be approximated by the continuous normal distribution. It is not clear a priori which distribution will be the better approximation, and therefore we consider below both scenarios. In any case, we do not expect major differences; the issue of whether a variable should be treated as continuous or discrete has received much attention in the randomized control trial literature and the main finding is that unless the outcome is binary or it is a "time to a first event", information will not be lost when treating a count variable as a continuous one, see, e.g. Herbison et al. (2015).

For the analysis, the headline estimate will be derived from an intention to treat effect (ITT), denoted by  $\beta$  in the regression:

$$g[E(Y_i|T_i)] = \alpha + \beta T_i,$$

where  $Y_i$  is the primary outcome for individual  $i$ , and  $T_i$  is an indicator variable equal to 1 if person  $i$  was assigned to treatment and 0 if person  $i$  was assigned to the control group. The function  $g$  is the link function and is equal to the natural logarithm function if the dependent variable is assumed to be Poisson or Negative Binomial.

Hypothesis testing in the above regression complies with the power predictions of the study. The effect size will be measured by the Incidence Rate Ratio if the data are assumed to be Poisson or Negative Binomial, and by Cohen's  $d$  if they are assumed approximately normal. If the data are Poisson/Negative Binomially distributed, the necessary sample size to detect a "small" effect of at least 1.22, see e.g. Cohen (1992) and Olivier et al. (2017), with  $\alpha=0.05$  and

80% power is 322 CYP equally split between the treatment and control groups.<sup>2</sup> Alternatively, if the data are assumed to be normally distributed, then to detect a “small” effect size of at least 0.3 is 352 CYP. Therefore, a sample of 352 CYP will lead to as sufficiently powered study irrespective of the data distributional assumption.

We will additionally estimate a regression with covariates:

$$g[E(Y_i|T_i, Z_i)] = \alpha + \beta T_i + Z_i' \gamma,$$

where  $Z_i$  is a vector of individual characteristics that are not affected by the intervention (sex, age, ethnicity, district, the number of contacts with the police in the year prior to referral, and referral source). The analysis below depends on the over-arching assumption that the sample of observations are independent to each other. This is reasonable as the sample is only a small subset of the whole population.

The dependent variable here is a count variable and it is highly likely that the lower values will have much higher frequency, in other words it is heavily skewed to the right. This feature of the data suggests that they are best analysed using Poisson regression. Mathematically, the Poisson regression takes the form:

$$Y_i | \lambda_i \sim Poi(\lambda_i)$$

where  $\lambda_i = \alpha + \beta T_i + Z_i' \gamma$ .

An additional assumption that needs to hold for the Poisson regression is that conditionally on the regressors, the mean and variance of  $Y_i$  must be equal to each other. This is called equidispersion. If this assumption does not hold and the variance is greater than the mean, then we must estimate the model using the Negative Binomial regression, which is able to deal with what is called overdispersion.

The Negative Binomial regression works for the same type of data as the Poisson regression, however it also contains an additional term which captures the excess variance, as typically the variance is greater than the mean, when they are not equal. The Negative Binomial distribution is a Poisson-gamma mixture, where a gamma noise variable which has a mean of 1 and a scale parameter of  $\nu$  has been included resulting in the following distribution:

---

<sup>2</sup> This calculation is based on the average number of police contacts for the control group in the pilot study, which is 1.8. Given that the pilot study participants did not have a full six-month follow-up it is expected that in the efficacy study this average will be much higher. However, the higher the average, the fewer observations are necessary for the power calculations. Therefore, 322 observations is a conservative estimate given that even fewer may be required.

$$P(Y = y_i | \mu_i, a) = \frac{\Gamma(y_i + a^{-1})}{\Gamma(y_i + 1)\Gamma(a^{-1})} \left( \frac{a^{-1}}{a^{-1} + \mu_i} \right)^{a^{-1}} \left( \frac{\mu_i}{a^{-1} + \mu_i} \right)^{y_i}$$

Where  $\mu_i = t_i\mu$  and  $a = 1/v$ . The parameter  $\mu$  may be interpreted as the risk of a new occurrence of the event during a specified exposure period  $t$ .

If the dependent variable can be approximated by the continuous normal distribution, then the regression models above become linear and will be estimated by least squares. In this case, the coefficient  $\beta$  is numerically equal to the difference of means between the control and the treatment group.

Finally, as it is expected that compliance will not be perfect, the analysis will involve estimating a local average treatment effect (LATE), which will provide an estimate of the treatment effect for individuals who engage with the intervention ([Imbens & Angrist \(1994\)](#), [Imbens & Wooldridge \(2009\)](#)). In this case, treatment assignment will be used as an instrument for whether individuals get treated or not (see Compliance section).

### **Secondary outcome analysis**

In this part of the analysis, the aim is to answer the research question: did the Restorative Mentoring intervention (treatment) reduce difficulties and self-reported behaviour measures according to the SDQ and SRDS questionnaires, when compared to the Restorative Choices intervention (control). Furthermore, do these reductions differ by sex, ethnicity, and referral source.

The secondary outcomes for this analysis will include the CYP's emotional and behavioural difficulties and self-reported delinquency.

Emotional and behavioural difficulties are measured by the following three variables:

- 1) Internalizing score – measured on a scale from 0-20
- 2) Externalizing score – measured on a scale from 0-20
- 3) Total difficulties score – measured on a scale from 0-40

Self-reported delinquency is measured by:

4) Variety of delinquency - measured as the sum of the overall categories of delinquency  
Range: 0-19.

5) Volume of delinquency - measured as the sum of magnitudes of delinquency, over all  
categories Range: 0-198.

These variables will be compared across the treatment and control groups to test the null hypothesis of equal means against the two-sided alternative.

The regression analysis takes the form

$$Y_i = \alpha + \beta T_i + Z_i' \gamma + u_i \quad (2)$$

where  $Y_i$  is the secondary outcome for individual  $i$  (one of the aforementioned SDQ and SRDS variables).  $T_i$  is an indicator variable equal to 1 if  $i$  was assigned to treatment and 0 if  $i$  was assigned to the control group.  $Z_i$  is a vector of individual characteristics that are not affected by the intervention (sex, age, ethnicity, district, prior contact with the police, and referral source).  $u_i$  represents the random errors, which are not observed. Given that  $Y_i$  is a continuous variable we will employ a linear regression. The key assumptions that we will use is that the observations are independent to each other, there is mean independence  $E(u_i|T_i, Z_i) = 0$ , and that not all CYP have identical  $T_i$  and  $Z_i$  (no perfect multicollinearity). These assumptions are weak and easily satisfied.

This part of the analysis will be exploratory as we will only have outcome measures for CYP who complete the survey at the end of the treatment. Therefore, we will obtain an estimate of the treatment's effect for those who finish the treatment. Additionally, it is important to acknowledge potential limitations in terms of statistical power since the sample size in this case will be smaller.

### **Subgroup analyses**

Given the potential for treatment to have different effects on individuals or specific groups, tests for heterogenous treatment effects will be performed to analyse whether individuals respond differently to treatment based on three observable characteristics: sex, ethnicity, and referral source.<sup>3</sup> In this case, we will estimate the regressions specified above for subsamples based on these characteristics (male, female, white, non-white, police, youth services). Although the individual randomisation process was not stratified based on these

---

<sup>3</sup> The subgroup analysis will not be done in interactions of subgroups due to sample size limitations.

groups, if the sample size is sufficiently large to do a powered analysis, reporting results for groups defined post-randomisation could provide supplementary insights alongside the primary findings (Duflo et al. 2007). Therefore, this part of the analysis will also be exploratory, and we acknowledge potential limitations in terms of statistical power, since the sample size for each subgroup will be small.

### **Interim analyses and stopping rules**

No such analysis is planned as it will be underpowered. The pilot study gave sufficient evidence for the safety of the intervention.

### **Longitudinal follow-up analyses**

One follow up point for measurement is specified – 6 months after the end of the intervention for the secondary outcomes. This will enable the estimation of potential longer lasting effects of the intervention by performing the same statistical analysis described above with data covering a longer period of time.

### **Imbalance at baseline**

Balance tests will be performed to formally test the random assignment of CYP to treatment and control groups. We will perform t-tests to evaluate whether there is significant differences in mean characteristics between treatment and control groups. The baseline characteristics we will test are: sex, age, ethnicity, district, the number of contacts with the police in the year prior to referral, and referral source.

### **Missing data**

The primary outcome analysis will be performed according to the ITT principle (described above). The primary outcome is measured by police administrative data and therefore outcome variables for CYP that have disengaged with the intervention will still be collected, reducing concerns of missing values for the primary outcome variable. For the control variables that might be missing, we will follow YEF guidelines and in a first step we will attempt to establish the mechanism behind missingness using a logistic regression model. Depending on the pattern of missing data, either multiple imputation or sensitivity analyses will be conducted.

### **Compliance**

The evaluation has been designed to engage CYP in both treatment and control groups. Implementation of the intervention will be monitored to make sure that compliance is near

perfect by introducing a fidelity check agreement. However, it is expected that some CYP might disengage, justifying the ITT and LATE analysis proposed in the analysis section above.

The ITT approach analyses CYP based on their original group assignments, irrespective of whether they completed the treatment or not. The LATE approach estimates the effect of the treatment exclusively for CYP who comply with their assigned treatment, defined as (J-PAL, 2023):

$$LATE = \frac{E(Y_i|z_i = 1) - E(Y_i|z_i = 0)}{E(d_i|z_i = 1) - E(d_i|z_i = 0)} \quad (3)$$

Where  $Y$  is the outcome for CYP  $i$  (as specified above);  $z$  represents the treatment assignment, with a value of 1 if the CYP was assigned to treatment (RM) and 0 otherwise;  $d$  represents whether the treatment was received, equal to 1 if the CYP engaged with treatment and 0 otherwise. In this equation, random treatment assignment serves as an instrumental variable for the status of treatment. As we have available data on the number of sessions received by each CYP we will define  $d_i = 1$  if the CYP completes the full intervention.<sup>4</sup>

The LATE estimation relies on a few assumptions. The first is the independence assumption that is automatically satisfied because of random assignment of the treatment. The estimation also requires a positive fraction of compliers and that the monotonicity (being assigned to treatment does not reduce the likelihood of being treated) and exclusion restriction assumptions (CYP respond to the treatment directly, rather than just to their assignment to receive it, therefore treatment assignment does not change the outcome) hold (J-PAL, 2023). In addition, we assume that compliance does not vary with treatment assignment. The experimental design, which ensures perfect blindness, guarantees that this assumption holds true. In other words, CYP who consent to participate in the intervention do not know whether they are going to receive RM or RC (Gerber, Green, Kaplan and Kern, 2010; EGAP).

### **Presentation of outcomes**

The effect size will be reported based on the Incidence Rate Ratio or Cohen's  $d$  formulas. The incidence rate ratio (IRR) is given by:

---

<sup>4</sup> This means that the CYP is classified by REMEDI as a closed case, which indicates that the CYP has completed all the sessions (occasionally 1 or 2 sessions under or over the specified amount). While there is no agreed definition of compliance, we follow this strict definition after discussion with the intervention provider.

$$\text{IRR} = \exp(\hat{\beta}),$$

where  $\hat{\beta}$  is the maximum likelihood estimate from the Poisson and Negative Binomial distributions. The confidence interval is estimated by applying the delta method to the inverse of the estimated Fisher Information matrix.

Cohen's d is equal to the difference in the average of the variable in the treatment and control group divided by the standard deviation of the pooled sample:

Cohen's

$$d = \frac{\bar{Y}_t - \bar{Y}_c}{\text{s.e.}(Y)}$$

The confidence interval for d is given by

$$[d - 1.96 \times \sigma(d), d + 1.96 \times \sigma(d)]$$

where

$$\sigma(d) = \sqrt{\frac{N_1 + N_2}{N_1 N_2} + \frac{d^2}{2(N_1 + N_2)}}.$$

The t-tests and confidence interval formulas are based on the normal distribution as the sample size is expected to be large. These will be used to examine the statistical significance of the estimated coefficients. The formula for t-statistic is given by

$$t = \frac{\hat{\beta}}{\text{s.e.}(\hat{\beta})}.$$



## References

Cohen, J. (1992). A power primer. *Psychol. Bull.* 112:155–159.

Duflo, E., Glennerster, R., & Kremer, M. (2006). "Using Randomization in Development Economics Research: A Toolkit" (Technical Working Paper No. 0333). Retrieved from National Bureau of Economic Research website: <https://doi.org/10.3386/t0333>

Evidence in Governance and Politics (EGAP) i. (n.d.). 10 Things to Know About the Local Average Treatment Effect. Retrieved April 23 from <https://egap.org/resource/10-things-to-know-about-the-local-average-treatment-effect/>.

Gerber, A. S., Green, D. P., Kaplan, E. H., & Kern, H. L. (2010). Baseline, Placebo, and Treatment: Efficient Estimation for Three-Group Experiments. *Political Analysis*, 18(3), 297-315. Hilbe, J.M. (2014) Modeling Count Data. Cambridge: Cambridge University Press.

Herbison, P., Robertson, M.C. & McKenzie, J.E. Do alternative methods for analysing count data produce similar estimates? Implications for meta-analyses. *Syst Rev* 4, 163 (2015).

Imbens, G. W., & Angrist, J. D. (1994). "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, 62(2), 467-475.

Imbens, G. W., & Wooldridge, J. M. (2009). "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature*, 47(1), 5-86.

J-PAL. (2023). Research resources. Abdul Latif Jameel Poverty Action Lab. Retrieved April 23, 2024, from <https://www.povertyactionlab.org/resource/data-analysis>.

J-PAL. (2021). Research resources. Abdul Latif Jameel Poverty Action Lab. Retrieved January 20, 2024, from <https://www.povertyactionlab.org/resource/implementation-monitoring>.

Olivier, J., May, W.L., & M.L. Bell (2017) Relative effect sizes for measures of risk, *Communications in Statistics - Theory and Methods*, 46:14, 6774-6781



[youthendowmentfund.org.uk](https://youthendowmentfund.org.uk)



[hello@youthendowmentfund.org.uk](mailto:hello@youthendowmentfund.org.uk)



[@YouthEndowFund](https://twitter.com/YouthEndowFund)

The Youth Endowment Fund Charitable Trust

Registered Charity Number: 1185413

---