UPO
University of Piemonte Orientale

Master Class

Why, When and How to Estimate (Causal) Effects on Territories CS-SEQDD Method and Other Options

Daniele Bondonio(*)

(*) *Professor of Statistics for Economic, University of Piemonte Orientale (Ph.D. Carnegie Mellon University)*
E-mail: daniele.bondonio@uniupo.it.   https://upobook.uniupo.it/daniele.bondonio

Evaluation Helpdesk Summer School, Prague 23-24 June 2022

Causal-Effect Estimation with Statistical/Econometric Analysis on Ex-Post Data

- Two important aspects to consider when estimating casual effects with statistical analysis on ex-post data:

  -Advanced analysis should **be designed also to estimate heterogeneous causal effects, if this is related to relevant Evaluation Questions based on ToC**= different causal contributions may be estimated for different types of beneficiaries, territories, socio-economic contexts, intensity of support, etc…

  - Causal effect/impact/contribution = difference between Δ Y with intervention (or a particular element of it, territorial context, intensity, a particular composition of beneficiaries etc) and the ("counterfactual" or "spontaneous") ΔY what would be achieved without the intervention (or without a particular element of it, territorial context, intensity etc.), holding everything else constant

What is the relation with ToC or TBE?

- Estimating casual effects with statistical analysis on post-implementation data should not be seen as unrelated or alternative to ToC or TBE (otherwise it produces black-box evaluations with uniform average effects for every beneficiary, context, intensity and type of treatment that may be not informative)

A)

Counterfactual impact evaluation /experimental designs = black box eval. do not explain why, how when it works

TBE, discovers program mechanisms, useful evidence no black box

Evaluation project

B)

II) Statistical/econometrics analysis on ex-post data tests some key findings from I), if data and implementation conditions are suitable

I) ToC and TBE discovers program mechanisms, how, what and when (contextual element) works/doesn't work

Evaluation project

- **Advanced scientific research (in most fields) works as in B) not A)!:** statistical analysis on post implementation data is aimed at testing specific theories, mechanisms, influence of contexts etc. (the ones that are most relevant and/or uncertain and that can be addressed given the available data)

Why and When to Estimate Effects on Territories

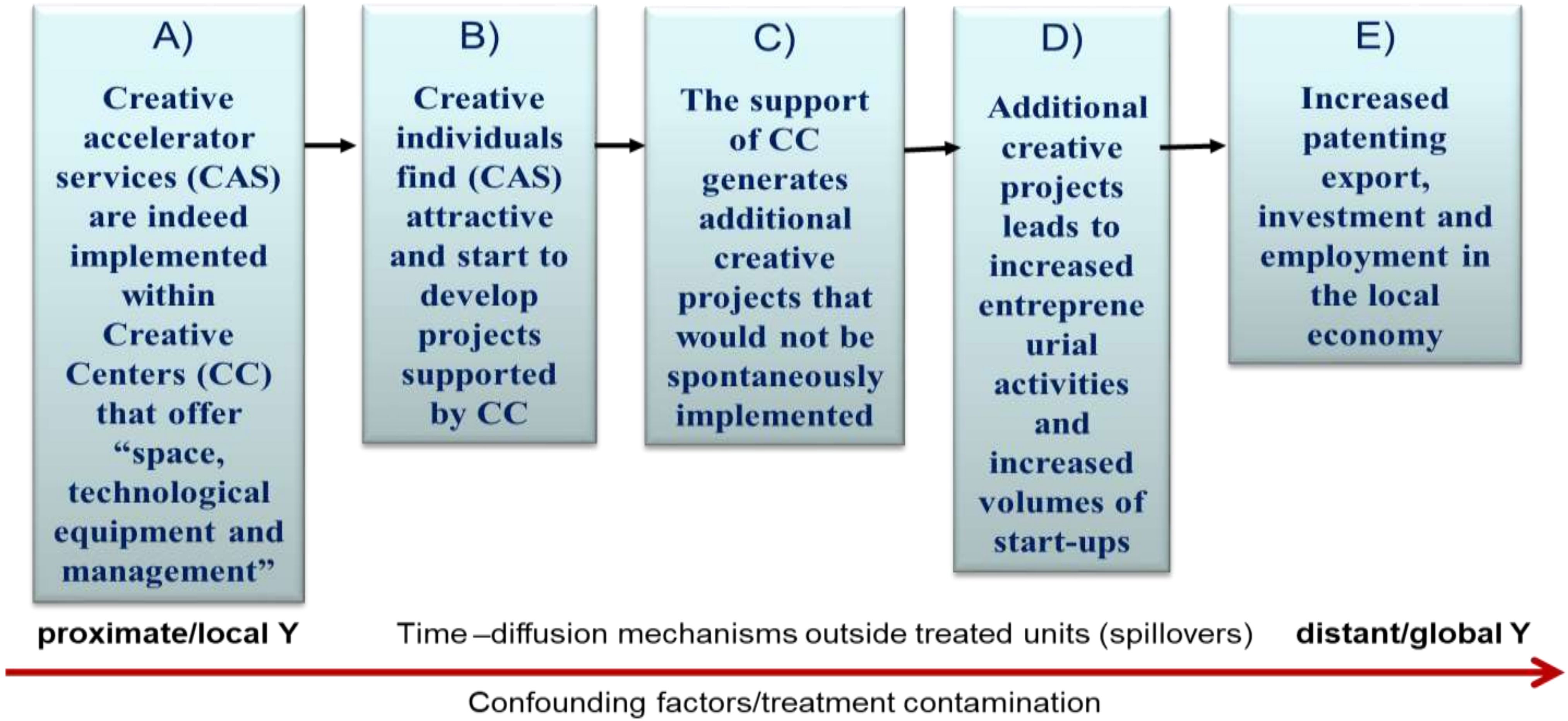
- Following **intervention logic/ToC**, evaluations of causal effects with statistical analysis can focus either on **proximate/local** or more **distant/global** result indicators **(Y)**:

Proximate	= close (causal links and temporally) to initial implementation
Local	= affected primarily by a single OP programme
Distant	= far away (causal links and temporally) from initial implementation
Global	= affected by multiple OP programmes

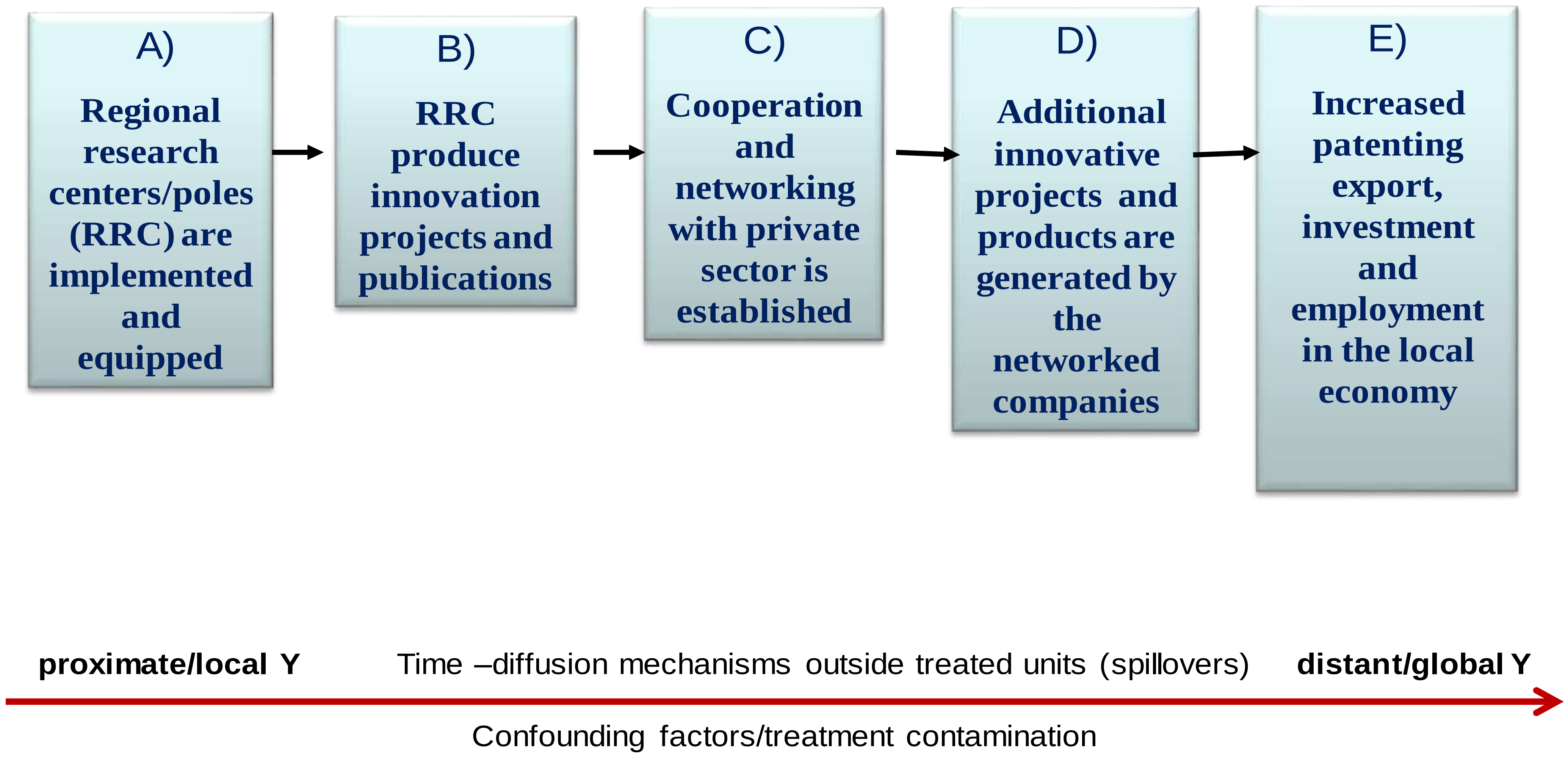
- The rationale to estimate causal effects with statistical analysis on ex-post data on territories is linked to estimating impacts on distant/global result indicators (embedding ultimate desirable outcomes of the programme interventions, close to Thematic Objectives)

Chain of causality links from Intervention-Logic / ToC

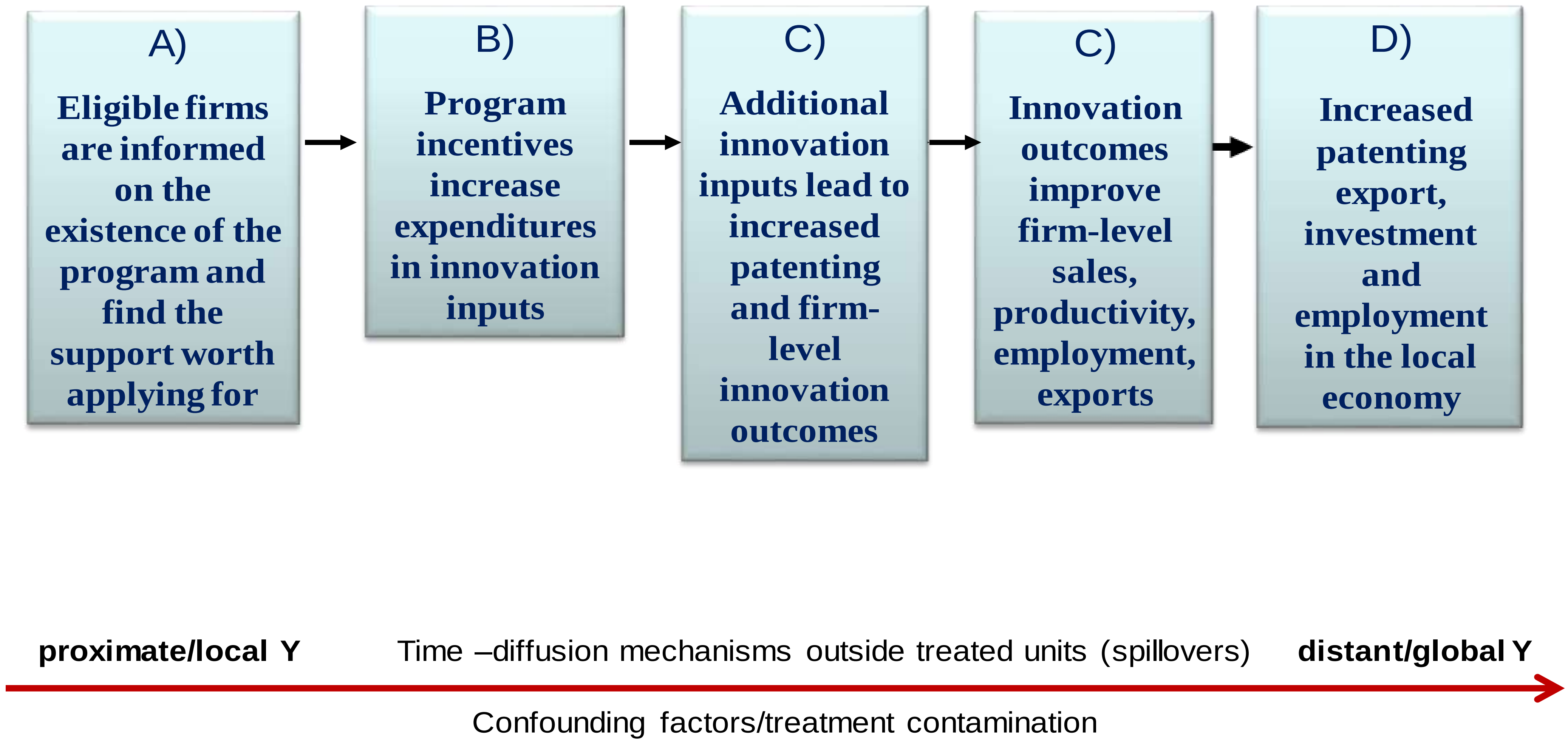
Example I (creative accelerator services)



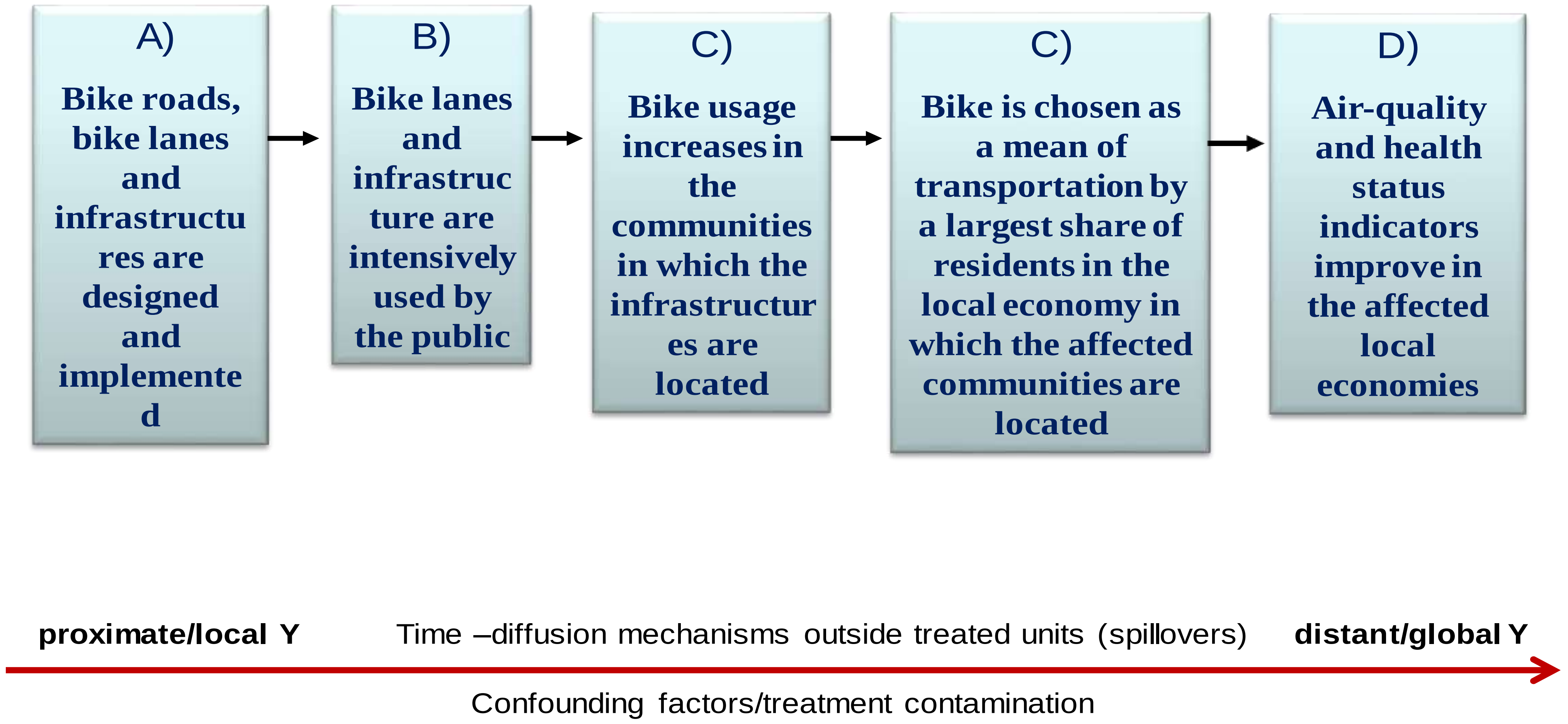
Example II (regional research centers)



Example III (support to SME innovation)



Example IV (bike roads, lane, infrastructures)



Focus on Distant/Global Y Leads to Consider Territorial Effects

- When the focus is on **distant/global** Y, close to TO goals, it is necessary to:
  -**pool together all OP projects** for which the intervention logic/ToC indicates a possible effect on a same distant/global Y
  -**aggregate the data** at the level of the territorial units that (according to ToC) contain all the diffusion mechanisms from the first-line direct-beneficiaries to other units of observation (generating treatment spillovers)

- For **distant/global Y,** statistical analyses by means of ex-post micro-data (e.g. firm-level-, research-center-level data etc..) **do not work well**: in the medium long- run non-treated units become also affected by the programme and they cannot detect anymore what would have been the spontaneous change due to other factors

Example:

- A maximally-effective SME support programme, **in the long run**, could generate large positive spillovers within a same industrial district (i.e. positive effects of the support extended also to non-treated firms). In this scenario a rigorous quasi experimental design, applied to a distant/global Y, would generate a zero impact estimate! (no difference between Y of treated and similar non-treated firms **= treatment contamination bias**)

Which Territorial Level to Consider?

- **Golden rule:**

The Intervention Logic/ ToC should guide the choice of the territorial level so that it does contain most of the diffusion-mechanisms/spillovers from the immediate beneficiaries/treated units to others. **Rule of thumb:**

-limited spillover potential = smaller territorial level

-high spillover potential = larger territorial level

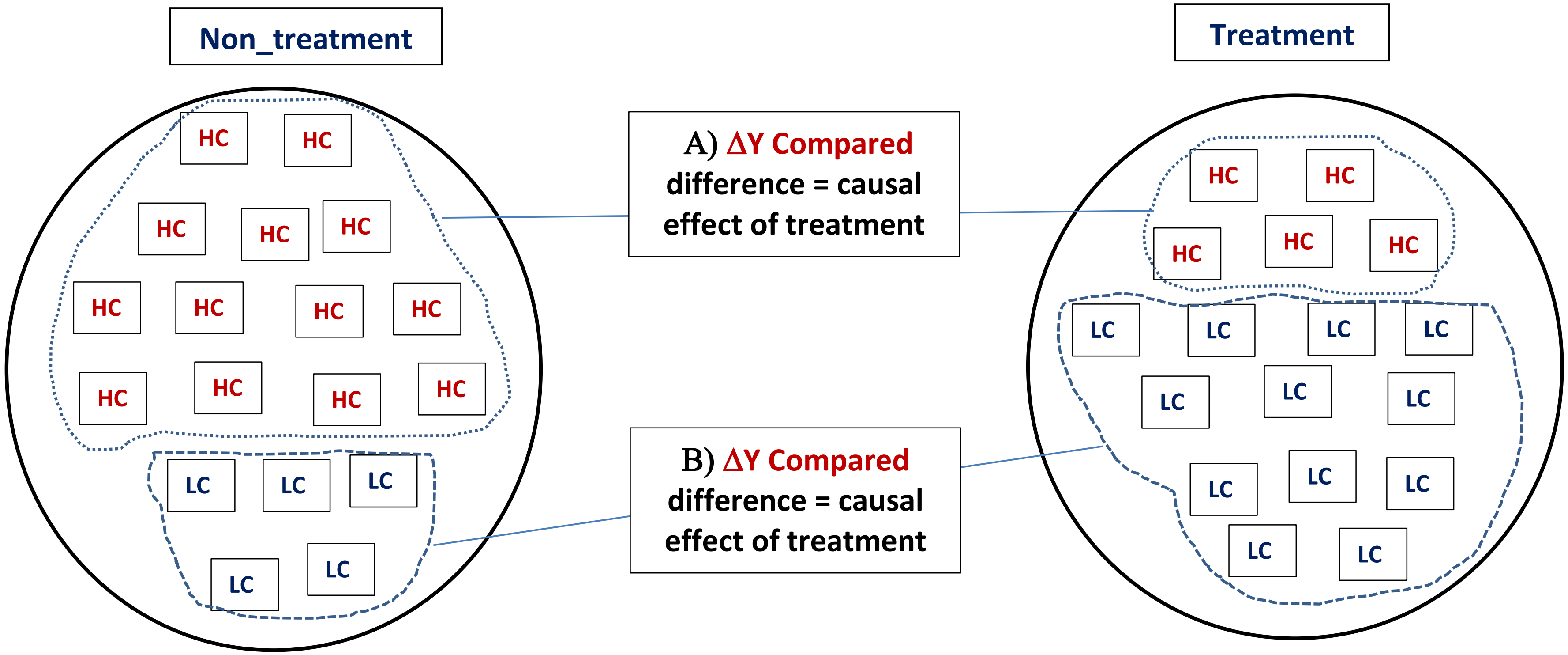
- **Possible hybrid solutions:**

-sometimes (e.g. **SME support**) ToC may indicate that diffusion/spillover effects may occur also at the level of sectors. In this case the data have to be aggregated at an hybrid level: **sector/territory** (e.g. sec_code-1/reg-1, sec-code-2/reg-1, …..sec-code-1/reg-2 etc.). E.g. Bondonio (2006, Provinces/Sectors)

-similar solution for **social/welfare/active labour policies** in terms of Hybrids **group of individuals/territory** (e.g. workers_U35 / reg-1, workers_O35 /reg-1……. workers_U35 / reg-2 etc…

Main Challenge in Estimating Causal Effects

- Estimating causal effects = separating the changes of (Y) caused by the programme interventions vs **spontaneous (counterfactual)** change (caused by other unrelated factors)
- **Key challenge** = how to estimate the spontaneous change. This is normally done by using data on similar non-treated units: if they display a change in Y, this can only be due to spontaneous change (not treatment).



- **If ToC or TBE indicates that high pre-existing research capacity may affect how the OP projects work,** comparing casual effect A) to B) is how to use statistical analysis (heterogenous causal effects) on ex-post data to test predictions from ToC or findings form TBE

The CS-SEQDD Method: Basic Principles and Intuitive Concepts

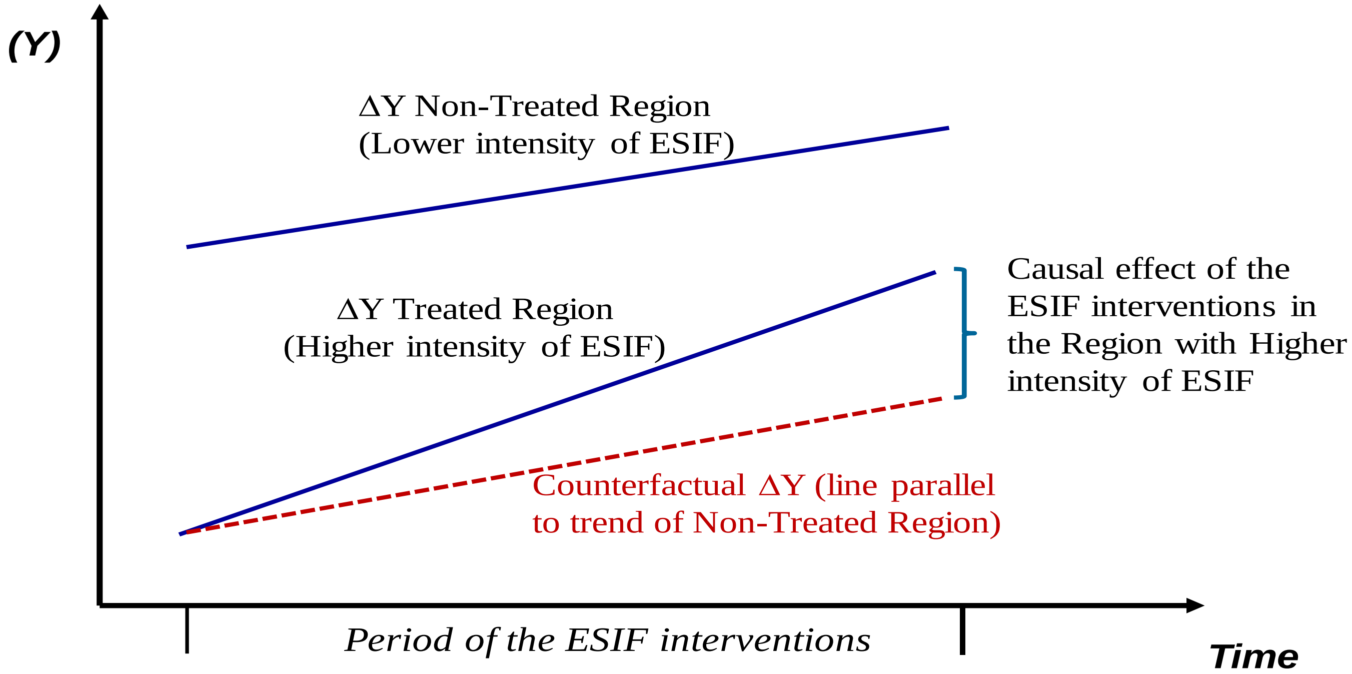
- At territorial level on distant/global Y, **all units are somehow treated**. Standard comparison treated vs untreated not feasible. The intuitive idea behind CS-SEQDD is to exploit the cross-territorial **variation in the intensities** of OP projects in order to assess whether or not the territories with higher intensities also display a larger ΔY, controlling for territorial differences

- This is done by considering **sequential pair-wise difference in difference** between territories: the largest the difference in the intensity of the OP interventions between two territories, the largest should be the difference in ΔY

- If this is the case, spontaneous change is low and most of the **Δ**Y at the territorial level is indeed caused by the OP interventions. If not, most of the **Δ**Y is instead likely to be caused by spontaneous change

How are territorial differences controlled for in CS-SEQDD?

- At the heart of CS-SEQDD are the sequential pairwise difference in difference (DD) that compare the before-after-intervention cross-territorial ΔY
- In the pairwise (DD) comparisons, **pre-post Δ Y** recorded in the low-intensity territory are assumed to be the counterfactual change that would be recorded in the higher-intensity territory in the presence of a lower intensity of treatment
- This assumption requires that the relevant different baseline characteristics of the territories are **fixed effects**: factors that exert a constant over-time effect on the levels of Y. This is also referred to as the "parallel trend assumption":



Fixed Effects /Parallel Trends Assumptions

- Suppose that **Region A** (receiving a low intensity of OP interventions) is structurally different than **Region B** (higher intensity of the OP interventions)
- Example: **Region A** has **higher R&D capacity** than **Region B** (> universities, > n. of existing R&D labs and facilities, > residents with higher education). These structural differences entails that **Region A** tends to have, in any given year, an higher value of Y (e.g. n. of patent applications) than **Region B**
- In standard quasi-experimental counterfactual impact evaluations it would be required to find a **comparison group of other regions** with very **similar structural characteristics of region A and B**, but different intensities of treatment (or no treatment at all)
- With a DD comparison, instead, the way in which the different structural characteristics of the regions are taken into account in the analysis is by means of **transforming the values of Y into changes between the beginning and the end of the OP projects** (e.g. 2014-2020)
- **Rationale of DD comparison:** if the differences between the two regions are structural, these different features are constantly in existence in any given year during the period of observation. For this reason , they are referred to as “fixed effects” (e.g. if one Regions has a larger number of universities this feature tend to be always in existence).
- **Fixed effects** by definition do have an influence on the levels of Y (e.g. yearly number of per-capita patent applications), but they **cannot have an influence on Δ Y between different years** (“fixed effects” are always in existence and therefore they cannot induce a change in Y between different periods)

CS-SEQDD Data Requirements

- The intensities of the OP projects have to be allocated at the territorial level within the programming period

- Data on Y have to cover at least the beginning and the end of the actual implementation on the ground of the projects

- OP projects and the result indicator/s (Y) have to measured in terms of intensities with respect to a same baseline size-indicator

  Example (support to R&D):

  The Intensity of OP intervention can be measured as: (economic value of OP projects)/ (residents);

  While a result indicator can be (number of yearly patent applications) / (residents)

  This size-indicator (in this case, residents or residents with higher education degree) is used to control for scale-effect differences that can lead to obvious different potentials for the absolute changes of Y along the estimation period of interest

- The intensities of OP projects need to have a sufficiently large degree of variation across region

An Example of CS-SEQDD Estimation

- The technical note presents four application examples of the CS-SEQDD estimation procedure *https://ec.europa.eu/regional_policy/en/information/publications/evaluations-guidance-documents/2021/cross-regional-sequential-difference-in-difference-cr-seqdd*
- For ease of comparability, all examples are related OP projects and distant/global Y related to Thematic Objective (TO) 1, *Strengthening research, technological development and innovation*.
- The sample of territories is *N=15 (in the example regions)*, and the available regional-level data concerns
- $\Delta Y_i = (Y_{ipost} - Y_{ipre})$ = Pre-post-intervention change in the yearly number of patent applications per million of residents recorded in region (*i*)
- T_i= Per-capita intensity of the OP interventions pertaining support to R&D

Example I: *Ideal data-availability and strong causal effect of the OP projects.*

Region	Pop.	TO1 OP support (1=€Mil.)	T [Intensity of TO1 support] 1 = (1 €Mil.) / (Mil. Residents)	Y_{pre} 1= No. Pat. Appl. / Mil. Residents	Y_{post} 1= No. Pat. Appl. / Mil. Residents	$\Delta Y = (Y_{post}) - (Y_{pre})$
A	500,000	0	0	65.5	66.0	0.5
B	1,200,000	24	20	58.4	62.8	4.4
C	800,000	36	45	55.3	64.1	8.8
D	2,400,000	120	50	52.3	62.0	9.7
E	3,000,000	165	55	50.1	60.8	10.7
F	1,400,000	86.8	62	48.6	61.2	12.6
G	2,000,000	130	65	53.5	66.7	13.2
H	1,500,000	102	68	52.3	65.7	13.4
I	2,200,000	154	70	55.7	69.8	14.1
L	1,200,000	88.8	74	58.9	73.5	14.6
M	600,000	45.6	76	60.2	75.3	15.1
N	1,400,000	109.2	78	56.4	71.8	15.4
O	2,000,000	160	80	57.3	73.5	16.2
P	1,100,000	93.5	85	60.1	76.9	16.8
Q	1,600,000	137.6	86	56.3	73.7	17.4
Nation	22,900,000	1452.5	63.4	54.8	67.5	12.7

Pairwise DD variations (DDY) between Comparison and Baseline Regions

		Baseline Region (Lower T)														
		A	B	C	D	E	F	G	H	I	L	M	N	O	P	Q
	A	-														
	B	3.9	-													
	C	8.3	4.4	-												
	D	9.2	5.3	0.9	-											
	E	10.2	6.3	1.9	1	-										
	F	12.1	8.2	3.8	2.9	1.9	-									
Comparison Region (Higher T)	G	12.7	8.8	4.4	3.5	2.5	0.6	-								
	H	12.9	9	4.6	3.7	2.7	0.8	0.2	-							
	I	13.6	9.7	5.3	4.4	3.4	1.5	0.9	0.7	-						
	L	14.1	10.2	5.8	4.9	3.9	2	1.4	1.2	0.5	-					
	M	14.6	10.7	6.3	5.4	4.4	2.5	1.9	1.7	1	0.5	-				
	N	14.9	11	6.6	5.7	4.7	2.8	2.2	2	1.3	0.8	0.3	-			
	O	15.7	11.8	7.4	6.5	5.5	3.6	3	2.8	2.1	1.6	1.1	0.8	-		
	P	16.3	12.4	8	7.1	6.1	4.2	3.6	3.4	2.7	2.2	1.7	1.4	0.6	-	
	Q	16.9	13	8.6	7.7	6.7	4.8	4.2	4	3.3	2.8	2.3	2	1.2	0.6	-

1= [No. Pat. Appl. / Mil. Residents] in terms of pairwise Difference-in-difference variation of Y (DDY) between Comparison and Baseline Regions

Pairwise Cross-Regional Differences in the Intensities of OP Interventions (ΔT)

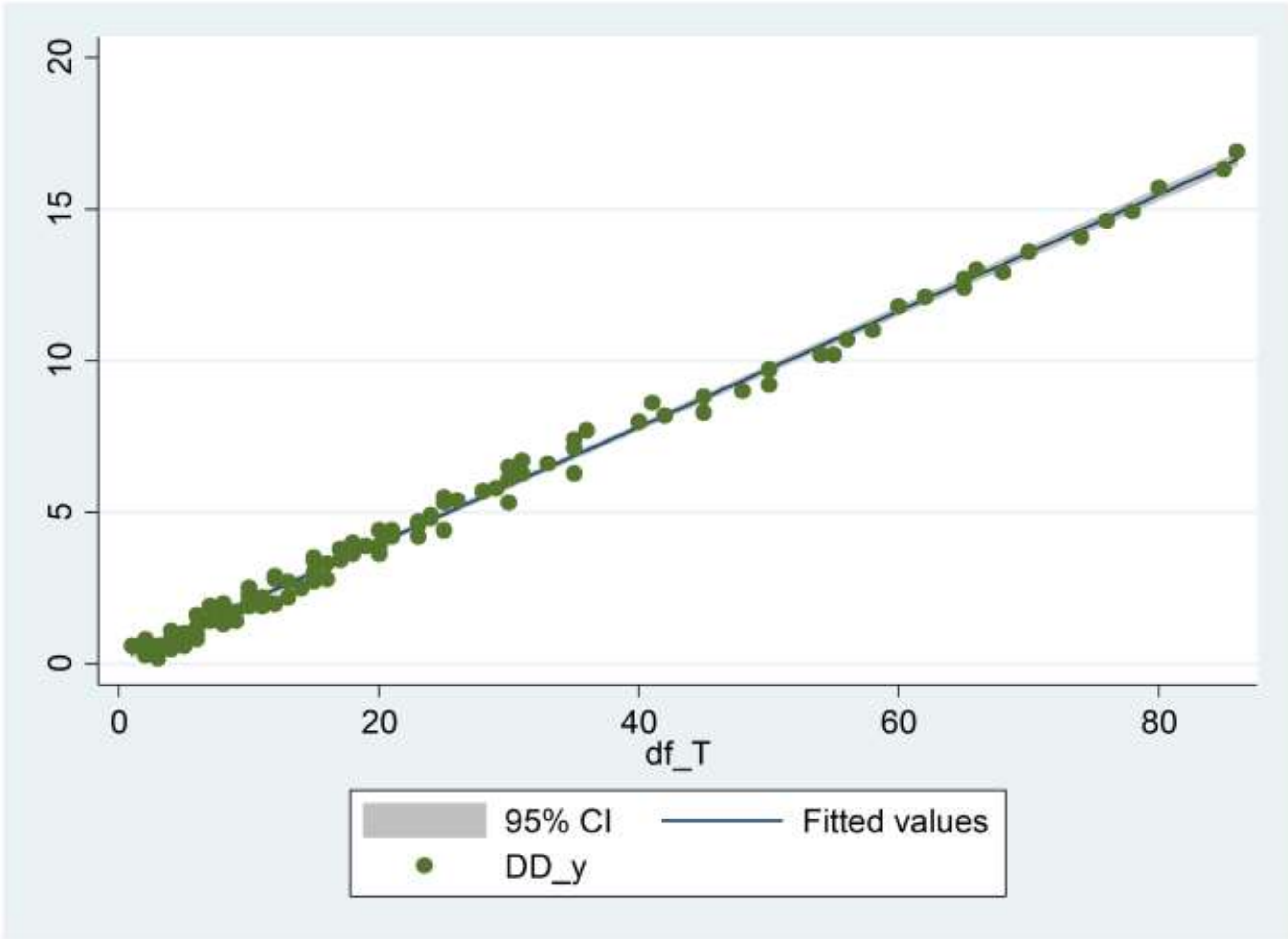
		Baseline Region														
		A	B	C	D	E	F	G	H	I	L	M	N	O	P	Q
Comparison Region	A	-														
	B	20	-													
	C	45	25	-												
	D	50	30	5	-											
	E	55	35	10	5	-										
	F	62	42	17	12	7	-									
	G	65	45	20	15	10	3	-								
	H	68	48	23	18	13	6	3	-							
	I	70	50	25	20	15	8	5	2	-						
	L	74	54	29	24	19	12	9	6	4	-					
	M	76	56	31	26	21	14	11	8	6	2	-				
	N	78	58	33	28	23	16	13	10	8	4	2	-			
	O	80	60	35	30	25	18	15	12	10	6	4	2	-		
	P	85	65	40	35	30	23	20	17	15	11	9	7	5	-	
	Q	86	66	41	36	31	24	21	18	16	12	10	8	6	1	-

1= [€Mil / Mil. Residents] in terms of cross-regional pairwise differences of OP-intervention intensities.

Two-way Scatter Plot Chart

Vertical Axis=Pairwise Cross-Regional Causal Impact Estimations DDY

Horizontal Axis=Pairwise Cross-Regional Variation of Treatment Intensity (ΔT)



- Estimated parameters of the linear dose-response function:

```
Number of obs      =        105
Prob > chi2        =     0.0000
Adj R-squared      =     0.9945
```

	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf. Interval]	
ΔT	.1915361	.0051216	37.40	0.000	.181498	.2015743
α	.1524562	.1015183	1.50	0.133	-.046516	.3514285

- Predicted value ($\widehat{DDY}$) for the nationally-recorded intensity of OP interventions = 0.1525 + 0.1915 * 63.4 = +12.3 yearly patent applications per million of residents caused by an intensity of € 63.4 Million worth of OP interventions in TO 1)

- Set of regional ($\widehat{DDY}$) can also be estimated:

REG B (€20 Mil OP int.)= 0.1525 + 0.1915 * 20 =+4 yearly patent applications per million of residents

…..

REG Q (€86 Mil OP int.)= 0.1525 + 0.1915 * 86 =+16.6 yearly patent applications per million of residents

Example II: *Ideal data-availability and absence of causal effect of the OP interventions.*

Region	Pop.	TO1 OP support (1=€Mil.)	T [Intensity of TO1 support] 1 = (1 €Mil.) / (Mil. Residents)	Y_{pre} 1= No. Pat. Appl. / Mil. Residents	Y_{post} 1= No. Pat. Appl. / Mil. Residents	$\Delta Y = (Y_{post}) - (Y_{pre})$
A	500,000	0	0	65.5	70.0	4.5
B	1,200,000	24	20	58.4	62.5	4.1
C	800,000	36	45	55.3	59.5	4.2
D	2,400,000	120	50	52.3	56.1	3.8
E	3,000,000	165	55	50.1	54.6	4.5
F	1,400,000	86.8	62	48.6	54.7	6.1
G	2,000,000	130	65	53.5	59.1	5.6
H	1,500,000	102	68	52.3	56.6	4.3
I	2,200,000	154	70	55.7	59.9	4.2
L	1,200,000	88.8	74	58.9	65.0	6.1
M	600,000	45.6	76	60.2	64.3	4.1
N	1,400,000	109.2	78	56.4	61.2	4.8
O	2,000,000	160	80	57.3	62.7	5.4
P	1,100,000	93.5	85	60.1	64.4	4.3
Q	1,600,000	137.6	86	56.3	61.0	4.7
Nation	22,900,000	1452.5	63.4	54.8	59.5	4.7

Pairwise DD variations (DDY) between Comparison and Baseline Regions

		Baseline Region (Lower T)														
		A	B	C	D	E	F	G	H	I	L	M	N	O	P	Q
Comparison Region (Higher T)	A	-														
	B	-0.4	-													
	C	-0.3	0.1	-												
	D	-0.7	-0.3	-0.4	-											
	E	0.0	0.4	0.3	0.7	-										
	F	1.6	2.0	1.9	2.3	1.6	-									
	G	1.1	1.5	1.4	1.8	1.1	-0.5	-								
	H	-0.2	0.2	0.1	0.5	-0.2	-1.8	-1.3	-							
	I	-0.3	0.1	0.0	0.4	-0.3	-1.9	-1.4	-0.1	-						
	L	1.6	2.0	1.9	2.3	1.6	0.0	0.5	1.8	1.9	-					
	M	-0.4	0.0	-0.1	0.3	-0.4	-2.0	-1.5	-0.2	-0.1	-2.0	-				
	N	0.3	0.7	0.6	1.0	0.3	-1.3	-0.8	0.5	0.6	-1.3	0.7	-			
	O	0.9	1.3	1.2	1.6	0.9	-0.7	-0.2	1.1	1.2	-0.7	1.3	0.6	-		
	P	-0.2	0.2	0.1	0.5	-0.2	-1.8	-1.3	0.0	0.1	-1.8	0.2	-0.5	-1.1	-	
	Q	0.2	0.6	0.5	0.9	0.2	-1.4	-0.9	0.4	0.5	-1.4	0.6	-0.1	-0.7	0.4	-

1= [No. Pat. Appl. / Mil. Residents] in terms of pairwise Difference-in-difference variation of Y (DDY) between Comparison and Baseline Regions

Pairwise Cross-Regional Differences in the Intensities of the OP Interventions (ΔT)

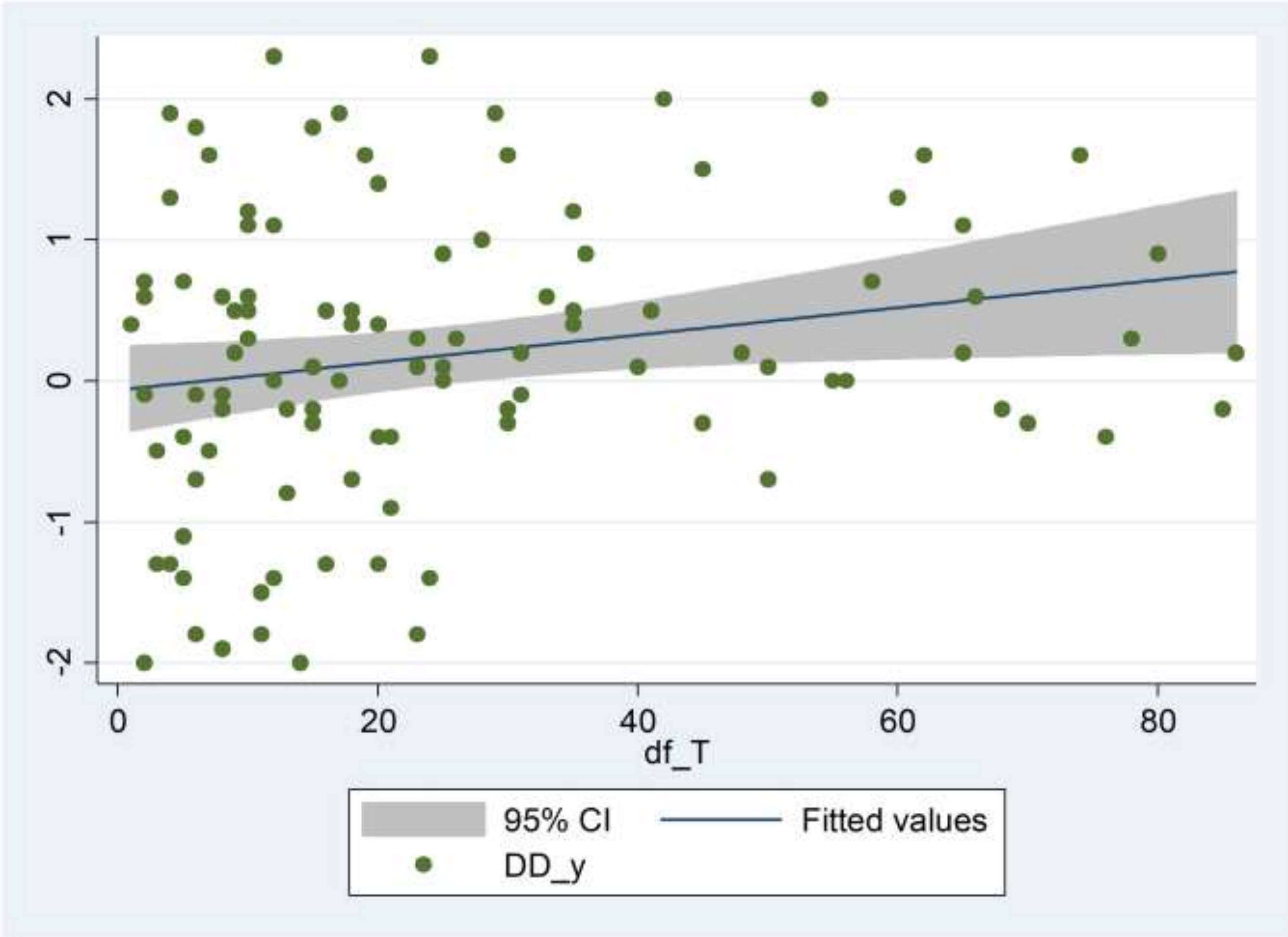
		Baseline Region														
		A	B	C	D	E	F	G	H	I	L	M	N	O	P	Q
Comparison Region	A	-														
	B	20	-													
	C	45	25	-												
	D	50	30	5	-											
	E	55	35	10	5	-										
	F	62	42	17	12	7	-									
	G	65	45	20	15	10	3	-								
	H	68	48	23	18	13	6	3	-							
	I	70	50	25	20	15	8	5	2	-						
	L	74	54	29	24	19	12	9	6	4	-					
	M	76	56	31	26	21	14	11	8	6	2	-				
	N	78	58	33	28	23	16	13	10	8	4	2	-			
	O	80	60	35	30	25	18	15	12	10	6	4	2	-		
	P	85	65	40	35	30	23	20	17	15	11	9	7	5	-	
	Q	86	66	41	36	31	24	21	18	16	12	10	8	6	1	-

1= [€ Mil. / Mil. Residents] in terms of cross-regional pairwise differences of OP-intervention intensities.

Two-way Scatter Plot Chart (Linear Fitting)

Vertical Axis=Pairwise Cross-Regional Causal Impact Estimations DDY

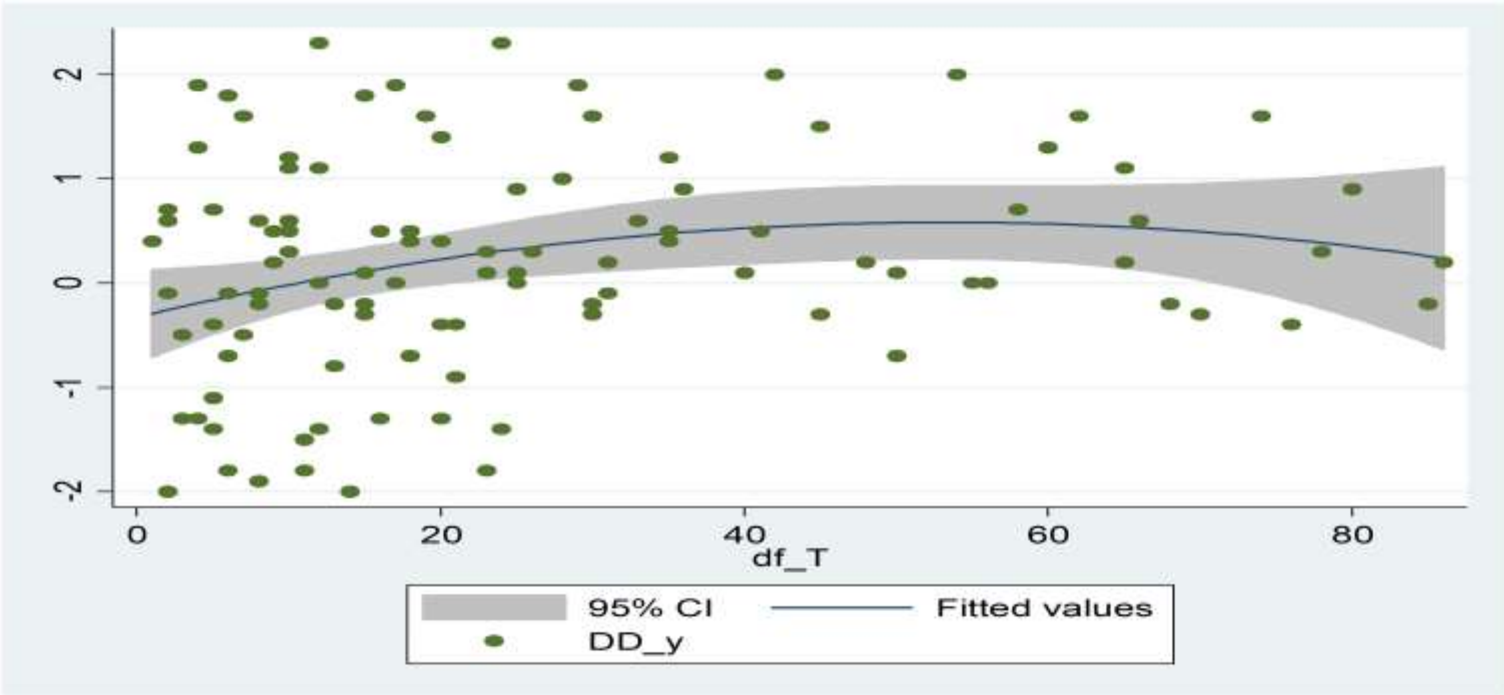
Horizontal Axis=Pairwise Cross-Regional Variation of Treatment Intensity (T)



Two-way Scatter Plot Chart (Quadratic Fitting)

Vertical Axis=Pairwise Cross-Regional Causal Impact Estimations DDY

Horizontal Axis=Pairwise Cross-Regional Variation of Treatment Intensity (T)



- Estimated parameters of the linear dose-response function:

Number of obs = 105
Prob > chi2 = 0.2600
Adj R-squared = 0.0346

	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
ΔT	.0097267	.0086355	1.13	0.260	-.0071984	.0266519
α	-.0635309	.343532	-0.18	0.853	-.7368413	.6097794

- Estimated parameters of the quadratic dose-response function:

Number of obs = 105
Prob > chi2 = 0.1963
Adj R-squared = 0.0487

	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
ΔT	.0341292	.0234206	1.46	0.145	-.0117744	.0800328
$(\Delta T)^2$	-.0003206	.000429	-0.75	0.455	-.0011615	.0005203
α	-.3249222	.3914696	-0.83	0.407	-1.092189	.4423441

- Interpretation of the results:

  - The estimated coefficients of both the linear and the quadratic functional forms have large standard errors and are not statistically significant at the 0.05 level

  - The predicted values ($\widehat{DDY}$), estimated at the national and regional intensities of the OP interventions (ΔT), is close to zero

  - The large standard errors and confidence intervals of the results do not stem from a data limitation in terms of insufficient cross-regional variation in the treatment intensities: CS-SEQDD indicates that the causal contribution of the OP projects to the territorial changes of Y is zero (any recorded Δ is due to spontaneous change)

Can CS-SEQDD take into consideration heterogeneous effects based on contexts/features of territories?

- Standard CS-SEQDD **estimates homogeneous effects** across different features of the territories. Can CS-SEQDD be used to test if different effects of OP projects are produced by different contexts/features of territories?
- **In principle yes:** data can be portioned into different groups of territories sharing the same relevant context/features (as indicated by ToC or TBE)
- CS-SEQDD can be then **applied separately to each group of territories** to assess if the estimated effects are different based on different contexts/features
- **When does this make sense?** When it is predicted by ToC or suggested by TBE
- **Is it feasible?** Only if:

  -ample number of territories for each group

  -large variation of intensities of OP projects within each group

Other Options I:
(GPSM and Longitudinal Econometrics Models)

- In (rare) cases of:

  -generous data availability also on factors and characteristics of the territories (control variables X)

  -large number of territories

  Two other (econometric) options are available

  I) Generalized propensity score matching (GPSM)

  II) Standard longitudinal econometric models

- Intuitively: both options make use of data on X that could cause spontaneous change

  I) Adapts PSM to data in which all units are treated with different intensities

  II) Separate spontaneous change from the effect of the OP projects by estimating the “causal contribution” of X (under strong functional form assumptions)

Other Options II:
Macroeconomic Simulation Models (MSM)

- Macroeconomic simulation models (such as Rhomolo, Quest, REMI…) **Vs.** CS-SEQDD:

  -CS-SEQDD can work only under certain data-availability conditions that enable causal effect estimation, while MSM can be more easily implemented in a larger set of conditions (suffering less from data limitations)

  -MSM do not estimate causal effects from actual ex-post data and the results can be sensible to arbitrary changes in the parameters of the model

  -MSM can be applied only to few indicators Y, and to pre-set fixed territorial units, while CS-SEQDD can be applied to any type of Y and can be adapted to different territorial levels (or hybrid solutions: sector/territory etc..)