



SDG accelerator analyses: A step by step guide on what to do with observational data

GCRF Hub Meeting, Jan 26th
Will Rudgard and Mark Orkin



What is an SDG? Have you memorized all 17 yet?



What is an SDG Accelerator?



Actions (e.g interventions: cash transfers, parenting support) that will take us forward more quickly across a broad range of interlinked goals



How to choose SDG accelerators?



Accelerators may be:

- Interventions: cash transfer, schooling (no fees, meals)
- Practices for which interventions could be implemented: positive parenting, safe school

Accelerators shouldn't be too vague:

- Reduce poverty (too broad)
- Improve food security (better)

A study may investigate one, or multiple (even six!) potential accelerators





Outcomes should be:

- Operationalisable (i.e. validly measurable)
- Modifiable (something that can be improved)
- Salient in countries' national development agendas
- Emerging priorities
- Responsive to the needs of the traditionally marginalized



Source: UNDP (2018). SDG Accelerator and Bottleneck Assessment



A motivating example: Identifying accelerators of child violence



An example accelerator

Accelerators: Actions that impact on multiple outcomes.

Cash transfers impacting on:

- a) Abuse
- b) School progression
- c) HIV care retention





1. Map potential accelerators
2. Map potential SDG outcomes
3. Select control variables
4. Descriptive analysis
5. Single and multiple outcome regressions
6. Benjamini & Hochberg procedure
7. Evaluate synergy effects of accelerators





Prospective cohort study (2 time points)

Merged data from South Africa:

- Western Cape
- Mpumalanga
- Eastern Cape

Sample size: 4811

Six hypothesized accelerators

- Measured as access at both baseline and follow-up

Six adolescent outcomes linked to targets across 2 SDGs

- Measured at follow-up





Step 1: Map potential accelerator variables

Interventions or circumstances influencing adolescent's behaviors, nutrition, health, and education and work.

- Free schooling
- Free school meals
- Cash transfers
- Parenting support
- Food security
- Vocational training
- Access to basic necessities (clean energy and water)





Step 1: Example of INSPIRE analysis

Which potential accelerators were available after data merge?

Final selection:

1. Positive parenting
2. Parental monitoring and supervision
3. Food security at home
4. Non-food poverty at home
5. Free schooling
6. Free school meals



Step 2: Map potential outcome variables

Assess availability of relevant indicators within SDGs.

For adolescents important SDGs are likely to fall across:





Step 2: Example of INSPIRE analysis

Which SDG indicators were available after data merge?

SDG 5:

1. Sexual abuse (Target 5.2)
2. Transactional sex (Target 5.6)

SDG 16:

1. Youth violence (Target 16.1)
2. Emotional abuse (Target 16.2)
3. Physical abuse (Target 16.2)
4. Community violence victimisation (Target 16.1)





Step 3: Control variables

- Control variables are important for (i) increasing precision of results and (ii) reducing bias
- Important controls should be well thought out and usually informed by prior knowledge / literature
- For interventions (e.g. cash transfers) one should control for reasons beyond the research question why families' might receive them (Caregiver education, geographical location)
- Other common control variables: socio-demographic factors (age and sex) and external shocks (drought, unemployment)





Step 3: Example of INSPIRE analysis

Which controls were available after data merge?

1. Age and Gender
2. Informal housing
3. Household size
4. Maternal orphanhood
5. Paternal orphanhood
6. Province of residence
7. HIV status

All of these controls were measured at baseline.



Step 4: How are potential “accelerators” and SDG outcomes associated?



Check 1. Are potential accelerators correlated with each other?

- Correlation matrices (categorical and continuous variables)
- Follow-up with 2x2 cross tabulation (categorical variables)

An assumption of regression methods is that independent variables are not too strongly correlated (i.e. values less than 0.7)

Check 2. How are potential accelerators associated with SDG outcomes?

This will add further detail when interpreting regression output.





Step 4: Example of Check 1 in INSPIRE analysis

Commands in STATA 15 (Mark O the expert in SPSS)

```
. spearman PosParentHighT1T2 MonitHighT1T2 FoodSecT1T2 SchFreeT1T2 SchMealsT1T2 ///
> NesitiesT1T2 , pw stats( rho p )
(obs=4640)
```

| |
|-------------------|
| Key |
| <i>rho</i> |
| <i>Sig. level</i> |

| | Po~hT1T2 | MonitH~2 | FoodSe~2 | SchFre~2 | SchMea~2 | Nesiti~2 |
|--------------|-------------------|-------------------|-------------------|-------------------|-------------------|----------|
| PosParentH~2 | 1.0000 | | | | | |
| MonitHighT~2 | 0.1620 0.0000 | 1.0000 | | | | |
| FoodSecT1T2 | 0.1347 0.0000 | 0.0799 0.0000 | 1.0000 | | | |
| SchFreeT1T2 | -0.0474 0.0012 | -0.0708 0.0000 | -0.0409 0.0054 | 1.0000 | | |
| SchMealsT1T2 | -0.0087 0.5546 | 0.0526 0.0003 | -0.0366 0.0127 | 0.3191 0.0000 | 1.0000 | |
| NesitiesT1T2 | 0.1538 0.0000 | 0.1307 0.0000 | 0.2861 0.0000 | -0.0842 0.0000 | -0.1183 0.0000 | 1.0000 |

```
. tab SchFreeT1T2 SchMealsT1T2 , row m chi
```

| |
|-----------------------|
| Key |
| <i>frequency</i> |
| <i>row percentage</i> |

| T1T2: teen went to a free school goes to free school? | T1T2: Teen had at least one meal a day at school at T1 & T2? | | Total |
|---|--|----------------|-----------------|
| | 0. No | 1. Yes | |
| 0. No | 838 29.80 | 1,974 70.20 | 2,812 100.00 |
| 1. Yes | 71 3.88 | 1,757 96.12 | 1,828 100.00 |
| Total | 909 19.59 | 3,731 80.41 | 4,640 100.00 |

Pearson chi2(1) = 472.3719 Pr = 0.000





Step 4: Example of Check 2 in INSPIRE analysis

Commands in STATA 15 (Mark O the expert in SPSS)

1.

| T1T2: teen went to a free school goes to free school? | ViolncPerpT2 | | Total |
|---|----------------|----------------|-----------------|
| | 0. No | 1. Yes | |
| 0. No | 2,141 76.14 | 671 23.86 | 2,812 100.00 |
| 1. Yes | 1,468 80.31 | 360 19.69 | 1,828 100.00 |
| Total | 3,609 77.78 | 1,031 22.22 | 4,640 100.00 |

Pearson chi2(1) = 11.1377 Pr = 0.001

2.

| T1T2: teen went to a free school goes to free school? | AbuEmoT2 | | Total |
|---|----------------|----------------|-----------------|
| | 0. No | 1. Yes | |
| 0. No | 2,002 71.19 | 810 28.81 | 2,812 100.00 |
| 1. Yes | 1,288 70.46 | 540 29.54 | 1,828 100.00 |
| Total | 3,290 70.91 | 1,350 29.09 | 4,640 100.00 |

Pearson chi2(1) = 0.2904 Pr = 0.590

3.

| T1T2: teen went to a free school goes to free school? | AbuPhysT2 | | Total |
|---|----------------|----------------|-----------------|
| | 0. No | 1. Yes | |
| 0. No | 1,921 68.31 | 891 31.69 | 2,812 100.00 |
| 1. Yes | 1,182 64.66 | 646 35.34 | 1,828 100.00 |
| Total | 3,103 66.88 | 1,537 33.13 | 4,640 100.00 |

Pearson chi2(1) = 6.6754 Pr = 0.010

Here we've shown one potential accelerator (free schooling) with three of our six violence outcomes of interest.



Step 5: Run single outcome regressions with potential accelerators and controls



Testing for accelerators, each regression model should include:

1. Specified set of potential “accelerators”
2. Specified set of controls

With data that has more than one timepoint also include as a control:

- Experience of SDG outcome at an earlier timepoint (e.g. baseline).

Models should adjust for similar covariates. This is necessary for valid comparisons across outcomes.





Logistic regression in STATA

```
* Sexual abuse (with baseline control)
logit ///
/*Outcome*/ AbuSexT2 ///
/*Accelerators*/ c.PosParentScoreT1T2 c.MonitParentScoreT1T2 i.FoodSecT1T2 ///
/*Accelerators*/ i.SchFreeT1T2 i.SchMealsT1T2 i.NesitiesT1T2 ///
/*T1 Control*/ AbuSexT1 ///
/*Covariates*/ c.AgeT1 i.GenderT1 i.RuralT1 i.InformalT1 ///
/*Covariates*/ i.MOrphanT1 i.POrphanT1 c.HHsizeT1 i.HIVstatusT1 i.ProvinceT1 ///
, or baselevels
```



Step 5: Example from INSPIRE analysis



Logistic regression

Number of obs = 4,640

LR chi2(17) = 111.05

Prob > chi2 = 0.0000

Pseudo R2 = 0.0817

Log likelihood = -623.71107

| Outcome | AbuSexT2 | Odds Ratio | Std. Err. | z | P> z | [95% Conf. Interval] | |
|----------------------|----------------------|------------|-----------|-------|-------|----------------------|----------|
| Continuous predictor | PosParentScoreT1T2 | .9556141 | .0154406 | -2.81 | 0.005 | .9258253 | .9863613 |
| | MonitParentScoreT1T2 | .9564474 | .0210559 | -2.02 | 0.043 | .9160562 | .9986195 |
| Binary predictor | FoodSecT1T2 | | | | | | |
| | 0 | 1 | (base) | | | | |
| | 1 | .6297968 | .1129774 | -2.58 | 0.010 | .4431042 | .8951484 |
| | SchFreeT1T2 | | | | | | |
| | 0. No | 1 | (base) | | | | |
| | 1. Yes | 1.427732 | .295289 | 1.72 | 0.085 | .9519145 | 2.141389 |
| SchMealsT1T2 | 0. No | 1 | (base) | | | | |
| | 1. Yes | .7929251 | .1770071 | -1.04 | 0.299 | .5119362 | 1.228142 |
| NesitiesT1T2 | 0 | 1 | (base) | | | | |
| | 1 | .9562849 | .2120449 | -0.20 | 0.840 | .6192177 | 1.476832 |

Output for controls not shown...





Step 5: Example of INSPIRE analysis

Accelerators: Actions that impact on multiple outcomes.

Regression 1

| | Sexual Abuse |
|-------------------------------------|--------------|
| Positive parenting | |
| Parental monitoring and supervision | |
| Food security | |
| Free schooling | |
| Free school meals | |
| Non-food poverty | |





Logistic regression in STATA

```
* Transactional sex (with no baseline control)
logit ///
/*Outcome*/ TST2 ///
/*Accelerators*/ c.PosParentScoreT1T2 c.MonitParentScoreT1T2 i.FoodSecT1T2 ///
/*Accelerators*/ i.SchFreeT1T2 i.SchMealsT1T2 i.NesitiesT1T2 ///
/*T1 Control*/ ///
/*Covariates*/ c.AgeT1 i.GenderT1 i.RuralT1 i.InformalT1 i.MOrphanT1 ///
/*Covariates*/ i.POrphanT1 c.HHsizeT1 i.HIVstatusT1 i.ProvinceT1 ///
, or baselevels
```



Step 5: Example from INSPIRE analysis



Logistic regression

Number of obs = 4,640

LR chi2(16) = 606.60

Prob > chi2 = 0.0000

Log likelihood = -987.25786

Pseudo R2 = 0.2350

| Outcome | → TST2 | Odds Ratio | Std. Err. | z | P> z | [95% Conf. Interval] |
|----------------------|----------------------|------------|-----------|-------|-------|----------------------|
| Continuous predictor | PosParentScoreT1T2 | .9960464 | .0119491 | -0.33 | 0.741 | .9728998 1.019744 |
| | MonitParentScoreT1T2 | .9277421 | .0141108 | -4.93 | 0.000 | .9004937 .9558151 |
| Binary predictor | FoodSecT1T2 | | | | | |
| | 0 | 1 | (base) | | | |
| | 1 | .6635351 | .0854829 | -3.18 | 0.001 | .5154708 .8541295 |
| | SchFreeT1T2 | | | | | |
| | 0. No | 1 | (base) | | | |
| | 1. Yes | .7083414 | .1079006 | -2.26 | 0.024 | .5255092 .9547835 |
| SchMealsT1T2 | 0. No | 1 | (base) | | | |
| | 1. Yes | .7732244 | .1215722 | -1.64 | 0.102 | .5681633 1.052296 |
| NesitiesT1T2 | 0 | 1 | (base) | | | |
| | 1 | .8157752 | .1420056 | -1.17 | 0.242 | .5799601 1.147474 |

Output for controls not shown...





Step 5: Example of INSPIRE analysis

Accelerators: Actions that impact on multiple outcomes.

| | Regression 1 | Regression 2 | Regression 3 | Regression 4 | Regression 5 | Regression 6 |
|-------------------------------------|--------------|-------------------|----------------|-----------------|----------------|------------------------|
| | Sexual Abuse | Transactional Sex | Physical Abuse | Emotional Abuse | Youth Violence | Violence Victimization |
| Positive parenting | Yellow | | Yellow | Yellow | Yellow | |
| Parental monitoring and supervision | Yellow | Yellow | Yellow | Yellow | Yellow | Yellow |
| Food security | Yellow | Yellow | Yellow | Yellow | | Yellow |
| Free schooling | | | | | | Yellow |
| Free school meals | | Yellow | | | Yellow | |
| Non-food poverty | | | | | | Yellow |

Red line serves to highlight accelerator effect of action on multiple outcomes.





Step 6. Benjamini and Hochberg correction

When conducting many hypothesis tests some will have p-values less than 0.05 purely by chance (even when there is no association).

100 hypothesis tests = 5 significant results

False positive

Benjamini-Hochberg minimizes the number of significant results that are false positives.





Step 6. Benjamini and Hochberg correction

False discovery rate (FDR) is the expected proportion of false positives in significant results.

Using the Benjamini & Hochberg correction we specify the FDR we are willing to tolerate.

For example, if $FDR = 0.1$

- 100 hypothesis tests

- 10 significant results, we expect to have 1 false positive





More to come





A quick recap of concepts



Methodological approach



What is definite for accelerators?

- Multiple outcomes
- Consider possible correlation between accelerators
- Need to control for potential confounding
- Need to control for multiple outcome testing (B&H correction).

What is desirable?

- More than one time point
- All of the general criteria for a rigorous study design





A motivating example:
Identifying synergies across
accelerators of child violence



An example synergy across accelerators



The combined effect of two or more accelerators:

| | Regression 1 | Regression 2 | Regression 3 | Regression 4 | Regression 5 | Regression 6 |
|-------------------------------------|--------------|-------------------|----------------|-----------------|----------------|------------------------|
| | Sexual Abuse | Transactional Sex | Physical Abuse | Emotional Abuse | Youth Violence | Violence Victimization |
| Positive parenting | | ↑ | ↑ | ↑ | ↑ | |
| Parental monitoring and supervision | ↑ | | | | | ↑ |
| Food security | ↑ | | | | | ↑ |
| | | | | | | |

Red arrows serve to highlight synergy effect of multiple interventions on one outcome (e.g. sexual abuse).



Recap of steps in accelerator analysis

For existing secondary data:

1. Map potential accelerators
2. Map potential SDG outcomes
3. Select control variables
4. Descriptive analysis
5. Single and multiple outcome regressions (multiple outcome regression saved for another time!)
6. Benjamini & Hochberg procedure
7. Evaluate synergy effects of accelerators (marginal effects estimation – saved for later!)



Thank you!

william.rudgard@spi.ox.ac.uk

mark.orkin@gmail.com

