THE CAMPBELL COLLABORATION
What helps? What harms? Based on what evidence?

Impact Evaluation for Development in a Global Context

Howard White

CEO, Campbell Collaboration

@c2update @HowardNWhite

More opportunities, a better future

More opportunities, a better future… But how do we know?

- Results agenda and New Public Management – such as US Government Results and Performance Act 1993 and UK *Modernizing Government* White Paper 1997

- Focused on outcome monitoring

- Which has its place, but DOES NOT TELL YOU WHAT DIFFERENCE YOU ARE MAKING (eg USAID)
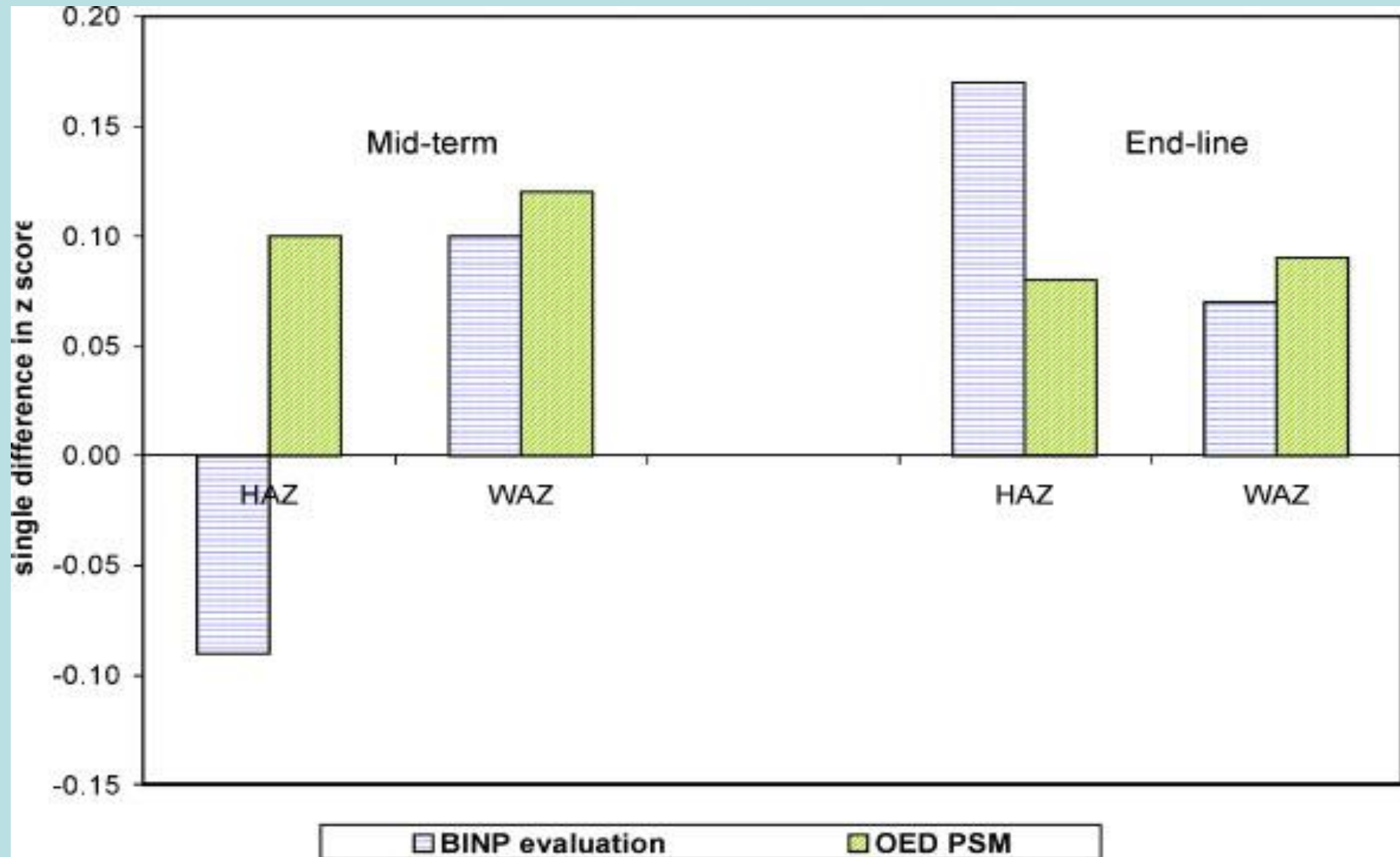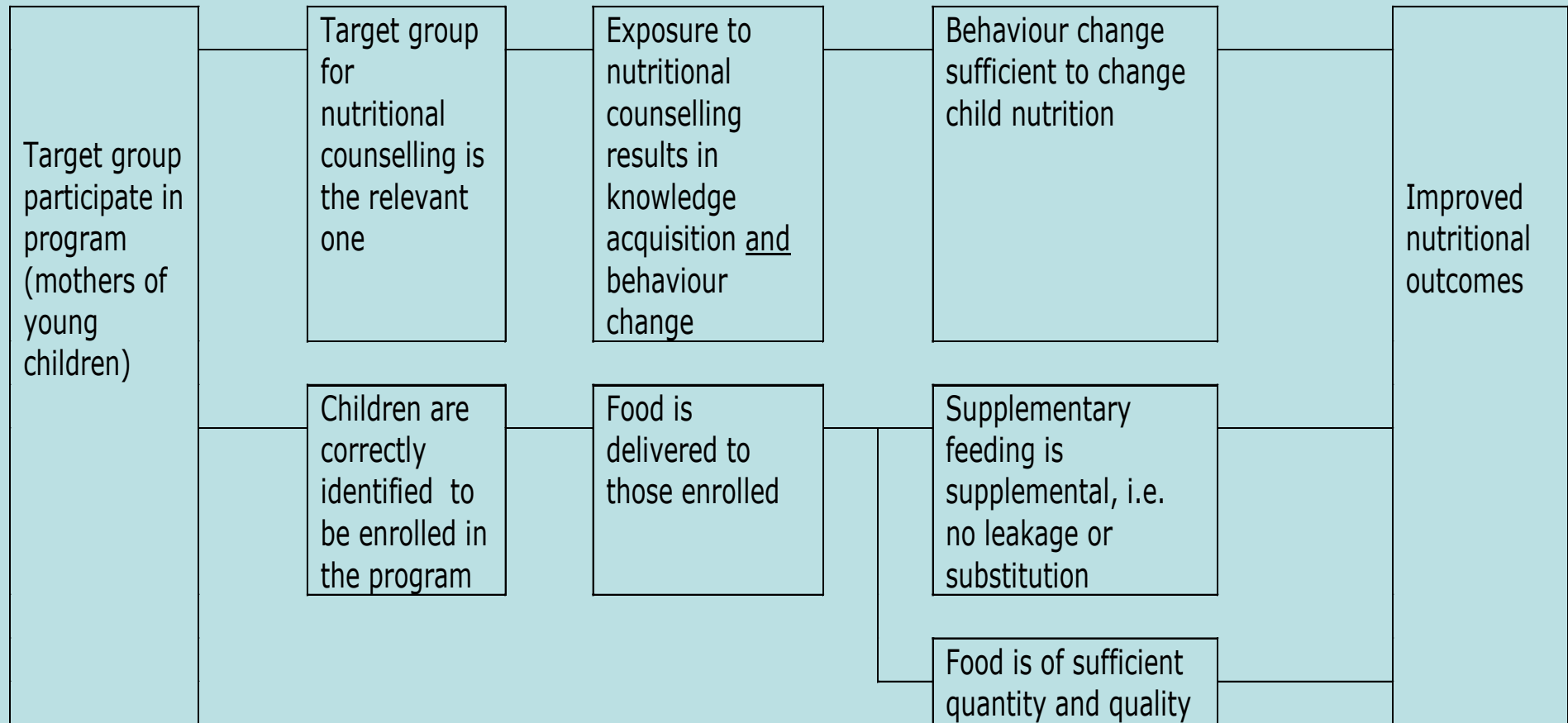
The misguided results agenda: a cautionary tale

The case of the Bangladesh Integrated Nutrition Project (BINP)

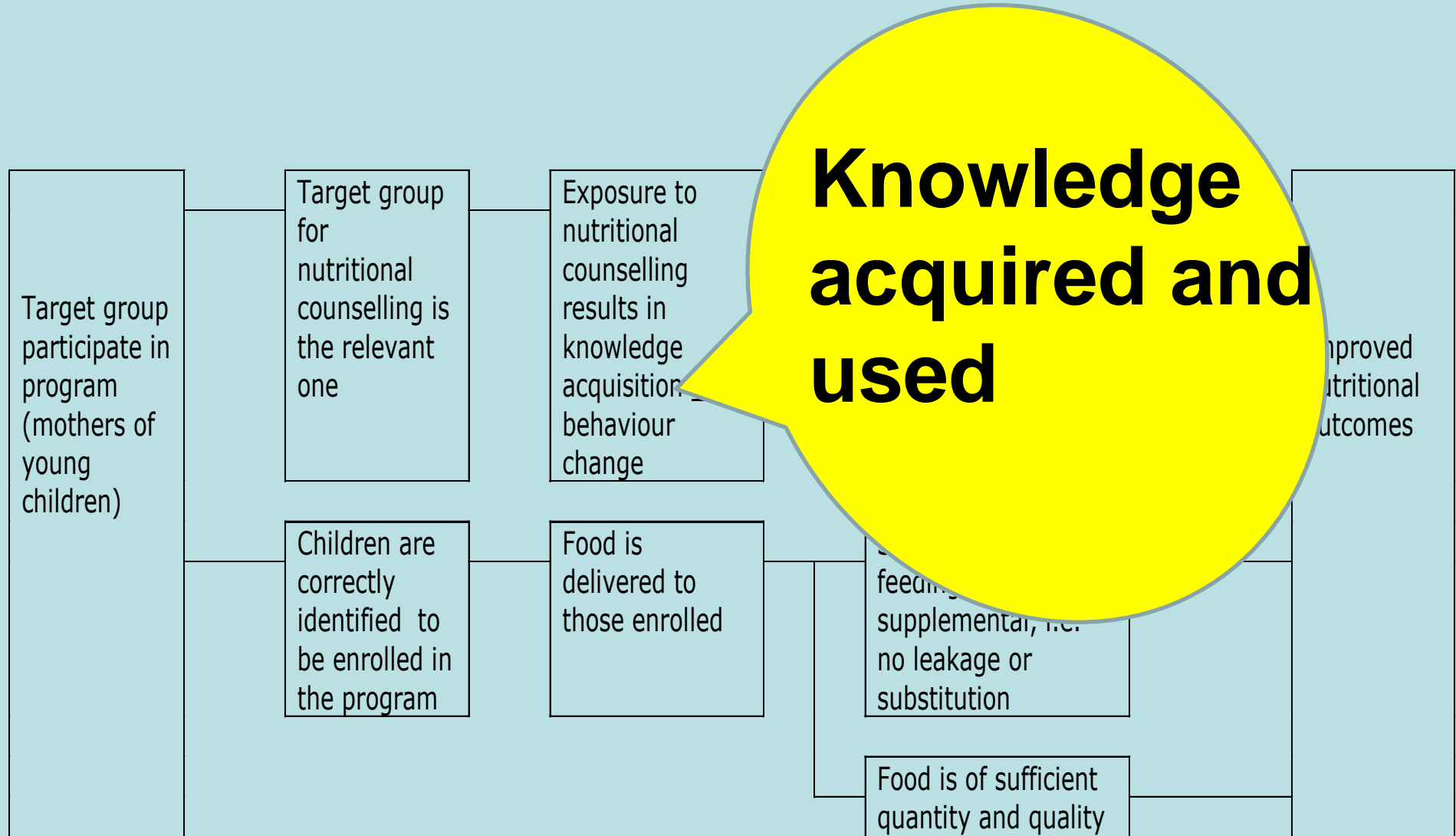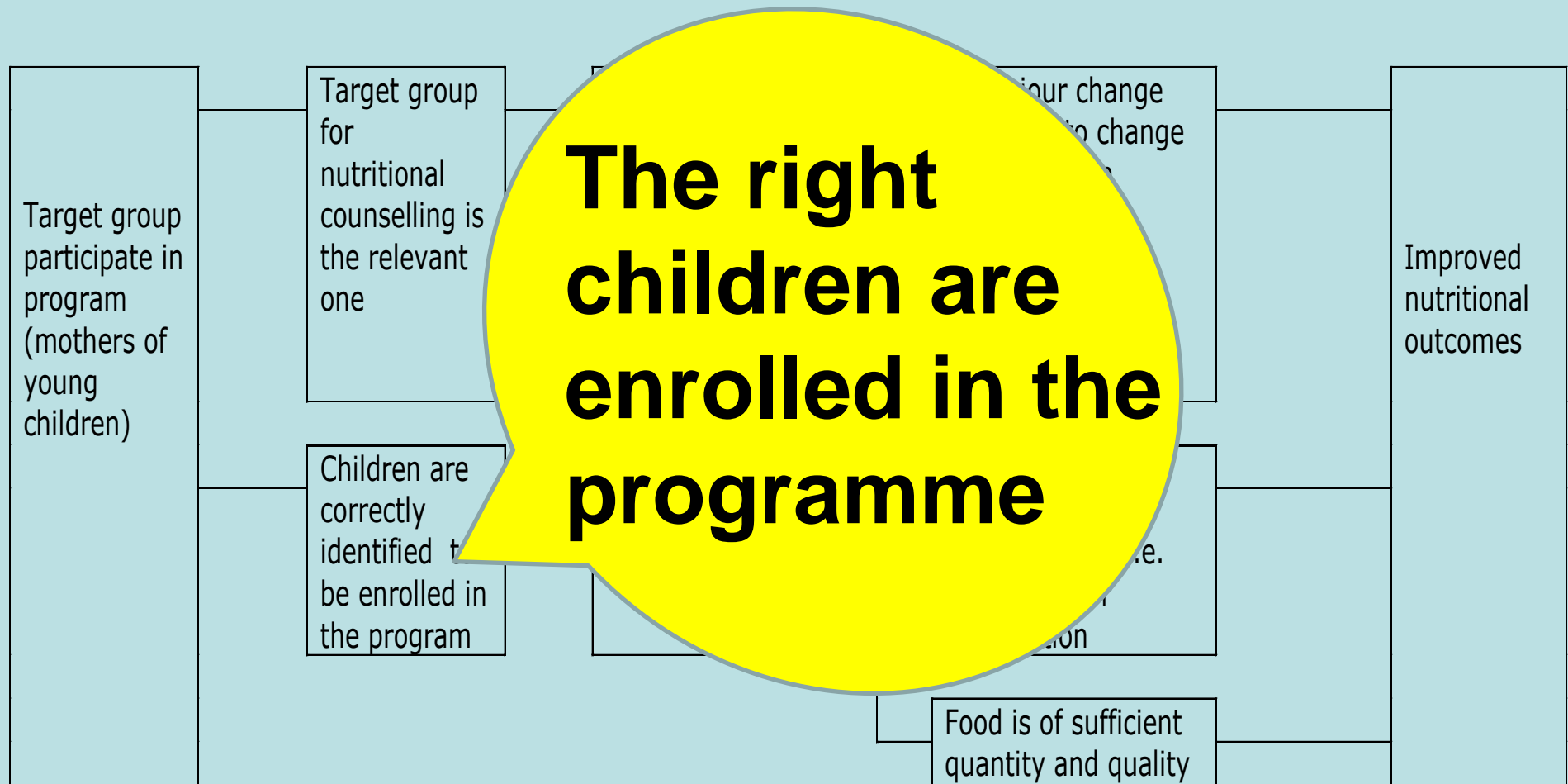# BINP: Comparison of impact estimates

Summary of theory


```
Target group          Target group        Exposure to          Behaviour change
participate in        for                 nutritional          sufficient to change
program               nutritional         counselling          child nutrition
(mothers of           counselling is      results in
young                 the relevant        knowledge
children)             one                 acquisition and
                                          behaviour
                                          change

                      Children are        Food is              Supplementary
                      correctly           delivered to         feeding is
                      identified  to      those enrolled       supplemental, i.e.
                      be enrolled in                           no leakage or
                      the program                              substitution

                                                               Food is of sufficient
                                                               quantity and quality

                                                                                        Improved
                                                                                        nutritional
                                                                                        outcomes
```

The theory of change

Target group participate in program (mothers of young children)

Target group for nutritional counselling is the relevant one

Exposure to nutritional counselling results know acc

Child are corre identif be enro

Right target group for nutritional

PARTICIPATION RATES WERE UP TO 30% LOWER FOR WOMEN LIVING WITH THEIR MOTHER-IN-LAW

The theory of change

Target group participate in program (mothers of young children)

Target group for nutritional counselling is the relevant one

Exposure to nutritional counselling results in knowledge acquisition & behaviour change

Knowledge acquired and used

Children are correctly identified to be enrolled in the program

Food is delivered to those enrolled

feeding supplemental, i.e. no leakage or substitution

Food is of sufficient quantity and quality

Improved nutritional outcomes

The theory of change

Target group participate in program (mothers of young children)

Target group for nutritional counselling is the relevant one

Children are correctly identified to be enrolled in the program

...iour change
...to change...

The right children are enrolled in the programme

...e.
...ion

Food is of sufficient quantity and quality

Improved nutritional outcomes

The theory of change

Target group participate in program (mothers of young children)

Target group for nutritional counselling is the relevant one

Exposure to nutritional counselling results in knowledge acquisition and behaviour change

Children are correctly identified to be enrolled in the program

Food is delivered to those enrolled

Supplementary feeding is supplemental, i.e. no leakage or substitution

Food is of sufficient quantity and quality

Supplementary feeding is supplementary

So nutrition was improving in BINP areas... but it was improving all over Bangladesh

Lessons from BINP

- Apparent successes can turn out to be failures

- Outcome monitoring does not tell us impact and can be misleading: **only rigorous impact evaluation does this**

All this is some years ago. Since then large rise in impact evaluation….

Grown rapidly but focused health, then education…

Studies published each year
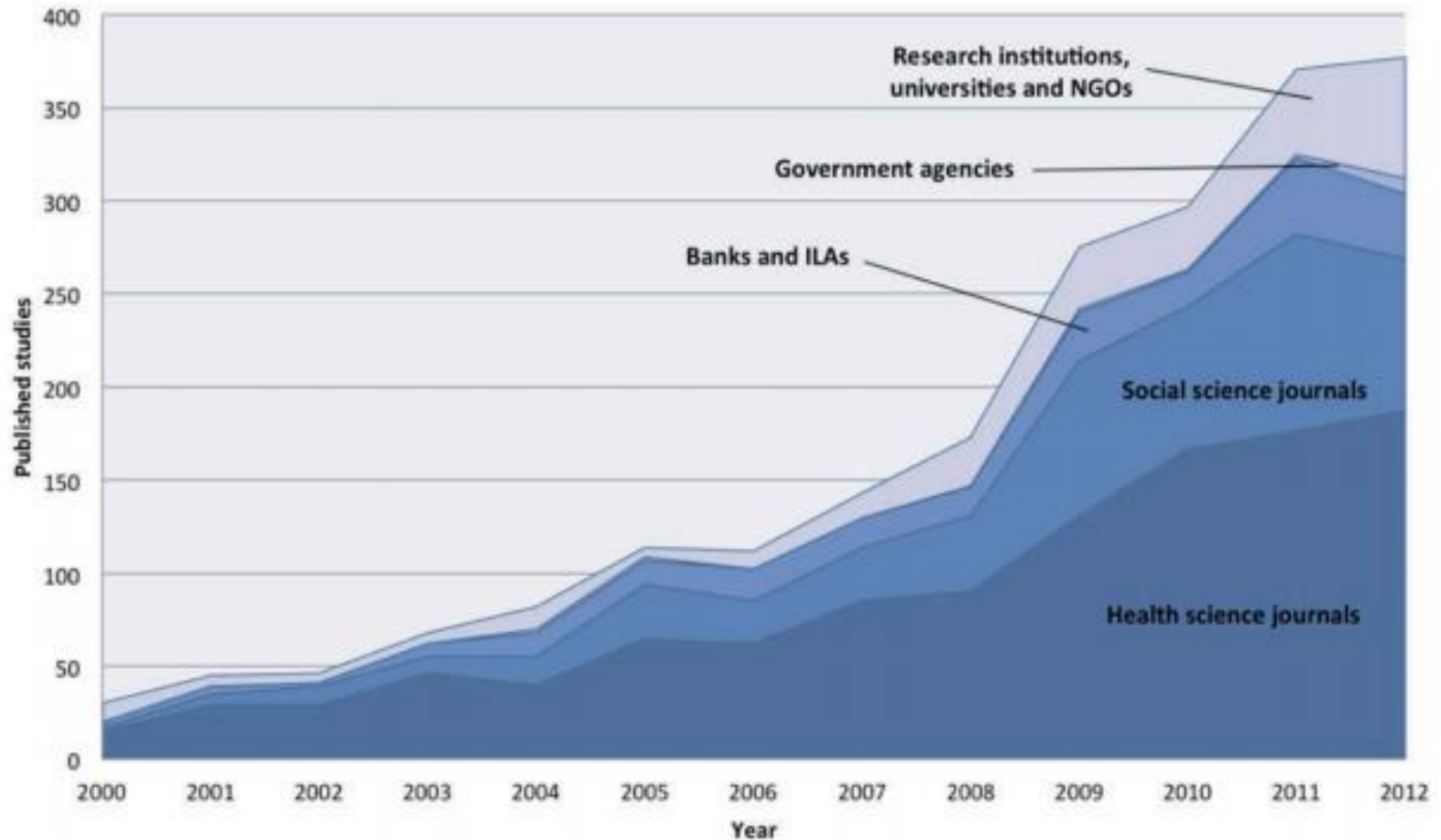
So what about infrastructure?

Is randomization possible?

Yes, for many interventions it is

Off-grid electricity, can randomize at the household level (e.g. solar home systems) or community level (e.g. micro-hydo)

Urban development: slum upgrading, can randomize at household level (improved housing) or settlement level (community-level services, e.g. street lighting)

Water supply and sanitation: community (e.g. standpipes) or household (latrines, point of use water treatment)

Even large scale infrastructure

Rural electrification

Matched pair randomization

V4

V3

V2

V1

Pipeline design

V10

V9

V8

V7

V5

V6
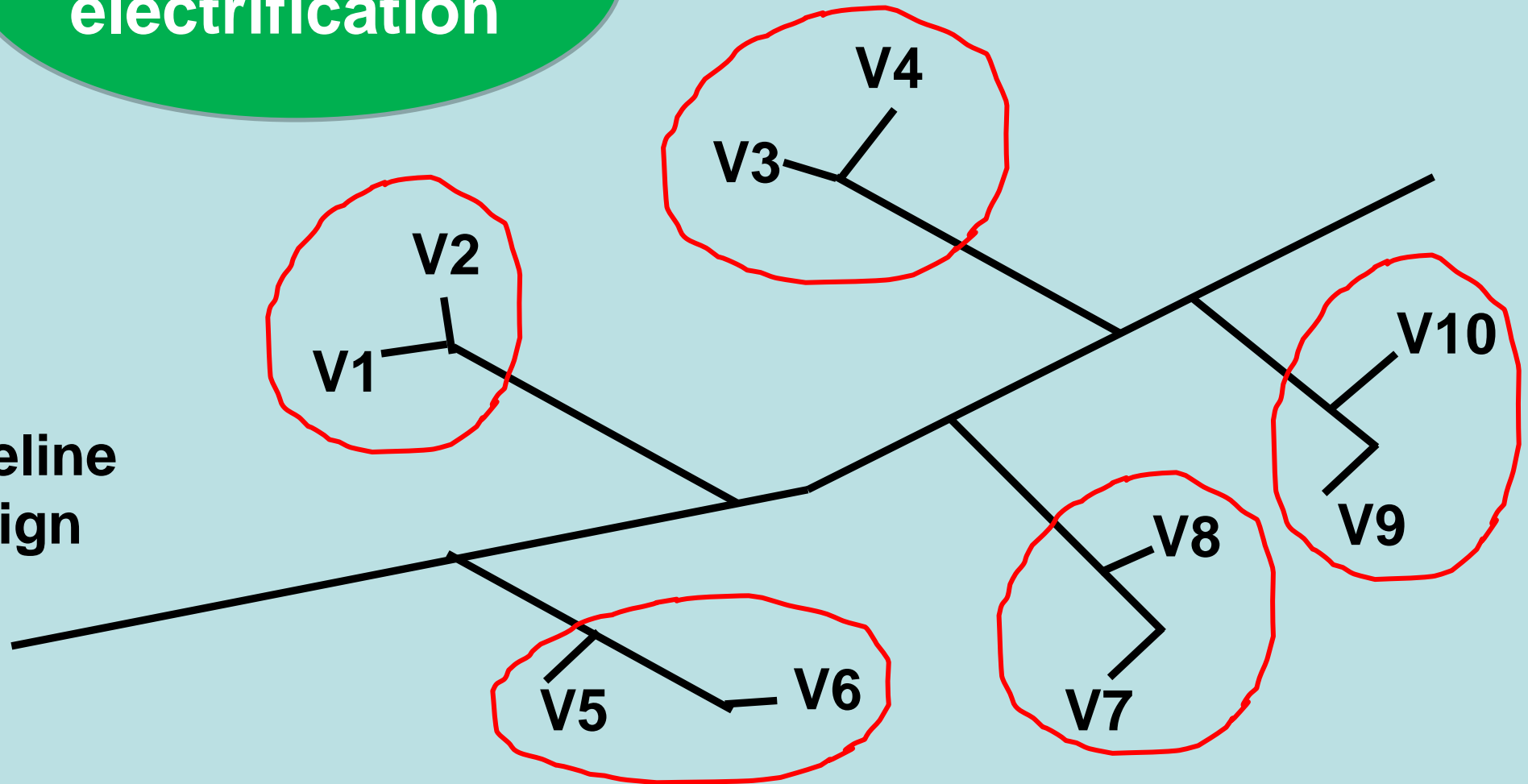
Other possible random roll out (pipeline design)

- Rural roads (or urban upgrading)

- Rehabilitation (e.g. piped water system)

- Secondary or tertiary irrigation canals

Or can examine policy reform

New system of pollution auditing in Gujarat

Agricultural water pricing in West Bengal

Can randomize pricing

- Subsidy for electricity connection charge

- Vouchers for

  – Utilities

  – Road and bridge tariffs

  – Using transport services

- Vary community-level subsidy for services

And you can use an encouragement design

- Everyone is exposed to the treatment
- Random assignment is of encouragement, not the treatment
- But the encouragement must not affect the outcomes

- **Examples of possible encouragements are information or reducing transaction costs (randomized pricing is an encouragement design)**

Non-experimental methods can be used

- If randomization not possible because
  - Study is ex-post
  - Administrative or political difficulties
- There may be a natural experiment
  - John Snow water supply in London 1850s cholera outbreak
  - Water privatization in Peru
- Can use ex-post design
  - Instrumental variables have been used for large infrastructure e.g. dams and roads
  - E.g. Propensity score matching at household or community level
  - Regression discontinuity of eligibility criteria for service subsidy
  - Preferably always double difference

Balance achieved through PSM
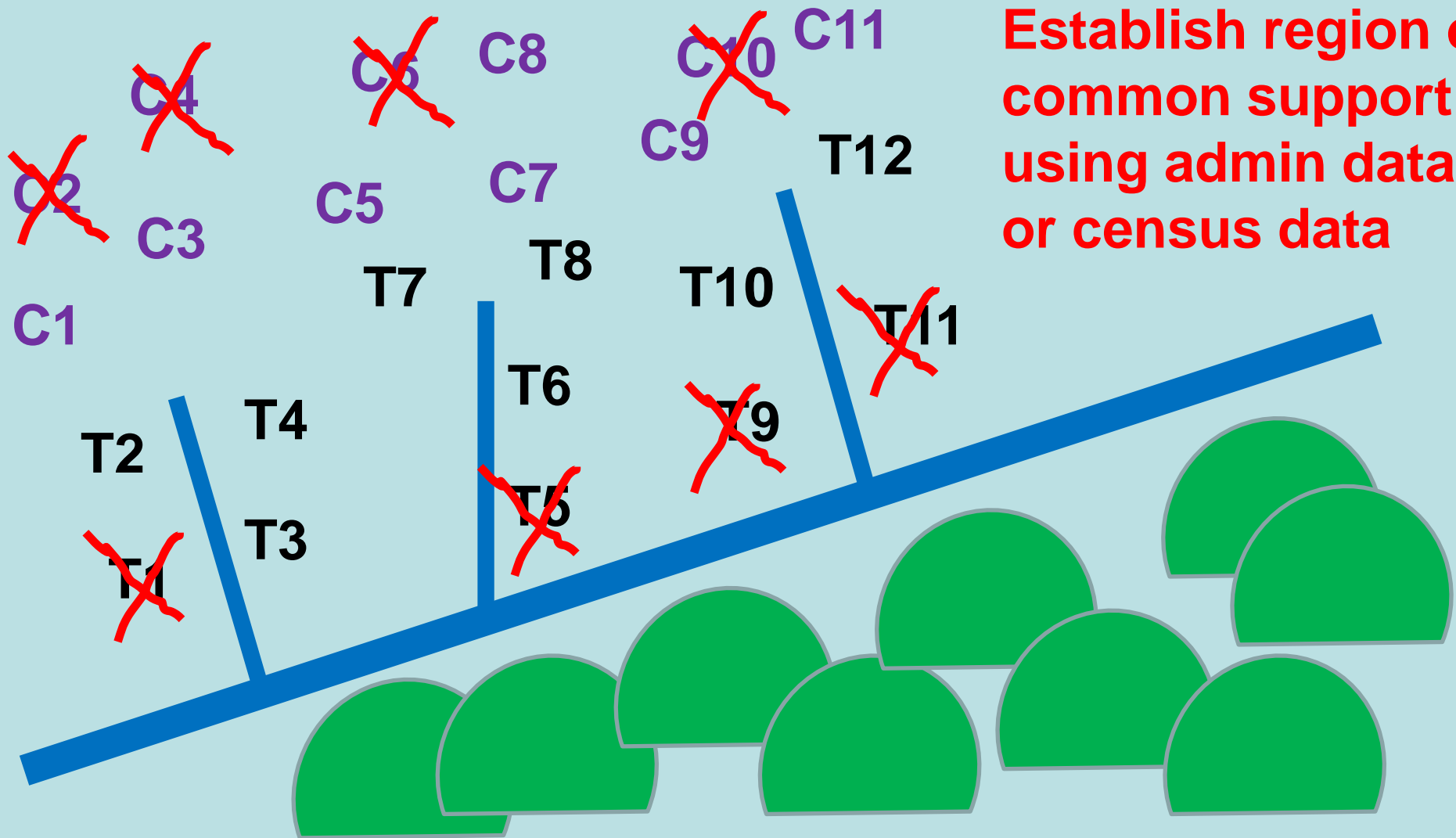
Access to improved water source in Nepal

Variable	Before matching	After matching
Rural resident	Treatment: 29% Comparison: 78%	Treatment: 33% Comparison: 38%
Richest wealth quintile	Treatment: 46% Comparison: 2%	Treatment: 39% Comparison: 36%
H/h higher education	Treatment: 21% Comparison: 4%	Treatment: 17% Comparison: 17%
Outcome (diarrhea incidence children<2)	Treatment: 18% Comparison: 23% OR = 1.28	Treatment: 15% Comparison: 23% OR = 1.53

Strengthening weak designs

- Triangulate
  - Different impact estimates from different sources
  - Qualitative sources
- Use theory of change to think who benefits and how
- Check the causal chain

New bridge	→	Migration and commuting Access to markets Increased goods availability	→	Higher income Lower prices

Example of Andhra Pradesh Irrigation

- Triangulation

    – Own survey double difference

    – Government data on irrigated and unirrigated *mandals* in treatment districts

    – Baseline report

    – Expert opinion

    All agreed 2-2.5 t/ha without

    and 4-4.5 t/ha with cf 7 t/ha

    with in appraisal report

- Causal chain

    – Construction delays (20 years before any water received)

    – Interrupted water supply

And what about large scale single investments e.g. ports and major bridges?

- Is the impact question the most important one? (also quality of construction, cost-effectiveness etc.)

- Will have made benefit estimates for ex-ante cost-benefit analysis, can test these with 'best available double difference' (going beyond before versus after)

- May well need computable general equilibrium analysis

In summary

- We need impact evaluation to properly allocate resources and justify investments

- Randomization is often possible
  - Of the intervention itself
  - Of a related policy issue
  - Or using an encouragement design

- Non-experimental methods will otherwise often serve

- If weak designs, buttress them

- And can apply these principles to large-scale infrastructure