Estimating Causal Effects in Evaluation – Opportunities and Challenges

Barbara Befani

XII International Evaluation Conference, Wroclaw 21-22 June 2017

European Funds
Technical Assistance

MINISTRY OF ECONOMIC DEVELOPMENT

European Union
Cohesion Fund

Outline

- (Multiple) Causal questions
- (Multiple) Approaches to and Models of, **Causal Inference**
- Why did I attend this conference? How can I answer this question using different models of causal inference?
- What kind of causal evaluation questions can I answer?
- Which methods allow me to answer which questions?
- What do these methods require?

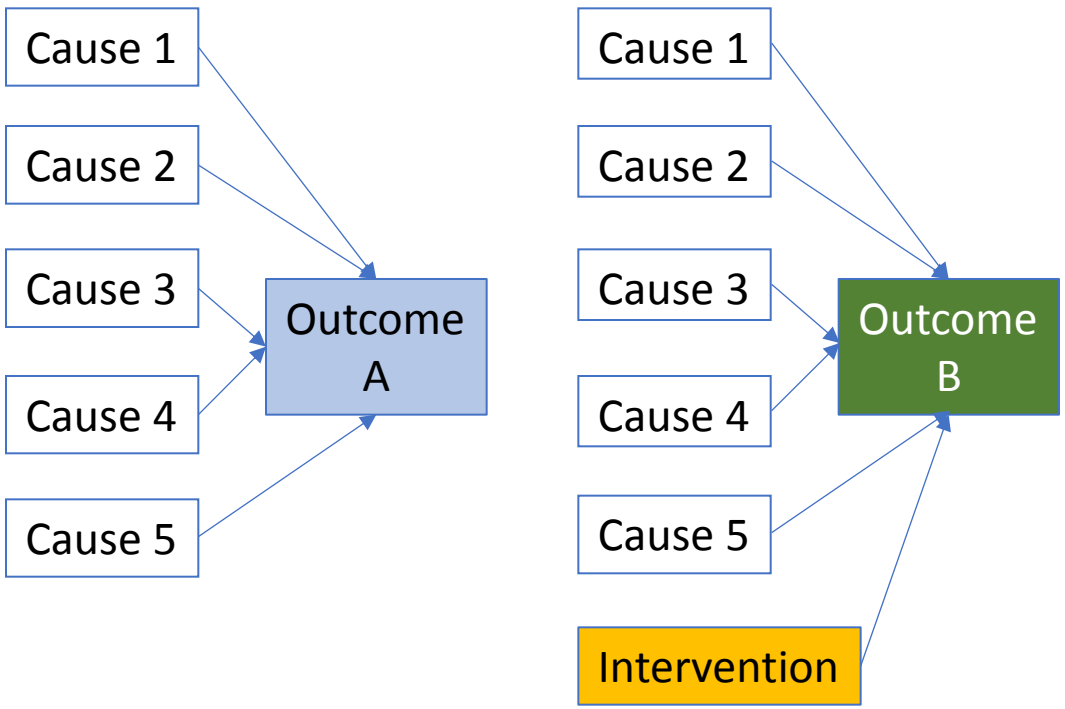
Did the intervention make a difference?

- **That's the overarching impact question**
  - *Which factors made a difference for the outcome?*
- **Counterfactual Analysis & Correlational Studies**
  - *How much of a difference did the intervention (or other factors) make?*
    - *usually on average*
- **Configurational / Multiple-Conjunctural Causation**
  - *Did the intervention (or other factors) make a difference, for whom and under what circumstances?*
    - *not on average, whether*
- **Generative / Mechanism-Based Causality**
  - *How did the intervention make a difference?*
  - *What is it in the intervention that made it (not) work?*
- **We'll see some of the requirements later**

Counterfactual Analysis

- Underpins experiments and quasi-experiments:
  - Mostly (and wrongly) considered "the only rigorous way" to establish causality
  - is also part of configurational approaches
- Literal meaning: comparing observations with assumptions
  - on "what would have happened without the causal factor"
- Based on **Mill's Method of Difference (MoD)**
- The **one difference** in the causes must be responsible for the difference in the outcomes.
- <u>How much</u>, but also <u>whether</u>

MILL'S METHOD OF DIFFERENCE



Counterfactual Logic/Method of Difference

- **Why did I attend this conference?**
- How do I know what is the real "cause"/explanation?

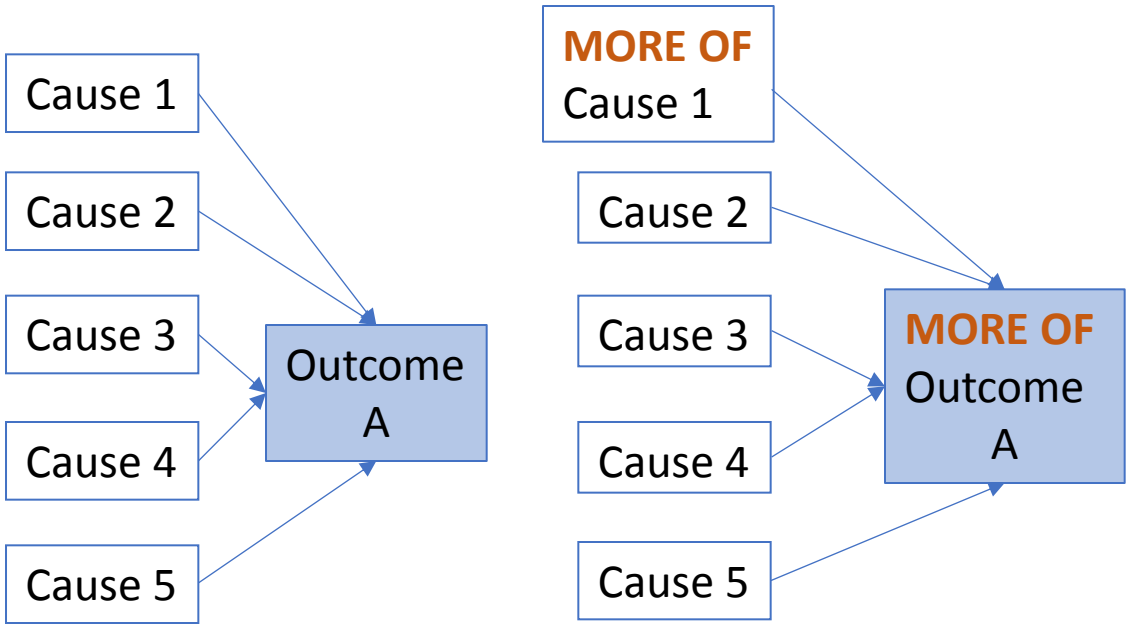
- With **Counterfactual Logic**/**Method of Difference/case-control**, I would reach this conclusion by examining the differences between this conference and all the other conferences I did **not** attend.
- Ideally I would find <u>one similar conference, one difference only</u>.

"same location, same cost, the timing also worked for me. The only difference is that I wasn't interested, so (that must be why) I didn't attend."

Correlational Studies

- Observational (no counterfactual).
- Regression Analysis, Econometric Models.
- Based on **Mill's Method of Concomitant Variation.**
- If change in one factor results in change in the outcome, and all other factors stay the same, the outcome is causally related to the factor.
- Correlation =/ Causation.
- Proportionality of change: <u>how much, not whether.</u>

Mill's Method of Concomitant Variation



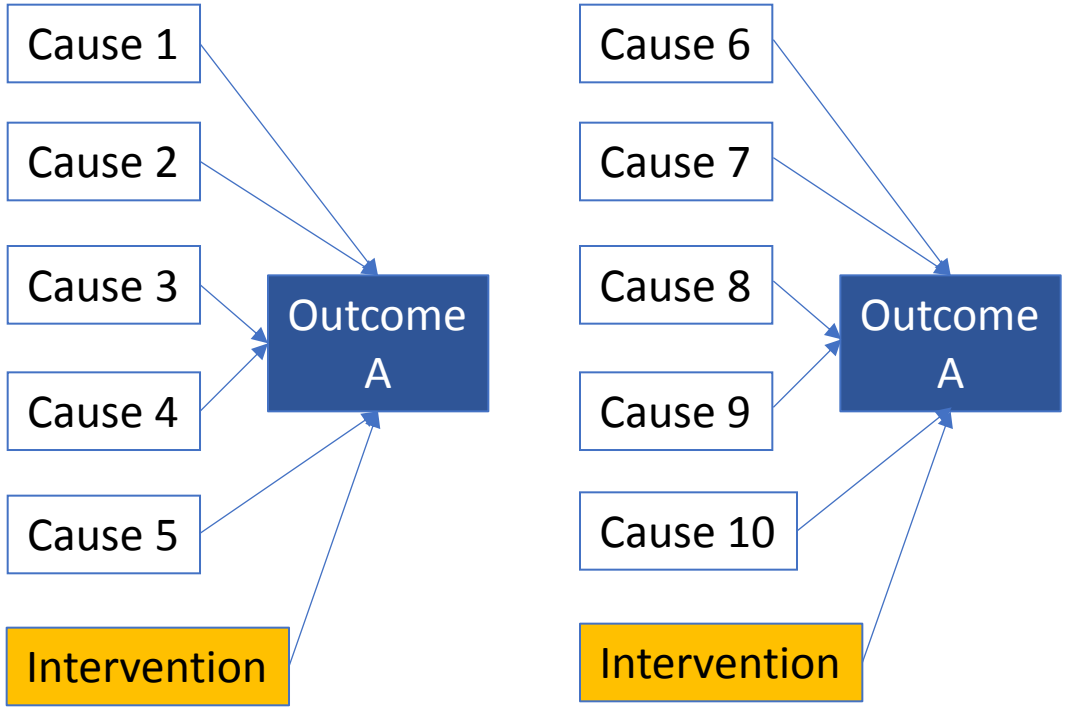
Correlational Logic/Mill's Method of Concomitant Variation

- **Why did I attend this conference?**
- How do I know what is the real "cause"/explanation?

- With **Correlational Logic/Method of Concomitant Variation**, I would measure the net effect of a series of factors that make me attend conference.
- "In the model that best explains why I attend conferences, "interest" has the largest effect on the outcome".
- "the net effect of the intervention is […]".

Mill's Method of Agreement (MoA)

- Used mostly in synthesis studies
  - What do different cases have in common?
- Also part of configurational approaches.
- Comparing **diverse situations with the same outcome.**
- Spotting limited similarities (instead of limited differences) in the causes, ideally just one.
- The **one similarity** must be the cause of the consistently observed outcome.
- <u>Whether, not how much.</u>



Mill's Method of Agreement (MoA)

- **Why did I attend this conference?**
- How do I know what is the real "cause" / explanation?

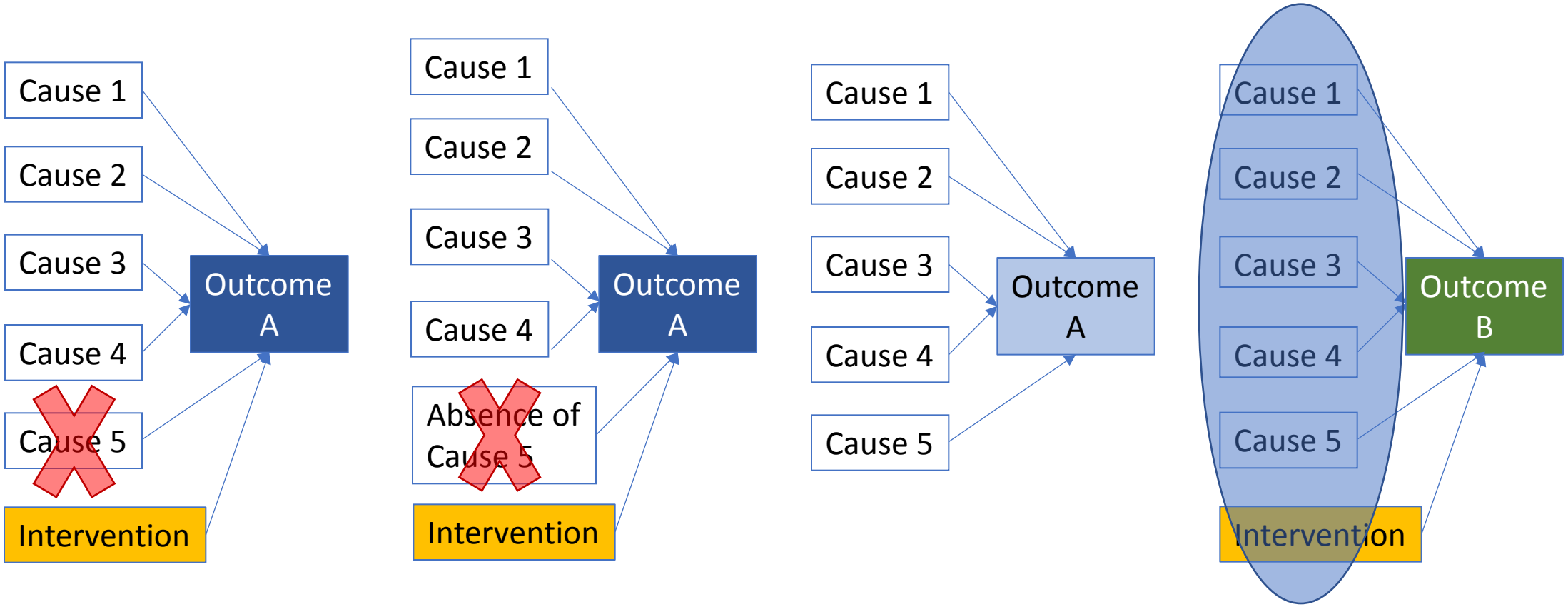
- With the **Method of Agreement**, I would compare the circumstances of all the conferences I ever attended and hopefully find one common element; e.g. that the only constantly present factor was "interest".
- "Different locations, different costs, sometimes I had more time to attend, other times less. But my interest was always, consistently, high".

Single-cause VS. Multiple-cause

- So far we have tried to attribute the effect to one cause.
- Or isolate the specific, net effect of one single cause.
- You can still call these multiple-cause frameworks in a way:
  - they consider different causes
  - They attribute causality through the successive elimination of causes
- HOWEVER **multiple-cause frameworks** properly refer to multiple causes being considered a "**package**", affecting the outcome altogether.
- It's the combination that has (consistent, identifiable) causal power.
- <u>The same single cause can have different effects</u> depending on which other causes it is combined with (depending on the <u>context</u>).
- "Context" makes a difference; changes the "causal power" of factors.

Configurational/Multiple-Conjunctural Causation

- It's a **mix** of Mill's Methods of Difference and Agreement.
- Applied to combinations of causes, rather than single causes.
- Same outcome VS. Different outcome.



Configurational Causality
a.k.a. Multiple-Conjunctural Causality (MCC)

- **Why did I attend this conference?**

- In all previous cases I was looking for “the real cause” or the “one best explanation” of my willingness to attend the conference.
- There might be **more that ONE cause** why I decided to attend and in most cases it’s hard to establish the unique contribution of each cause to my decision to attend:

  “because I was interested, had time available, and I could afford the costs”.

Configurational Causality
a.k.a. Multiple-Conjunctural Causality (MCC)

- I start comparing the configurations explaining why I did or did not attend conferences in the past:
  1. I was interested and had time => I attended
  2. I was interested but didn’t have time => I didn’t attend
  3. I had time but was not interested => I didn’t attend
  4. I was not interested and didn’t have time => I didn’t attend
- By comparing 3 and 1, I discover that interest made a difference when I had time:
  – The addition of “interest” made me attend.
- By comparing 4 and 2, I discover that interest didn’t make any difference when I didn’t have time:
  – With the addition of interest, I still did not attend.

Configurational Causality
a.k.a. Multiple-Conjunctural Causality (MCC)

- Having time made a difference when I was interested (2 to 1).
- Having time didn't make any difference when I wasn't interested (4 to 3).
- **Different factors make the difference under different circumstances:**
  - We don't seek the one "real"/best/average cause, but different causes that make a difference under different circumstances.
- **Conjunctural**: **causes act in packages, not by themselves**. They might have a unique contribution, make a difference, but not in general; more likely in a specific context (in combination with specific other causes).
  - In counterfactual thinking, there is only one cause which makes the difference, on average

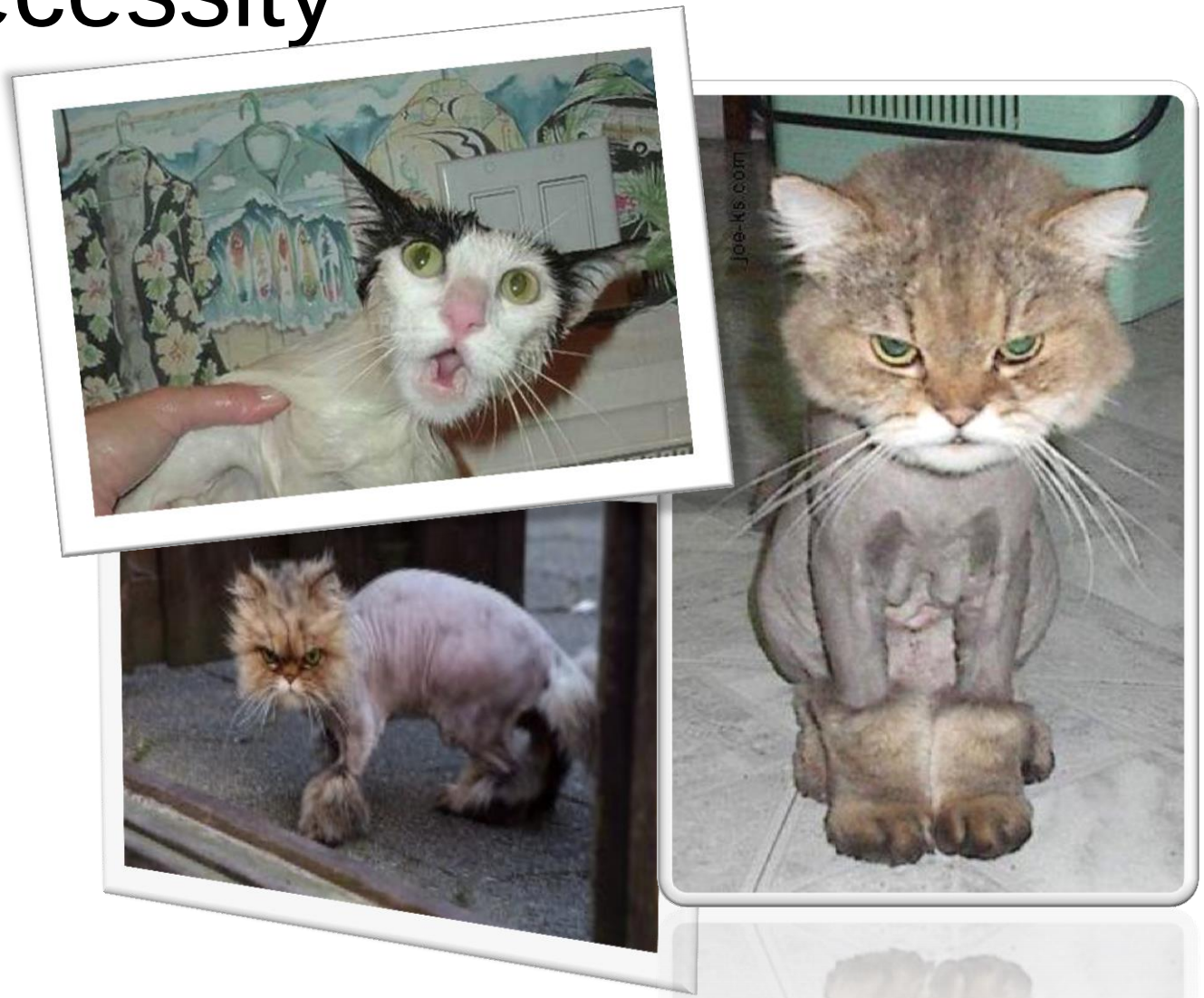
Necessity Without Sufficiency

- Which ingredients are really necessary/we can't do without, for success?
  - Is the intervention (or another factor) necessary for a positive outcome?
  - "Conjunctural"
- **B requires A.**
- **but A alone does not lead to B**.
- "All cases eventually found to be successful implemented the intervention; but there were cases where the intervention was implemented that did not become successful."

"You can lead a horse to water but you can't make it drink"

Sufficiency Without Necessity

- Which combinations of ingredients/factors are good enough?
  - What are the different ways/contexts the intervention can achieve the outcome?
  - (multiple/equifinality)
  - **A leads to B but**
  - **absence of A does not lead to absence of B**.
  - “Every time the intervention was implemented correctly, success was observed; but some cases were successful even without the intervention”.



”There are many ways to skin a cat”

Case ID	Measures adopted to withstand decrease of demand in times of crisis			Success?	Risk level of the combination
	Decrease prices (PRICE)	Set up Promotional offers (PROMO)	Decrease costs (COSTS)		
A	1	0	1	1	Low to Medium
B	1	0	1	1	Low to Medium
C	0	1	0	1	Low to Medium
D	0	1	1	1	Low to Medium
E	0	1	1	1	Low to Medium
F	0	1	1	1	Low to Medium
G	0	0	1	1	Low to Medium
H	0	0	1	1	Low to Medium
I	0	0	1	1	Low to Medium
J	**1**	**1**	**1**	**0**	**High**
K	**1**	**1**	**1**	**0**	**High**
L	**1**	**1**	**1**	**0**	**High**
M	0	0	1	0	Low to Medium
N	0	0	1	0	Low to Medium
O	0	0	1	0	Low to Medium
P	0	0	1	0	Low to Medium
Q	0	1	1	0	Low to Medium
R	0	1	1	0	Low to Medium
S	**0**	**0**	**0**	**0**	**High**
T	**0**	**0**	**0**	**0**	**High**
general average	0.25	0.45	0.85	0.45	
correlation with outcome	-0.058	-0.010	0.099		
Averages for the successful	0.222	0.444	0.889	1	
Averages for the unsuccessful	0.273	0.455	0.818	0	
Average difference	-0.051	-0.010	0.071	1	
Regression coefficients	-0.099	-0.007	0.167		

Findings from the sufficiency analysis

Combination of Strategies	Success?	Number of cases covered	% of cases covered
PRICE*promo*COSTS	Y	2	10%
price*PROMO*costs	Y	1	5%
price*PROMO*COSTS	Y	3	15%
price*promo*COSTS	Y	3	15%
PRICE*PROMO*COSTS	N	3	15%
price*promo*costs	N	2	10%

Young Farmers' Strategies to demand crisis

- How do farmers withstand crisis of demand? Three options:
  1. Lower prices
  2. Lower costs
  3. Launch promotions

- The winning strategies are balanced/mixed; implement some options but not others.
- The losing strategies are extreme; they either implement all options or none.
- Correlational analysis does not indicate the above; but merely that there is no association between any of the single strategies/options and the outcome.

*Befani, B. (2013)"*Between Complexity and Generalisation: Addressing Evaluation Challenges with QCA"*, Evaluation 19(3)*

Results of the regression analysis with a triple interaction effect

Variable	Regression Coefficient	Standard Error	t	P>|t|	95% Confidence Interval	
Var1	0.6356589	0.3720715	1.71	0.108	-0.1573927	1.428711
Var2	0.3255814	0.2472856	1.32	0.208	-0.2014954	0.8526582
Var3	0.1395349	0.3028618	0.46	0.652	-0.5059997	0.7850695
Var123	-1.325581	0.4937118	-2.68	0.017	-2.377903	-0.2732595
constant	0.2248062	0.2825508	0.80	0.439	-0.3774365	0.8270489

Generative/Mechanism-Based Causality

- Correlations and associations are not good enough
- Open the "black box" and investigate the "inner workings" that "generate" the effect
- High degree of precision is required
- "Magnifying lens"
- Ideally we want to observe the effect while it is being "caused"
- If not possible, we seek evidence that a specific process took place…

©Peter Chong

Generative/Mechanism-Based Causality

- **Why did I attend the conference?**
- How do I know what is the real "cause"/explanation?

- With **Mechanism-Based or Generative Causality**, I would try to understand the in-depth reasons that ultimately made me attend:
  - *"attending the conference allowed me to stay up to date with the latest developments in this topic (which I need because […]), network with other professionals interested in the same topic, develop new ideas for my current work, etc. all things I was very interested in".*
  - **Fine-grained narrative.**

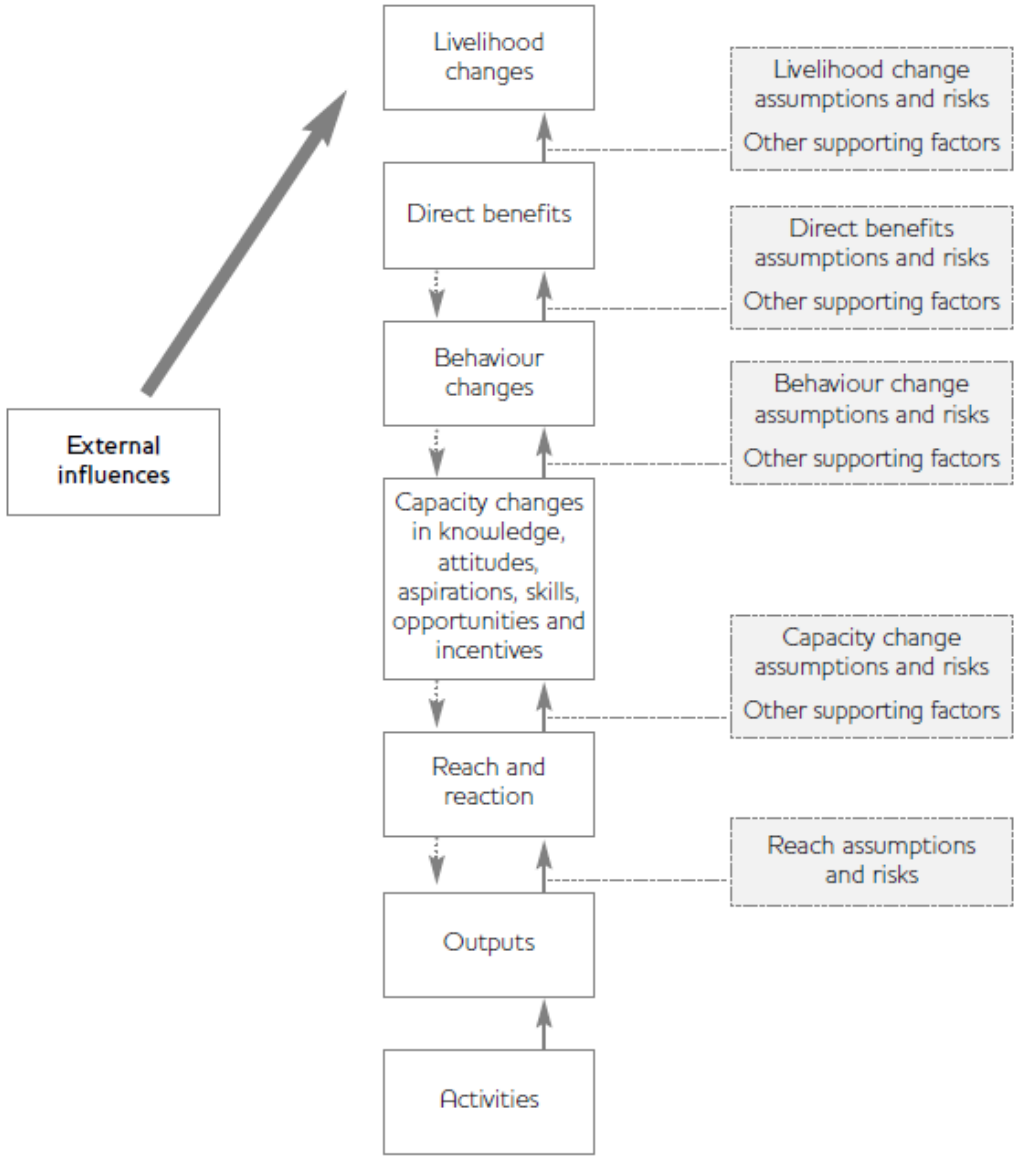
Different forms of Theory-Based Evaluation

- **Contribution Analysis**
- **Realist Evaluation**
- **Systems-Based Evaluation**
- **Process Tracing**

- The opportunity to study cases in-depth (immerse yourself in the case) and represent intermediate outcomes in a causal chain, together with risks and assumptions (**CA**).
- Analytical skills to understand the role of the context (for **RE**).
- Ability to describe complex relations and incorporate different perspectives (**SBE**).
- Ability to measure/assess the “probative power” of given pieces of evidence for your hypothesis (**PT**).

Contribution Analysis

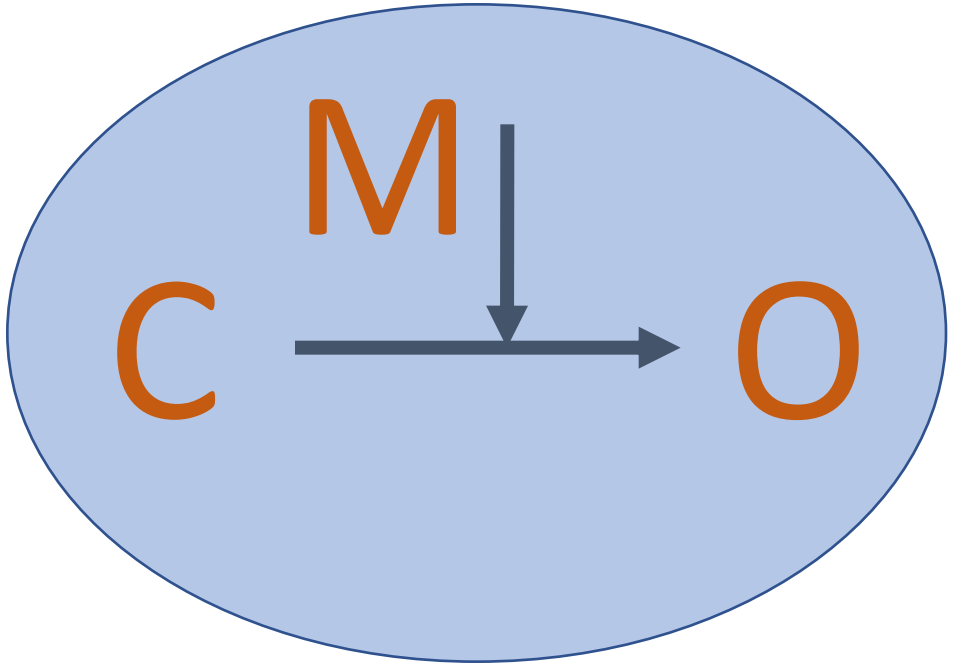
- A "contribution story".
- A causal chain.
- Each step is uncertain, comes with risks and assumptions.
- Each step is necessary for the outcome to occur:
  - That's the assumption under test.
- The whole "package" of steps is sufficient, for the outcome to occur.



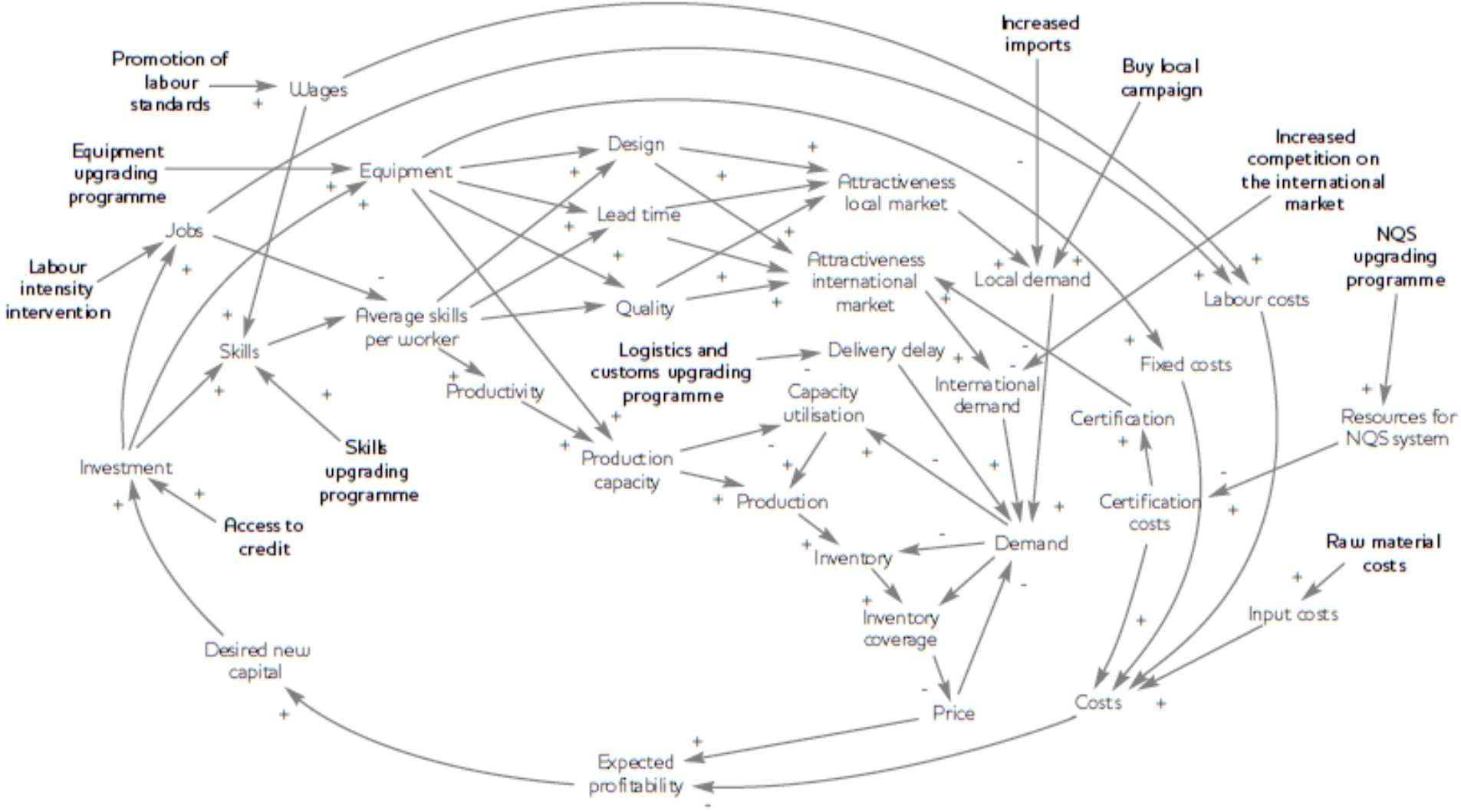
Befani & Mayne 2014, *IDS Bulletin* 45.6

Context-Mechanism-Outcome Configuration

- Detailed explanation of why each arrow holds.
- Explanation of the behaviour of specific stakeholders:
  - Thinking, decision-making, action
  - “The girl did not attend school because her parents did not allow it”
- On the basis of a context with specific resources/opportunities/ constraints:
  - Financial benefits, skills, social rewards, institutional structures, prestige, anything that constitutes an incentive or an obstacle for a specific behaviour or decision.

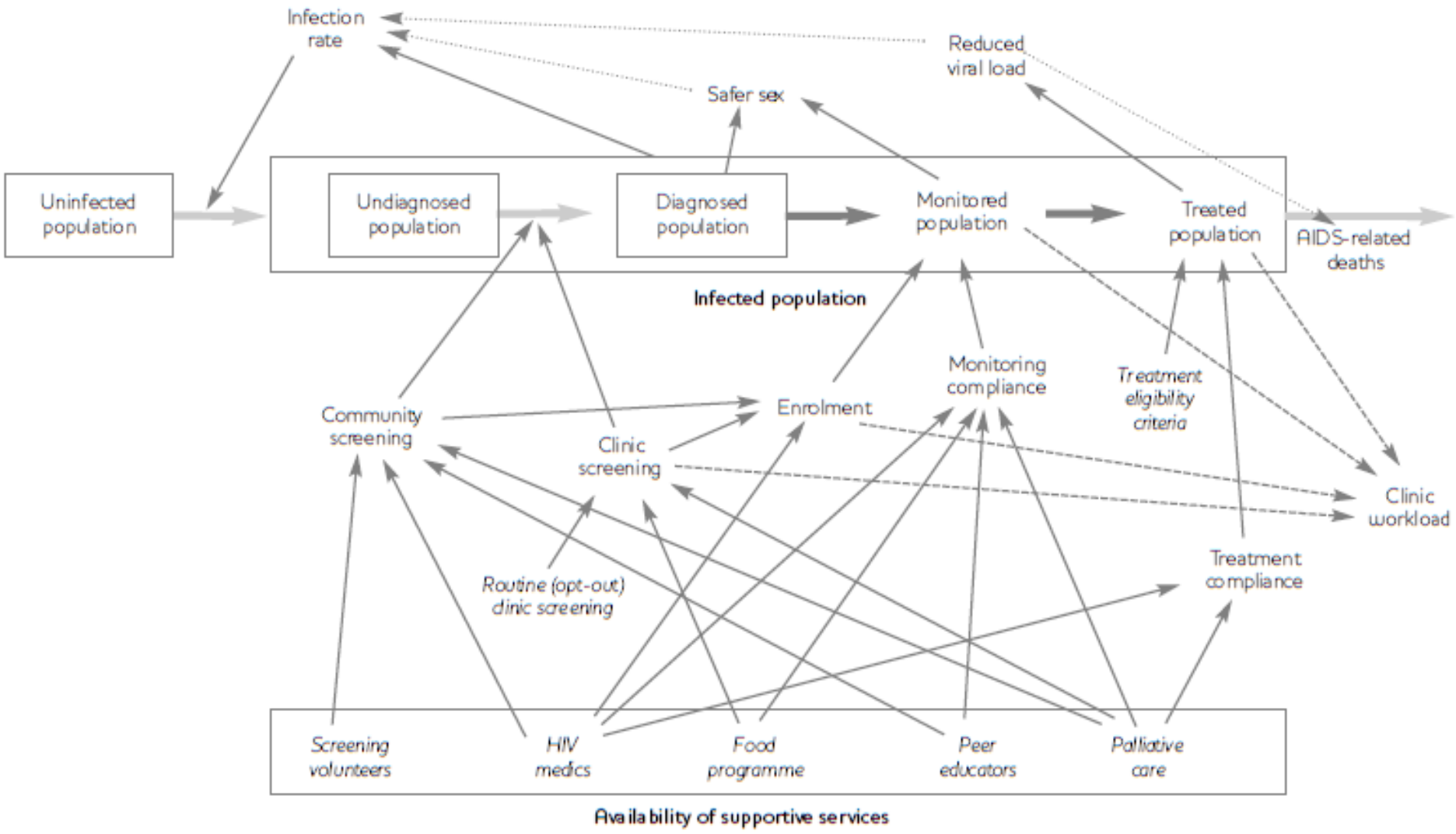


A complex system… the leather shoes sector in Ethiopia



Derwisch & Loewe (2015) “Systems Dynamics Modelling in Industrial Development Evaluation” *IDS Bulletin* 46.1

The Health System in Zambia… and how it can reduce HIV deaths



John Grove (2015) "Aiming for Utility in 'Systems-based Evaluation': A Research-based Framework for Practitioners" *IDS Bulletin* 46.1

What happens when we can't directly observe the mechanism at work?

- The existence of the mechanism is an "ontological reality".
- The hypothesis that the mechanism exists is an "idea in your head"; you have a degree of confidence that the mechanism exists or not
- Evidence/empirical observations are facts that can support or weaken your confidence in your idea (that something is the case).
- Ways to assess the strength of the evidence to prove/disprove your hypothesis.

Befani & Stedman-Bryce (2016) 'Process Tracing and Bayesian updating for impact evaluation', *Evaluation*, currently available "OnlineFirst"

There is no single or best way to infer causality

- Neither attribution nor contribution: causality is "**inferred**".

- Causal inference frameworks should **minimise "inferential error"**
  - The only general recommendation you can give: it's not just about the counterfactual, the Method of Difference is only **one** of Mill's Methods and Mill's Methods are just **one** group of causal models! (the single-cause group).

- They can be **combined**; NO NEED to select ONLY ONE:
  - They all can be applied to the same case/intervention.

Choice is based on:

- **Causal QUESTIONS:**
  - How much
  - Whether
  - On average or In which context (for whom, under what circumstances)
  - How/Why

- **REQUIREMENTS:**
  - How many **cases**? (comparative, cross-case, within-case)
  - Field **expertise** and theoretical **skills** (how good a theory can you build? How well can you understand the context?)
  - How plausible of a **counterfactual** can you build? (What would have happened without the intervention?)

- Does not end here: METHODS underpinned by the same CAUSAL FRAMEWORK come with their own selection criteria….
  - see *Befani, B. (2016)* "Choosing Appropriate Evaluation Methods: a Tool for Assessment and Selection", *Bond*