Effect Size Calculation for experimental & quasi-experimental methods

Jorge Garcia Hombrados

&

Hugh Waddington

Our goal is to get from here…

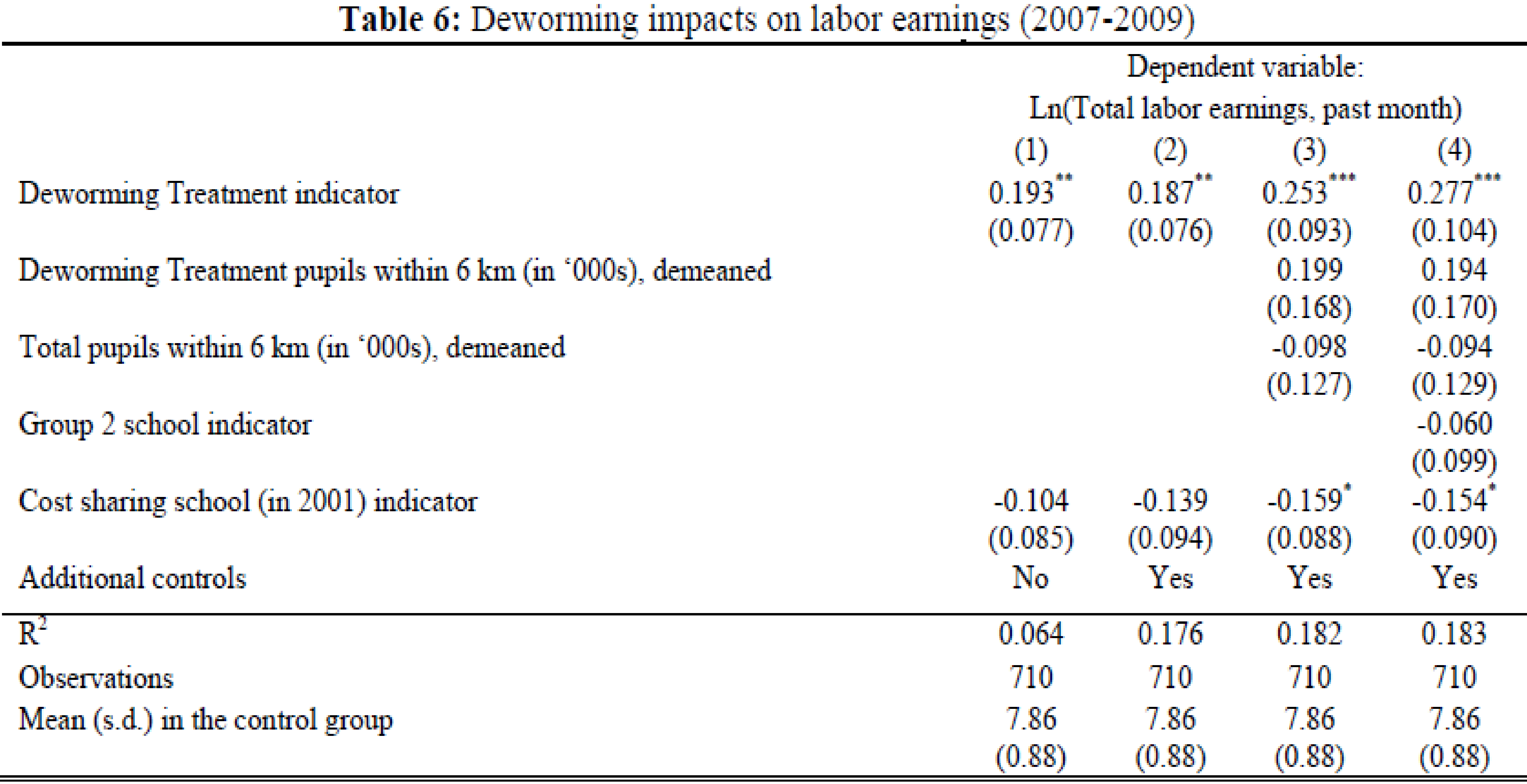
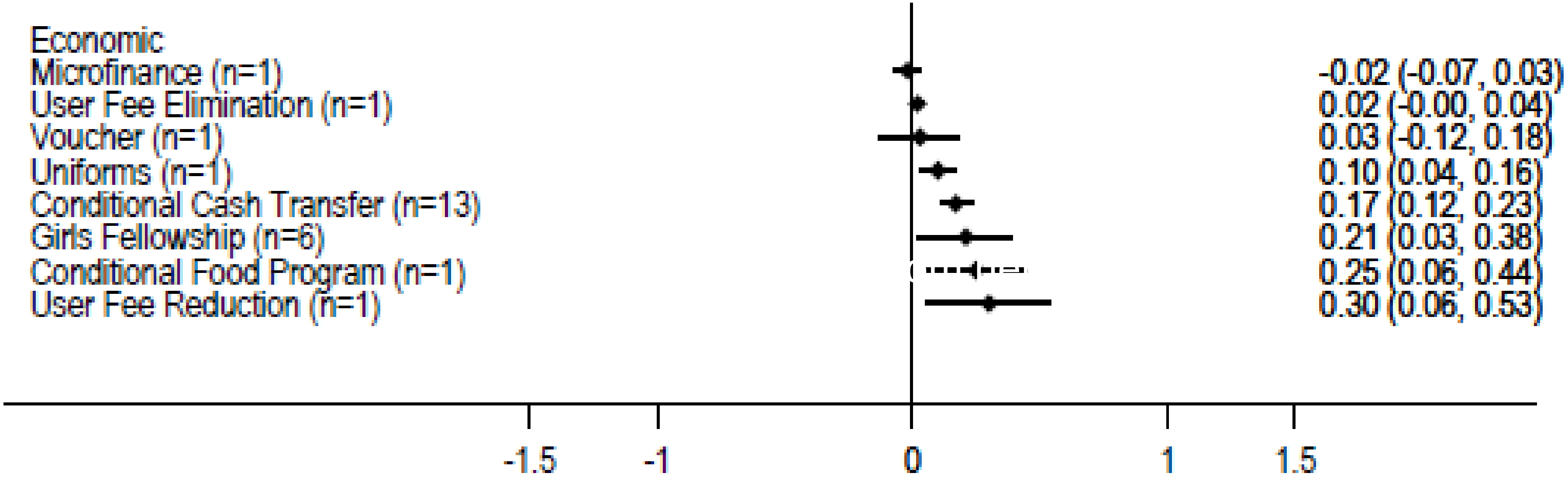


Table 6: Deworming impacts on labor earnings (2007-2009)

	Dependent variable: Ln(Total labor earnings, past month)			
	(1)	(2)	(3)	(4)
Deworming Treatment indicator	0.193**	0.187**	0.253***	0.277***
	(0.077)	(0.076)	(0.093)	(0.104)
Deworming Treatment pupils within 6 km (in ‘000s), demeaned			0.199	0.194
			(0.168)	(0.170)
Total pupils within 6 km (in ‘000s), demeaned			-0.098	-0.094
			(0.127)	(0.129)
Group 2 school indicator				-0.060
				(0.099)
Cost sharing school (in 2001) indicator	-0.104	-0.139	-0.159*	-0.154*
	(0.085)	(0.094)	(0.088)	(0.090)
Additional controls	No	Yes	Yes	Yes
R^2	0.064	0.176	0.182	0.183
Observations	710	710	710	710
Mean (s.d.) in the control group	7.86	7.86	7.86	7.86
	(0.88)	(0.88)	(0.88)	(0.88)

Source: Baird et al. 2011 Worms at work

... to here



- Source: Petrosino et al. (2012) 'Interventions in Developing Nations for Improving Primary and Secondary School Enrollment of Children: A Systematic Review' Campbell Systematic Reviews

Working in small groups, you will learn:

- To calculate effect sizes: odds ratios, risk ratios, response ratios, standardized mean differences (SMDs)

- Effect sizes for different types of data
  - based on data reported as group comparisons from raw and adjusted data
  - based on data reported from statistical inference (eg regression analysis)

- For different types of rigorous evaluation design:
  - randomized control trials (RCTs),
  - cross-section design with Propensity Score Matching (PSM),
  - longitudinal design with Difference-in-Differences (DID) analysis

Concepts we will use liberally

- Mean
- Standard deviation
- Standard error
- Confidence interval
- Average treatment effect on the treated (ATET)

What is an effect size?

- Impact evaluation tells us ‘what difference’ an intervention makes to outcomes, measured by the treatment effect (eg regression coefficient on treatment variable)

- The effect size scales the treatment effect in units which tell us the <u>magnitude</u> of this difference and its statistical significance (as indicated by 95% confidence interval)

- There are different units and measures in which we can quantify this difference but it needs to be <u>consistent</u> and <u>comparable</u> across studies.

- Effect sizes are the unit of analysis in meta-analysis but they are also important to indicate policy relevance of findings

Example: Height-for-age

	Mean Z-score	Percentage below z=-2
2004 Bangla-DHS	-2.0	51%
2011 Bangla-DHS	-1.7	41%

Examples of different effect sizes

- Standardized mean difference
  - Group contrast
  - Continuous outcome variable e.g. test scores

- Odds ratio/ risk ratio
  - Group contrast
  - Dichotomous outcome variable e.g. mortality

- Correlation coefficient (pearson's r)
  - Association between 2 variables e.g. value of micro-credit loan and income

- Proportion
  - Disease prevalence rates e.g. proportion suffering diarrhoea

Effect sizes measures

For continuous outcomes:

- Standardised Mean Differences (SMD)
- Response Ratios (RR)
- Regression coefficients, t-stat

d-based effect sizes

- R-based effect sizes (correlation coefficients)

For dichotomous outcomes:

- Odds ratios
- Risk ratios

Standardized mean difference

$$SMD = \frac{Y_t - Y_c}{S_p}$$

$$S_p = \sqrt{\frac{(n_t - 1) * S_t^2 + (n_c - 1) * S_c^2}{n_t + n_c - 2}}$$

- Uses the pooled standard deviation (some cases use control group standard deviation)

Exercise: calculate absolute and SMD

Income per capita	Treatment mean	Comparison mean	Treatment SD	Comparison SD	Sample size t	Sample size c
Bangladesh	770	500	300	200	50	50
Nepal	550	280	200	150	50	50

Odds ratio and risk ratio

	Frequencies	
	Success	Failure
Treatment Group	*a*	*b*
Control Group	*c*	*d*

$$\overline{OR} = \frac{a/b}{c/d} = \frac{ad}{bc}$$

- Ratio of odds of success in the treatment group relative to odds of success in the comparison

$$\overline{RR} = \frac{a/(a+b)}{c/(c+d)}$$

- Ratio of probability of success in the treatment group relative to probability of success in the comparison group

Exercise: calculate and interpret OR and RR

	Poor	Non-Poor
Treatment	315	685
Comparison	400	600

Selecting the effect size measure in practice

- We need to estimate the same effect size measure across all the included studies pooled in the meta-analysis

- However, very often not all the studies report enough information to compute all the effect size measures and contacting authors does not always work

- The selection of the effect size measure in practice takes into account:
  - Nature of outcome (dichotomous vs continuous)
  - Minimising the number of studies lost

Option 1: SMD

- Measures the impact of the programme in standard deviations of the outcome variable
- Can be computed consistently for both experimental and non-experimental studies
- Is the “less problematic” methodology
- However, its interpretation is not straightforward and the data required for its computation is not always available

Option 2: Response Ratio

- Measures the impact of the programme in percentage change
- Based on the Risk Ratio effect size used for dichotomous outcomes
- Data required for its computation is minimum
- Can be computed consistently for both experimental and non-experimental studies
- It is appropriate for both continuous and dichotomous outcomes (risk ratio) providing the outcome measure has a natural scale unit and natural zero points (but is not likely to equal zero)
- Synthesis uses logarithmic scales for both RR and SE(RR)

Other measures are problematic

- T-statistics are noisy and when applied to regression has some important shortcomings (Becker & Wu, 2007)
- Regression coefficients: data on covariance matrix (hardly reported) are required for appropriate synthesis
- R-based effect sizes do not seem to work properly for multivariate regression

Computation of effect sizes: SMD

- Standardised Mean Difference (SMD):

$$SMD = \frac{Y_t - Y_c}{S_p} \qquad SE = \sqrt{\frac{n_t + n_c}{n_c * n_t} + \frac{SMD^2}{2 * (n_c + n_t)}}$$

- How to estimate SMD from different study designs?
- An easy way of dealing with it is thinking separately in the numerator and denominator.

Computation of effect size: SMD

- The numerator Y_t-Y_c represents the causal raw impact of the programme in the outcome:
  - In a regression analysis is the coefficient of interest (Beta).
  - In a matching-based study this is the causal impact (ATT) or the difference in outcomes between groups after matching.

Computation of effect size: SMD

- The denominator S_p is a measure of the standard deviation of the outcome. It makes the effect size comparable across studies
  - In regression studies, we can use the standard deviation of the regression errors.
  - Alternatively, we can use the sample standard deviation or the treatment and control standard deviation to calculate S_p in matched-based studies or approximate it in regression based studies:

$$S_p = \sqrt{\frac{\left(\left(SD_y^2\right) * \left(n_t + n_c - 2\right)\right) - \left(\frac{\beta^2 * \left(n_t * n_c\right)}{n_t + n_c}\right)}{n_t + n_c}}$$

$$S_p = \sqrt{\frac{\left(n_t - 1\right) * S_t^2 + \left(n_c - 1\right) * S_c^2}{n_t + n_c - 2}}$$

Computation of effect size: SMD

- Diff-in-Diff models
  - In Diff-in-Diff model, the dependent variable is the change in the dependent variable of reference, or in semi log models, the growth rate of the variable of reference.
  - In these cases, Sp should measure the pool standard deviation of the change (or the growth rate) in the dependent variable of reference.
  - Unfortunately, this information is not reported very often and most of the times the calculation of SMDs requires assumptions.

Computation of effect size: SMD

- Tobit, Logit and Probit models

$$SMD = \frac{\Delta}{S_p}$$

- Where for:
  - Probit models $\Delta = \Phi\ (x_i\beta+\gamma) - \Phi(x_i\beta)$
  - Censored Tobit models

  $\Delta=((\Phi\ ((x_i\beta+\gamma)/\sigma))^*((\ x_i\beta+\gamma)+\ \sigma\ \varphi(x_i\beta+\gamma)))\ -\ ((\Phi\ ((x_i\beta)/\sigma))^*((\ x_i\beta)+\ \sigma\ \varphi(x_i\beta))$

  - Logit models

$$\Delta = \frac{\exp\ (xi\beta + \gamma)}{1 +\ \exp\ (xi\beta + \gamma)} - \frac{\exp\ (xi\beta)}{1 +\ \exp\ (xi\beta)}$$

SMD correction for small sample bias

- When the sample size is small, a correction in the effect size and its variance is needed.
- Although the correction is going to be almost imperceptible, we recommend to apply to all SMD calculations.
- For regression studies, there is a more efficient alternative (see Keef and Roberts, 2004)

$$SMD_{corrected} = SMD_{uncorrected} * \left[1 - \frac{3}{4 * (n_t + n_c - 2) - 1}\right]$$

$$SE_{corrected} = \sqrt{\frac{n_t + n_c}{n_t * n_c} + \frac{(SMD_{corrected})^2}{2 * (n_t + n_c)}}$$

Computation of effect size: Response Ratios

$$\mathrm{RR} = \frac{\mathrm{Y_t}}{\mathrm{Y_t}} \qquad SE(RR) = S_p^2 * \left(\frac{1}{n_t * Y_t^2} + \frac{1}{n_c * Y_c^2}\right)$$

- For matched-based studies (e.g., PSM), $Y_c = Y_t$ – ATT where Y_t is the outcome level in the treatment, ATT is the average treatment effect on the treated and Y_c is the outcome level in the control group after matching.

- For regression-based studies, $Y_t = Y_c + \beta$ where Y_c is the outcome level in the total sample and Y_t is the “ceteris paribus” average predicted outcome if the sample received the treatment.

Computation of Response Ratio Standard Error

- When the standard deviation of the dependent variable or the necessary information to calculate SD is not reported, we can approximate the SE for response ratios using the t statistics/p-value of the regression coefficient or of the results of the t test for equality of means between groups after matching:

$$SE(RR) = \exp(\frac{Ln\,(RR)}{t})$$

Computation of effect size: Response Ratios

- Semi-log difference-in-differences (DID):

$$RR = e^{\beta}$$

If Sp is computed to calculate the SE (RR), Sp should measure the pooled standard deviation of the change (or the growth rate) in the variable of reference.

Computation of effect size: Response Ratios

- Logit, Probit and Tobit model

$$RR = \frac{Y_c + \Delta}{Yc}$$

- Where for:
  - Probit models $\Delta = \Phi\,(x_i\beta+\gamma) - \Phi(x_i\beta)$
  - Censored Tobit models

  $\Delta=((\Phi\,((x_i\beta+\gamma)/\sigma))*((\,x_i\beta+\gamma)+\,\sigma\,\varphi(x_i\beta+\gamma)))\,-\,((\Phi\,((x_i\beta)/\sigma))*((\,x_i\beta)+\,\sigma\,\varphi(x_i\beta))$
  - Logit models $\Delta = \dfrac{\exp\,(xi\beta+\gamma)}{1+\,\exp\,(xi\beta+\gamma)} - \dfrac{\exp\,(xi\beta)}{1+\,\exp\,(xi\beta)}$

Computation of effect size: Response Ratios

- For logit model we can also estimate RR as:

$$RR = \frac{e^{\beta}}{1 - Y_c + Y_c * e^{\beta}}$$

Unit of analysis error correction

- Unit of analysis error (UoA) arises in impact evaluation studies in which programme placement and analysis are conducted at a different unit level and the researcher does not account for this within-cluster dependency.
- E.g. Programme placement at cluster level and outcomes analysed at household level.
- The consequences of UoA are false smaller variances and false narrower confidence intervals.
- If the study conduct analysis and programme placement at different levels and do not use cluster robust standard errors, we need to apply a correction to the standard errors to avoid potential Type II error:

$$SE_{corrected\ UoA} = SR * \sqrt{1 + (m - 1) * ICC}$$

- Where m is the cluster size and ICC is the intra-cluster correlation coefficient.

Formulae Table for ES computation

Effect size measure	Formulae for matched-based studies:	Information needed to be reported in matched-based studies:	Formulae for regression-based studies:	Information needed to be reported in regression-based studies:
Standardizes Mean Differences (SMD)	$SMD = \frac{Y_t - Y_c}{S_p}$ $SE(SMD) = \sqrt{\frac{n_t + n_c}{n_t * n_c} + \frac{SMD^2}{2 * (n_t + n_c)}}$ OR $SE(SMD) = \frac{SMD}{t}$ $S_p = \sqrt{\frac{(n_t - 1) * S_t^2 + (n_c - 1) * S_c^2}{n_t + n_c - 2}}$ OR $S_p = \sqrt{\frac{\left(SD_y^2 * (n_t + n_c - 1)\right) - \left(\frac{\beta^2 * (n_t * n_c)}{n_t + n_c}\right)}{n_t + n_c}}$	*SMD* -Sample mean outcome for the treated and control group after matching OR -Sample mean outcome for the treatment group AND Average Treatment Effect on the Treated. -Sample standard deviation for treatment and control group AND sample size for treatment and control group OR -Sample standard deviation for the total sample AND sample mean outcome for treatment and control group OR sample mean outcome for treatment group and ATT AND sample size for treatment and control group. *SE (SMD)* -Sample size of the treated and control group OR t statistics of the treatment effect.	$SMD = \frac{\beta}{S_p}$ $S_p = \sqrt{\frac{\left(SD_y^2 * (n_t + n_c - 1)\right) - \left(\frac{\beta^2 * (n_t * n_c)}{n_t + n_c}\right)}{n_t + n_c}}$ OR $S_p = SD\ of\ the\ regression\ residuals.$ $SE(SMD) = \frac{SMD}{t}$ OR $SE(SMD) = \sqrt{\frac{SMD^2}{v - 2} * \left(\frac{v}{t^2} + v[c(v)]^2 - v + 2\right)}$	*SMD* -Regression coefficient -Sample standard deviation of the dependent variable AND sample size for treatment and control group OR -Standard deviation of the error term in the regression. SE(SMD) -t statistics for the regression coefficient OR t statistics for the regression coefficient AND Number of covariates AND sample size for the total sample.

Formulae Table for ES computation

Effect size measure	Formulae for matched-based studies:	Information needed to be reported in matched-based studies:	Formulae for regression-based studies:	Information needed to be reported in regression-based studies:
Response Ratio (RR)	$RR = \frac{Y_t}{Y_c}$ $SE(RR) = \mathrm{Exp}\left(\frac{Ln(RR)}{t}\right)$ OR $SE(RR) = S_p^2 * \left(\frac{1}{n_t * Y_t^2} + \frac{1}{n_c * Y_c^2}\right)$ $S_p = \sqrt{\frac{(n_t - 1) * S_t^2 + (n_c - 1) * S_c^2}{n_t + n_c - 2}}$ OR $S_p = \sqrt{\frac{\left(SD_y^2 * (n_t + n_c - 1)\right) - \left(\frac{\beta^2 * (n_t * n_c)}{n_t + n_c}\right)}{n_t + n_c}}$	*RR* -Sample mean outcome for the treated and control group after matching OR -Sample mean outcome for the treatment group AND Average Treatment Effect on the Treated. *SE(RR)* -t statistics of the treatment effect. OR -Sample mean outcome for the treated and control group after matching OR Sample mean outcome for the treatment group AND Average Treatment Effect on the Treated AND Sample size for the treatment and control group AND sample standard deviation for the treatment and control group. OR sample standard deviation for all the sample.	$RR = \frac{Y_s + \beta}{Y_s}$ $SE(RR) = \mathrm{Exp}\left(\frac{Ln(RR)}{t}\right)$	*RR* -Mean outcome for the total sample. -Beta of the regression coefficient. SE(RR) -t statistics of the regression coefficient.

Definitions

- S_p is the pool standard deviation
- β is the coefficient or impact effect of interest.
- t is the t statistics of the regression coefficient or of the relevant treatment impact (t-test for equality of means).
- *Exp* is the exponential function ($e^{f(x)}$)
- Y_t, Y_c, Y_s, n_t , n_c and $\mathrm{n_s}$ are the mean outcome in the treatment group, control group and total sample and the sample size for the treatment group, control group and total sample.
- $\mathrm{SD_t}$, $\mathrm{SD_c}$ and $\mathrm{SD_Y}$ are the standard deviation for the treatment group, control group and total sample.
- *v is the degrees of freedom of the regression equation.*
- *ATT* Average treatment effect on the treated.
- *m* is the cluster size
- *ICC* is the intra-cluster correlation coefficient, an estimate of the relative variability within clusters.

Definitions

- In a logit, probit or tobit regression:
  - Δ is the impact effect
  - xiβ is the mean predicted outcomes for without participating in the programme.
  - γ is the coefficient of interest .
  - Φ is the cumulative distribution function.
  - φ is the probability distribution function.

Exercises

- Compute effect sizes for the following studies:
  - For *Banerjee et al. 2009*:
    - estimate RR and SMD of the impact of microfinance access on per capita expenditure
    - Estimate RR of the impact of microfinance access on non-food decision making by women
  - For *Chen et al. 2012*, estimate RR and SMD of the impact of tuition relief on the change in test scores (use the Diff-in-Diff multivariate regression specification)

Banerjee et al. 2009

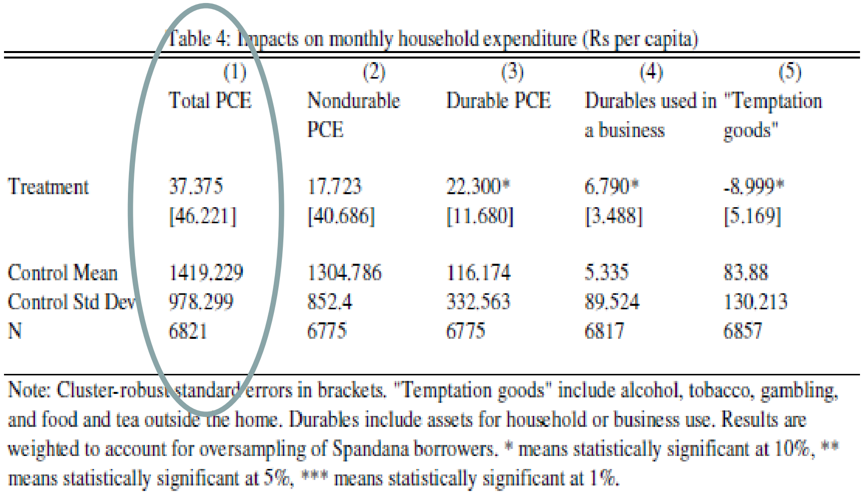


Table 4: Impacts on monthly household expenditure (Rs per capita)

	(1) Total PCE	(2) Nondurable PCE	(3) Durable PCE	(4) Durables used in a business	(5) "Temptation goods"
Treatment	37.375 [46.221]	17.723 [40.686]	22.300* [11.680]	6.790* [3.488]	-8.999* [5.169]
Control Mean	1419.229	1304.786	116.174	5.335	83.88
Control Std Dev	978.299	852.4	332.563	89.524	130.213
N	6821	6775	6775	6817	6857

Note: Cluster-robust standard errors in brackets. "Temptation goods" include alcohol, tobacco, gambling, and food and tea outside the home. Durables include assets for household or business use. Results are weighted to account for oversampling of Spandana borrowers. * means statistically significant at 10%, ** means statistically significant at 5%, *** means statistically significant at 1%.

Information coded:

In an RCT, at the baseline and assuming n_t=n_c:

$Y_s \approx Y_t \approx Y_c \approx 1419.229$

$SD_t \approx SD_c \approx 978.3$

Beta= 37.375

SE(Beta)= 978.299

n_s= 6821; we assumed n_c= 3410 and n_t = 3411

RCT with regression analysis.

Banerjee et al. 2009

- Information coded:

  In an RCT, at the baseline and assuming $n_t = n_c$:

  $Y_s \approx Y_t \approx Y_c \approx 1419.229$

  $SD_t \approx SD_c \approx 978.3$

  Beta= 37.375

  SE(Beta)= 46.221

  n_s= 6821; we assumed n_c= 3410 and n_t = 3411

  RCT with regression analysis

Estimate Standardised Mean Difference (SMD):

1. Compute t:

$$t = \frac{37.375}{46.221} = 0.809$$

2. Compute $SD_{y:}$:

$$SD_y = \sqrt{\frac{1}{4} * \left(SD_t^2 + SD_c^2 + 2 * Cov(SD_t\ , SD_c\ )\right)} = \sqrt{\frac{1}{4} * \left(978.3^2\ + 978.3^2\ + 2 * 0.8 * 978.3 * 978.3\right)} = 928.1$$

3. Compute S_p:

$$S_p = \sqrt{\frac{\left(SD_y^2 * (n_t + n_c - 1)\right) - \left(\frac{\beta^2 * (n_t * n_c)}{n_t + n_c}\right)}{n_t + n_c}} = \sqrt{\frac{(928.1^2 * (3411 + 3410 - 1)) - \left(\frac{37.375^2 * (3411 * 3410)}{3411 + 3410}\right)}{3411 + 3410}} = 927.8$$

Banerjee et al. 2009

- Information coded:

  In an RCT, at the baseline and assuming $n_t=n_c$:

  $Y_s \approx Y_t \approx Y_c \approx 1419.229$

  $SD_t \approx SD_c \approx 978.3$

  Beta= 37.375

  SE(Beta)= 46.221

  n_s= 6821; we assumed n_c= 3410 and n_t = 3411

  t=0.809

  SD_y=928.1

  S_p=927.8

  RCT with regression analysis

Estimate Standardised Mean Difference (SMD):

4. Compute SMD:

$$SMD = \frac{\beta}{S_p} = \frac{37.375}{927.46} = 0.04$$

5. Compute SE(SMD):

$$SE(SMD) = \frac{SMD}{t} = \frac{0.054}{0.809} = 0.05$$

6. Interpret the results

7. For yourself: correct for sample bias

Banerjee et al. 2009

- Information coded:

In an RCT, at the baseline and assuming n_t=n_c:

$Y_s \approx Y_t \approx Y_c \approx 1419.229$

$SD_t \approx SD_c \approx 978.3$

Beta= 37.375

SE(Beta)= 46.221

n_s= 6821; we assumed n_c= 3410 and n_t = 3411

t=0.809

SD_y=928.1

S_p=927.8

RCT with regression analysis

Estimate Response Ratio (RR):

1. Compute RR:

$$RR = \frac{Y_s + \beta}{Y_s} = \frac{1419.229 + 37.375}{1419.229} = 1.026$$

2. Compute SE(RR):

$$SE(RR) = \mathrm{Exp}\left(\frac{Ln(RR)}{t}\right) = Exp\left(\frac{Ln(1.026)}{0.809}\right) = 1.033$$

3. Interpret the results

Banerjee et al. 2009

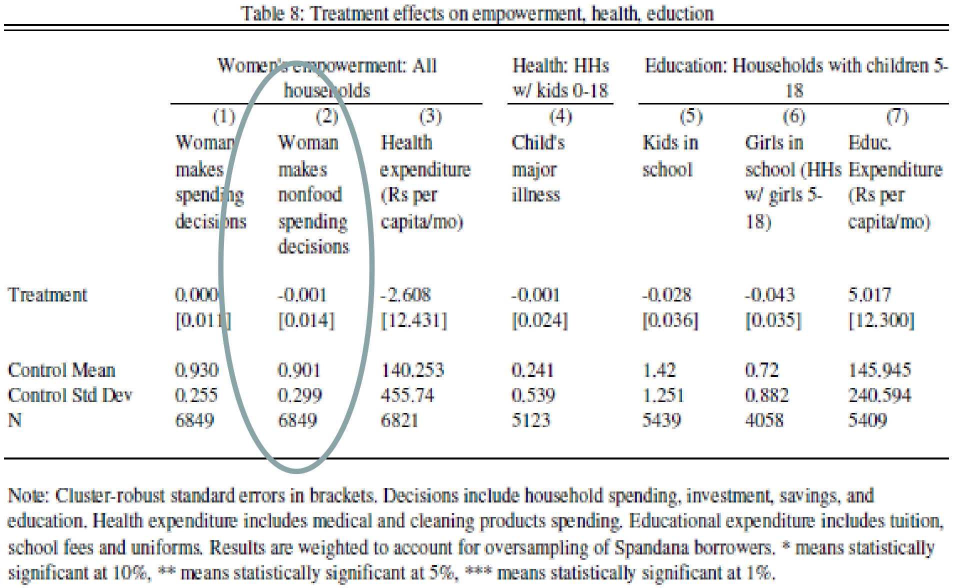


Table 8: Treatment effects on empowerment, health, eduction

	Women's empowerment: All households			Health: HHs w/ kids 0-18	Education: Households with children 5-18		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Woman makes spending decisions	Woman makes nonfood spending decisions	Health expenditure (Rs per capita/mo)	Child's major illness	Kids in school	Girls in school (HHs w/ girls 5-18)	Educ. Expenditure (Rs per capita/mo)
Treatment	0.000	-0.001	-2.608	-0.001	-0.028	-0.043	5.017
	[0.011]	[0.014]	[12.431]	[0.024]	[0.036]	[0.035]	[12.300]
Control Mean	0.930	0.901	140.253	0.241	1.42	0.72	145.945
Control Std Dev	0.255	0.299	455.74	0.539	1.251	0.882	240.594
N	6849	6849	6821	5123	5439	4058	5409

Note: Cluster-robust standard errors in brackets. Decisions include household spending, investment, savings, and education. Health expenditure includes medical and cleaning products spending. Educational expenditure includes tuition, school fees and uniforms. Results are weighted to account for oversampling of Spandana borrowers. * means statistically significant at 10%, ** means statistically significant at 5%, *** means statistically significant at 1%.

- Information coded:
  - In an RCT, at the baseline and assuming $n_t=n_c$:
    $Y_s \approx Y_t \approx Y_c \approx 0.901$
    $SD_t \approx SD_c \approx 0.299$
  - Beta= -0.001
  - SE(Beta)= 0.014
  - n_s= 6821; we assumed n_c= 3410 and n_t = 3411
  - t=-0.071
  - RCT with regression analysis

Banerjee et al. 2009

- Information coded:
  - In an RCT, at the baseline and assuming $n_t = n_c$:
    $Y_s \approx Y_t \approx Y_c \approx 0.901$
    $SD_t \approx SD_c \approx 0.299$
  - Beta= -0.001
  - SE(Beta)= 0.014
  - n_s= 6821; we assumed n_c= 3410 and n_t = 3411
  - t=-0.071
  - RCT with regression analysis

Estimate Standardised Mean Difference (SMD):

1. Compute t statistics:

$$t = \frac{\beta}{SE(\beta)} = \frac{-0.001}{0.014} = -0.07$$

2. Compute SD_y:

$$SD_y = \sqrt{\frac{1}{4} * \left(SD_t^2 + SD_c^2 + 2 * Cov(SD_t\ , SD_c\ )\right)} = \sqrt{\frac{1}{4} * \left(0.299^2 + 0.299^2 + 2 * 0.8 * 0.299 * 0.299\right)} = 0.284$$

3. Compute S_p:

$$S_p = \sqrt{\frac{\left((SD_y^2) * (n_t + n_c - 2)\right) - \left(\frac{\beta^2 * (n_t * n_c)}{n_t + n_c}\right)}{n_t + n_c}} = \sqrt{\frac{(0.284^2 * (3410 + 3411 - 1)) - \left(\frac{(-0.001)^2 * (3410 * 3411)}{3410 + 3411}\right)}{3410 + 3411}} = 0.284$$

Banerjee et al. 2009

- Information coded:
  - In an RCT, at the baseline and assuming n_t=n_c:

    $Y_s \approx Y_t \approx Y_c \approx 0.901$

    $SD_t \approx SD_c \approx 0.299$
  - Beta= -0.001
  - SE(Beta)= 0.014
  - n_s= 6821; we assumed

  n_c= 3410 and n_t = 3411
  - t=-0.071
  - SD_y=0.284
  - S_p=0.284
  - RCT with regression analysis

Estimate Standardised Mean Difference (SMD):

4. Compute SMD:

$$SMD = \frac{\beta}{S_p} = \frac{-0.001}{0.284} = -0.004$$

5. Compute SE(SMD):

$$SE(SMD) = \frac{SMD}{t} = \frac{-0.004}{-0.071} = 0.049$$

6. Interpret the results:

Banerjee et al. 2009

- Information coded:
  - In an RCT, at the baseline and assuming n_t=n_c:

    $Y_s \approx Y_t \approx Y_c \approx 0.901$

    $SD_t \approx SD_c \approx 0.299$
  - Beta= -0.001
  - SE(Beta)= 0.014
  - n_s= 6821; we assumed

  n_c= 3410 and n_t = 3411
  - t=-0.071
  - SMD=-0.004
  - SE(SMD)=0.207
  - RCT with regression analysis

Estimate Odds ratio (OR):

1. Compute OR:

$$OR = e^{\frac{SMD * \pi}{\sqrt{3}}} = e^{\frac{-0.004 * 3.142}{\sqrt{3}}} = 0.994$$

2. Compute SE(OR):

$$SE(OR) = e^{\sqrt{SE(SMD)^2 * \pi^2/3}} = e^{\sqrt{0.049^2 * 3.142^2/3}} = 1.094$$

3. Interpret the results:

Banerjee et al. 2009

- Information coded:
  - In an RCT, at the baseline and assuming $n_t = n_c$:

    $Y_s \approx Y_t \approx Y_c \approx 0.901$

    $SD_t \approx SD_c \approx 0.299$
  - Beta= -0.001
  - SE(Beta)= 0.014
  - n_s= 6821; we assumed n_c= 3410 and n_t = 3411
  - t=-0.071
  - RCT with regression analysis

Estimate Risk Ratio (RR):

1. Compute RR:

$$RR = \frac{Y_s + \beta}{Y_s} = \frac{0.901 - 0.001}{0.901} = 0.999$$

2. Compute SE(RR):

$$SE(RR) = \mathrm{Exp}\left(\frac{Ln(RR)}{t}\right) = Exp\left(\frac{Ln(0.999)}{-0.071}\right) = 1.016$$

3. Interpret the results

Chen et al. 2012

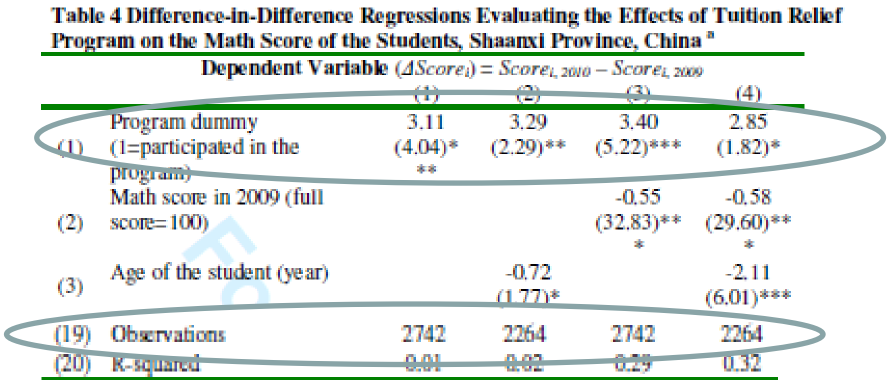


Table 4 Difference-in-Difference Regressions Evaluating the Effects of Tuition Relief Program on the Math Score of the Students, Shaanxi Province, China [a]

Dependent Variable ($\Delta Score_i$) = $Score_{i,\,2010} - Score_{i,\,2009}$

		(1)	(2)	(3)	(4)
(1)	Program dummy (1=participated in the program)	3.11 (4.04)***	3.29 (2.29)**	3.40 (5.22)***	2.85 (1.82)*
(2)	Math score in 2009 (full score=100)			-0.55 (32.83)***	-0.58 (29.60)***
(3)	Age of the student (year)		-0.72 (1.77)*		-2.11 (6.01)***
(19)	Observations	2742	2264	2742	2264
(20)	R-squared	0.01	0.02	0.29	0.32

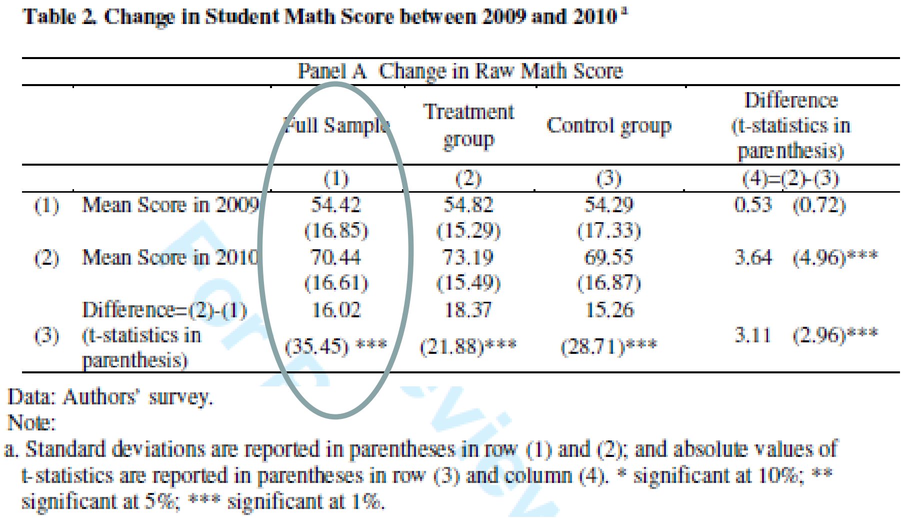


Table 2. Change in Student Math Score between 2009 and 2010 [a]

		Panel A Change in Raw Math Score			
		Full Sample	Treatment group	Control group	Difference (t-statistics in parenthesis)
		(1)	(2)	(3)	(4)=(2)-(3)
(1)	Mean Score in 2009	54.42 (16.85)	54.82 (15.29)	54.29 (17.33)	0.53 (0.72)
(2)	Mean Score in 2010	70.44 (16.61)	73.19 (15.49)	69.55 (16.87)	3.64 (4.96)***
(3)	Difference=(2)-(1) (t-statistics in parenthesis)	16.02 (35.45) ***	18.37 (21.88)***	15.26 (28.71)***	3.11 (2.96)***

Data: Authors' survey.

Note:

a. Standard deviations are reported in parentheses in row (1) and (2); and absolute values of t-statistics are reported in parentheses in row (3) and column (4). * significant at 10%; ** significant at 5%; *** significant at 1%.

- Information coded:

  Y_s= 16.02

  Beta=2.85

  t= 1.82

  n_t= 555

  n_c= 1709

  $SD_{y\ 2010}$=16.61

  $SD_{y\ 2009}$=16.85

  Diff-in Diff regression study

Chen et al. 2012

- Information coded:

  Y_s= 16.02

  Beta=2.85

  t= 1.82

  n_t= 555

  n_c= 1709

  $SD_{y\ 2010}$=16.61

  $SD_{y\ 2009}$=16.85

  Diff-in Diff regression study

Estimate Standardised Mean Difference (SMD):

1. Compute $SD_{y\cdot}$ on the change in the outcome:

$$SD_{y\,2010-2009} = \sqrt{SD^2_{y\,2009} + SD^2_{y\,2010} - 2 * Cov(SD_{y\,2010}, SD_{y\,2009})} = \sqrt{16.61^2 + 16.85^2 - 2 * 0.85 * 16.61 * 16.85} = 10.58$$

3. Compute S_p:

$$S_p = \sqrt{\frac{\left(SD_y^2 * (n_t + n_c - 1)\right) - \left(\frac{\beta^2 * (n_t * n_c)}{n_t + n_c}\right)}{n_t + n_c}} = \sqrt{\frac{(10.58^2 * (555 + 1709 - 1)) - \left(\frac{2.85^2 * (555 * 1709)}{555 + 1709}\right)}{555 + 2709}} = 10.51$$

Chen et al. 2012

- Information coded:

  Y_s= 16.02

  Beta=2.85

  SE(Beta)= 1.82

  n_t= 555

  n_c= 1709

  $SD_{y\ 2010}$=16.61

  $SD_{y\ 2009}$=16.85

  SD_y=10.58

  t=1.82

  S_p =10.51

  Diff-in Diff regression study

Estimate Standardised Mean Difference (SMD):

4. Compute SMD:

$$SMD = \frac{\beta}{S_p} = \frac{2.85}{10.51} = 0.27$$

5. Compute SE(SMD):

$$SE(SMD) = \frac{SMD}{t} = \frac{0.27}{1.82} = 0.149$$

6. Interpret the results
7. For yourself: correct for sample bias

Chen et al. 2012

- Information coded:

  Y_s= 16.02

  Beta=2.85

  SE(Beta)= 1.82

  n_t= 555

  n_c= 1709

  $SD_{y\ 2010}$=16.61

  $SD_{y\ 2009}$=16.85

  SD_y=10.58

  t=1.82

  S_p =10.51

  Diff-in Diff regression study

Estimate Response Ratios (RR):

1. Compute RR:

$$RR = \frac{Y_s + \beta}{Y_s} = \frac{16.02 + 2.85}{16.02} = 1.178$$

2. Compute SE(RR):

$$SE(RR) = \text{Exp}\left(\frac{Ln(RR)}{t}\right) = Exp\left(\frac{Ln(1.178)}{1.82}\right) = 1.094$$

3. Interpret the results

Practical tips

- David Wilson effect size calculator available here:

http://gunston.gmu.edu/cebcp/EffectSizeCalculator/index.html

but remember to apply corrections for sample bias!

- Code all the relevant information in a spreadsheet for all relevant studies before starting ES calculations. Decision on the selection of ES measure would depend on which ES measure would lead to the smaller study loss.

Main Bibliography Used

- Banerjee, A., Duflo, E., Glennerster, R., Kinnan, C. 2009, "The Miracle of Microfinance? Evidence from a Randomized Evaluation", *JPAL*
- Becker, B. & Wu, M. 2007, "The Synthesis of Regression Slopes in Meta-Analysis", *Statistical Science*, Vol. 22, No 3, 414-429.
- Borenstein, M., Hedges, L., Higgins, J., Rothstein,H. 2009, "Introduction to Meta-Analysis". WILEY
- Chen, X., Shi, Y., Yi, H., Zhang, L., Mo, D., Chu, J., Rozelle, S. 2012, "The Impact of a Senior High School Tuition Relief Program on Poor Junior High School Students in Rural China", *REAP Working Paper 239, University of Stanford*.
- Hansen, H., Klejntrup, N.R., Andersen, O.W. 2011, "A comparison of Model-based and Design-based Impact Evaluations of Interventions in Developing Countries", FOI Working Paper n o.16, University of Copenhagen.
- Hasler, S. *Forthcoming,* "The Effect of Multi-grade Teaching on Education Quality in Cameroon. A Propensity Score Matching Model", *Journal of Development Studies*.
- Hedges, L.V. 1981, "Distribution Theory for Glass's Estimator of Effect Size and Related Estimators", *Journal of Educational Statistics,* vol. 6, pp107-128.
- Keef, S.P. & Roberts, L.A. 2004. "The meta-analysis of partial effect sizes", *British J. Math. Statist. Psych*, Vol 57, pp 97-129.
- Lipsey, M.W. & Wilson, D.B. 2001, "Practical Meta-Analysis", *Applied Social Research Methods Series*, Volume 49. SAGE Publications.

Thank you very much

Visit:

www.3ieimpact.org

http://www.campbellcollaboration.org/international_development/index.php

Annex: standard errors and 95% CI for OR / RR

$$\overline{SE_{OR}} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}$$

95% confidence interval = exp[ln(OR))±1.96*ln(SE$_{OR}$)]

$$\overline{SE_{RR}} = \frac{1}{a} - \frac{1}{a+b} + \frac{1}{c} - \frac{1}{c+d}$$

95% confidence interval = exp[ln(RR))±1.96*ln(SE$_{RR}$)]