Cross-Regional Sequential Difference in Difference (CR-SEQDD): An Empirical Approach for Evaluating EU Thematic-Objective Interventions with Regional Data Aggregated at the National Level

Technical Note

Daniele Bondonio[(*)]

DG Regional and Urban Policy – "Evaluation Network Meeting"
Brussels, 10-11 December 2020

[(*)] *Professor of Statistics for Economic, University of Piemonte Orientale (Ph.D. Carnegie Mellon University). Via Cavour 84, 15100 Alessandria, Italy. E-mail: daniele.bondonio@uniupo.it. https://upobook.uniupo.it/daniele.bondonio*

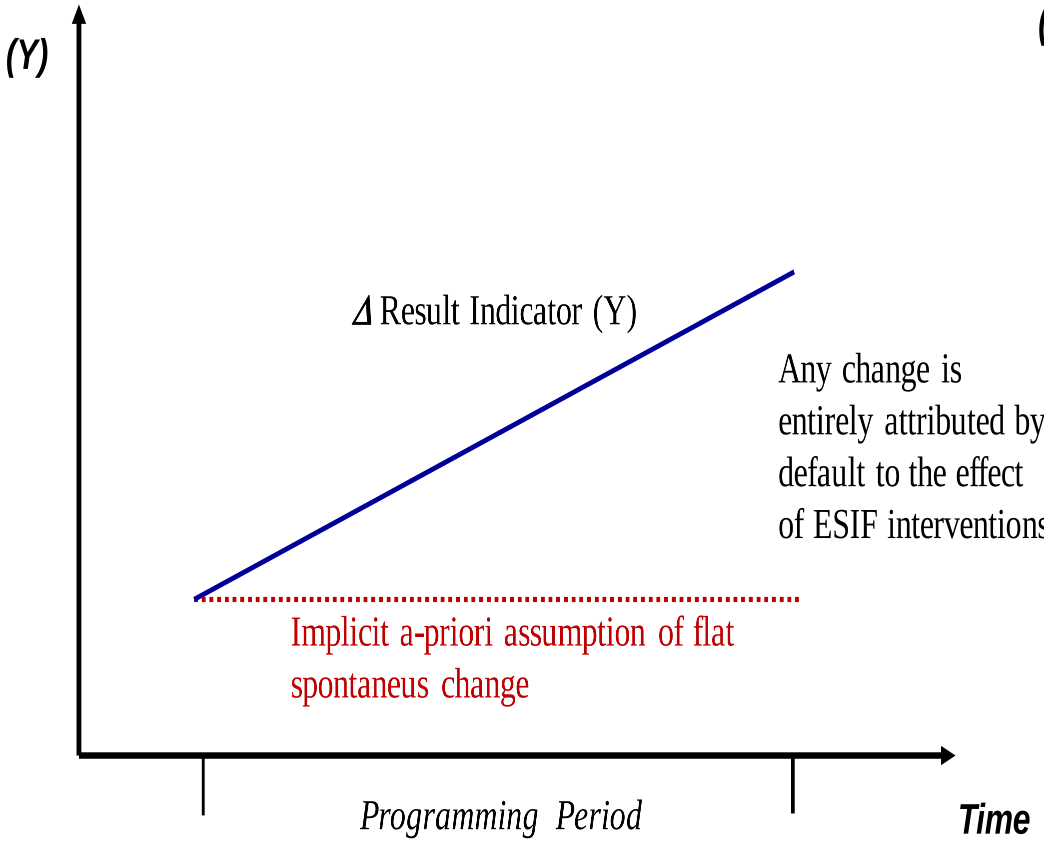
Impact Evaluations of TO Interventions with Regional Data Aggregated at the National Level

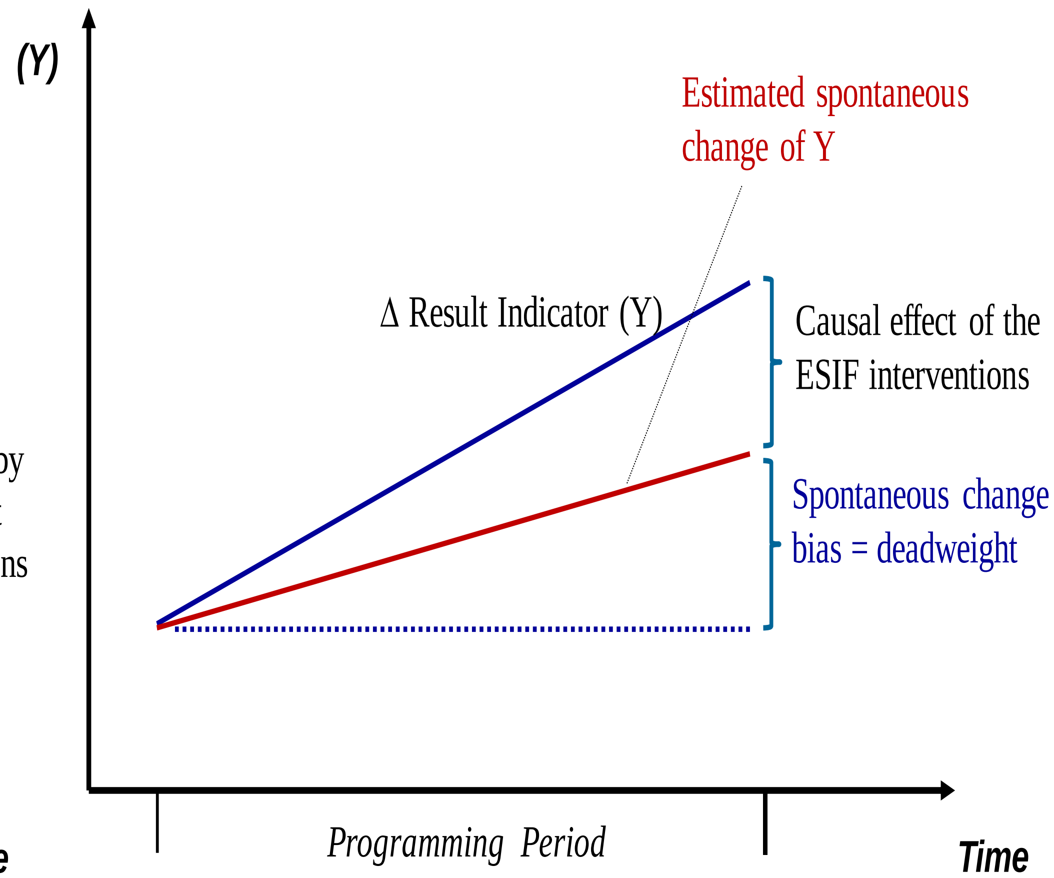
- CR-SEQDD can be applied to impact evaluations aimed at estimating the causal effects produced at the national level on result indicators pertaining to thematic objectives (TOs)

- TO evaluations imply a very high level of aggregation (all OP interventions are pooled together at the national level) and they are tools to inform policy makers, and the public on the overall contribution of the ESIF in achieving desirable results. Findings are useful at a very high decision level to support choices on where to allocate resources among largely-defined domains of program interventions

- Main challenges for TO evaluations: separate the part of the before-after-intervention change of a result indicator Y caused by the ESIF support from the part caused by other factors (unrelated to ESIF support)

Naïf Result-Indicator (RI) Analysis VS Causal-Effect Analysis of RI





The Challenge of Estimating Causal (Effects/Impacts)

- Estimating the spontaneous change would entail to acquire data on units of observations with similar characteristics and with no ESIF support. In the case of data aggregated at the regional or national level this is a nearly impossible tasks: EU member states and regions tend to have unique features and to receive some support from ESIF, so that no comparable "non-treated units" can be found for the analysis

- For this reason, producing reliable "causal effect/impact" estimates on nationally- (or regionally) aggregated Y has inherently a low degree of internal validity

- This is not due to shortcomings of available methods. Because of the challenging impact identification conditions it is scientifically proven that the empirical evidence will have limitations:

  -The more the evaluation is focused on broad TOs at the national level, the less rigorous is the internal validity of causal effects estimations

  -The more the analysis focuses on more specific interventions at a micro-level the higher the internal validity of the causal effect estimation tends to be

CR-SEQDD: Basic Principles and Intuitive Concepts

- The intuitive idea behind CR-SEQDD is to exploit the cross-regional variation in the intensities of the OP interventions in order to estimate (controlling for regional differences) a set of parameters that are subsequently used at the national level to establish what part of a change in Y is likely to be caused by the ESIF support and what part is instead caused by spontaneous change
- Regional differences are controlled for by means of difference-in-difference (DD) models
- Pairwise DD regional comparisons are performed sequentially (in a cross-sectional acception), following an ascending order of the intensity of the ESIF support
- If ESIF support was indeed the major factor in affecting the nationally-recorded before-after-intervention change of Y, these sequential DD comparisons should show that a more positive change of Y is recorded in the regions with higher intensities of ESIF support. If this is not the case, instead, it would be more likely that the nationally-recorded change of Y was the result of spontaneous change

CR-SEQDD Limitations

- The internal validity of CR-SEQDD estimates holds only under very strict causal identification assumptions and the approach combines together different standard econometric tools that have been in existence for decades, with known limitations in the range of applicability

- CR-SEQDD should not be regarded as a breakthrough methodological tool that produces findings with the same strong internal validity as a standard quasi-experimental approach implemented under more favourable scenarios in terms of causal identification conditions

- Unlike experts' opinions and meta-analyses, however, CR-SEQDD offers the advantage of being a fully replicable empirical tool, enabling a consistent comparison of the findings across different times and areas of interventions: the same identical strict casual identification assumptions would apply to the evaluations of different thematic areas, periods and or EU countries, enabling a suitable comparison of the results

CR-SEQDD Data Requirements and Assumptions

CR-SEQDD can be implemented under the following assumptions and data availability scenarios:

- The intensities of the OP interventions can be allocated at the regional level within the programming period of interest, and regional-level data on Y have to cover at least the beginning and the end of the programming period

- OP interventions and the result indicator/s (Y) are measured as intensities with respect to a same baseline size-indicator

  E.g. OP intensity = (TO 1 support)/ (residents); Y=(patent applications) / (residents)

  This size-indicator controls for scale-effect differences among regions that can lead to obvious different potentials for the absolute changes of Y along the estimation period of interest

- Intensities of OP interventions have a sufficiently large degree of variation across regions (necessary condition to achieve standard errors and confidence intervals of limited size)

- The result indicators (Y) have to be affected solely by the OP interventions pertaining to the TO considered in the analysis and not by the OP interventions pertaining to other TOs

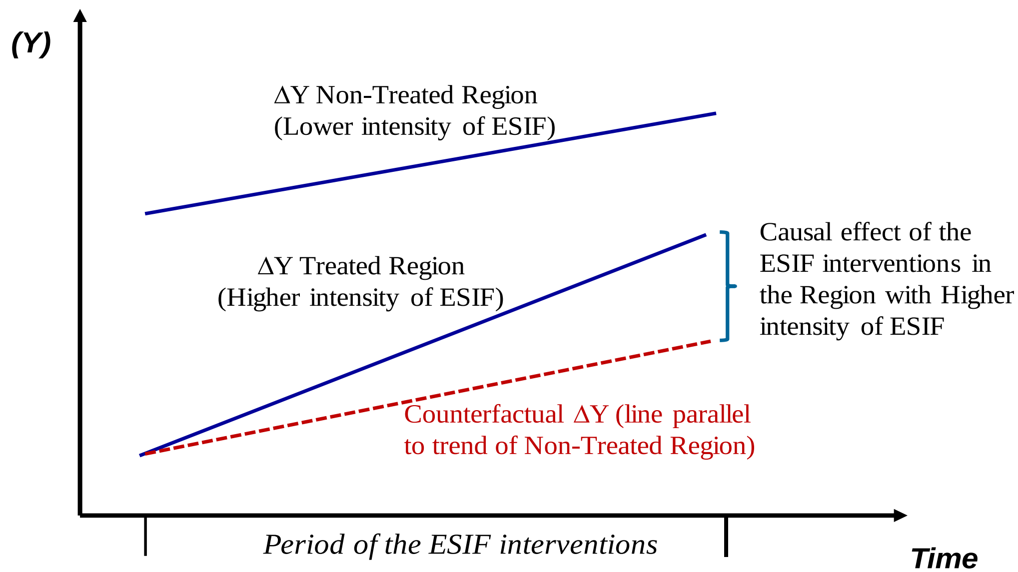
- The spatial spillovers produced by the OP interventions have to be contained within the same region in which they are implemented, not spanning across different regions

- The marginal return on Y of each additional unit of OP intensity is constant both cross-regionally and across the different values of the treatment intensity within the estimation period considered in the analysis.

Controlling for Regional Differences: DD Comparisons

- In the pairwise regional difference in difference (DD) comparisons, before-after-intervention change of Y recorded in the low-intensity region are assumed to be the counterfactual change that would be recorded in the higher-intensity regions in the presence of a lower intensity of the treatment

- This assumption requires that the relevant different baseline characteristics of the regions have to be 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":



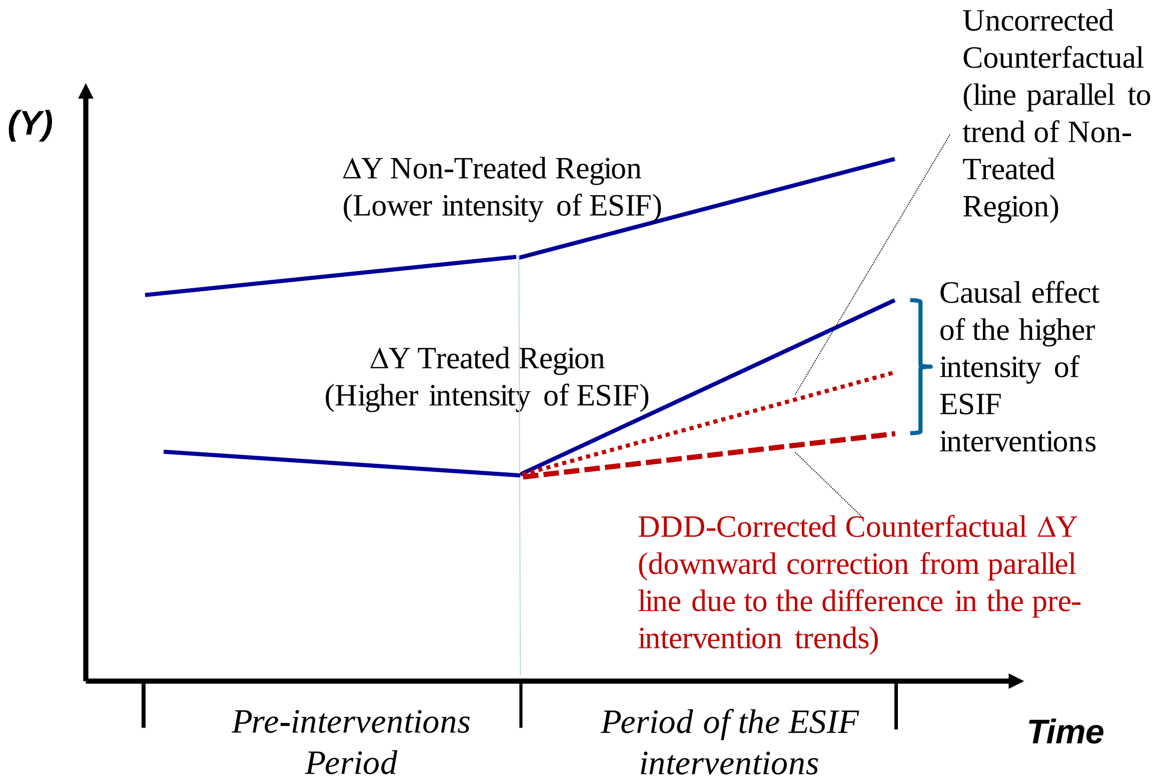
Fixed Effect /Parallel Trend Assumption

- Fixed-Effect Assumption:
  E.g. Region A (receiving a low intensity of OP interventions) is structurally different than region B (higher intensity of the OP interventions). For example, the R&D capacity of Region A is higher than Region B because Region A has more universities, larger number of existing R&D labs and facilities, stronger concentration of residents with higher education. These structural differences between Region A and Region B, entails that Region A tends to have, in any given year, an higher value of a result indicator Y (e.g. n. of patent applications) than Region B

- With a DD comparison, these different structural characteristics of the regions are controlled for by means of transforming the values of result indicator Y into changes between the beginning and the end of the OP interventions. The rationale behind this is the following: if the differences between the two regions are structural characteristics (fixed effects), these different features by definition do have an influence on the levels of result indicator Y (e.g. yearly number of pro-capita patent applications), but they cannot have an influence on the change of Y between different years (these “fixed effects” are always in existence and therefore they cannot induce a change in Y between different periods)

Difference-in-Difference-in-Difference (DDD) Comparisons

- In the case where the estimation period can be extended to include one additional pre-intervention time, in which the regional units of observation are all unexposed to the treatment (or exposed to a treatment of the same intensity), the CR-SEQDD model can be estimated with a difference-in-difference-in-difference scheme



Estimation Procedure

A) The before-after-intervention trend of Y and the intensity of the OP interventions are recorded for each region

B) Regions are sorted in ascending order of the intensity of the OP interventions

C) Sequentially, along the order of the treatment intensity, a series of pairwise cross-regional difference-in-difference (DD) comparisons [or DDD comparisons] are implemented. The results from the DD [or DDD] pairwise comparisons are in terms of casual impact parameters (*DDY*) or (*DDDY*) that (with all the limitations posed by the very strict impact identification assumptions previously mentioned) estimate the degree by which an higher intensity of the OP interventions generate a positive change of Y

D) The results from the pairwise cross-regional DD comparison are displayed in a two-way scatter plot that contains on the vertical axis the causal impact parameters *DDY* and on the horizontal axis the corresponding difference of treatment intensity (ΔT) between the pair of regions

E) Based on the plot chart of step D), a linear or quadratic dose-response function is estimated:

$$DDY = \alpha + \beta \Delta T + \varepsilon \qquad (1) \qquad \text{or}$$

$$DDY = \alpha + \beta \Delta T + \chi (\Delta T)^2 + \varepsilon \qquad (2)$$

This linear (1) or quadric function (2) is then used to predict what would be the expected contribution on *DDY* of the ΔT registered at the national level

F) The predicted *DDY* value ($\widehat{DDY}$), estimated in step E) for the nationally-recorded intensity of the OP interventions, is then compared with the raw national change (ΔY):

- When $\widehat{DDY}$ reaches similar values of ΔY, the CR-SEQDD findings are indicative of a causal impact of the OP interventions being responsible for most of the nationally-recorded ΔY
- When $\widehat{DDY}$ is largely lower than ΔY, the CR-SEQDD findings are indicative of a strong component of spontaneous change being responsible for most of the nationally-recorded ΔY

Application Examples

- Nationally-aggregated OP interventions pertaining to TO1.

- Data:

  - Sample of regions *N=15*

  - $\Delta Y_i = (Y_{ipost} - Y_{ipre})$ = Pre-post-intervention change in the yearly number of patent applications per million of residents recorded in region (*i*) [*t=pre* (pre-intervention year) and *t=post* (post-intervention year]

  - T_i= Per-capita intensity of the ESIF monetary resources spent in the (pre-post) period for all the OP interventions pertaining to TO1

Example I): Ideal data-availability scenario, strong 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	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
Comparison Region (Higher T)	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	-									
	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)

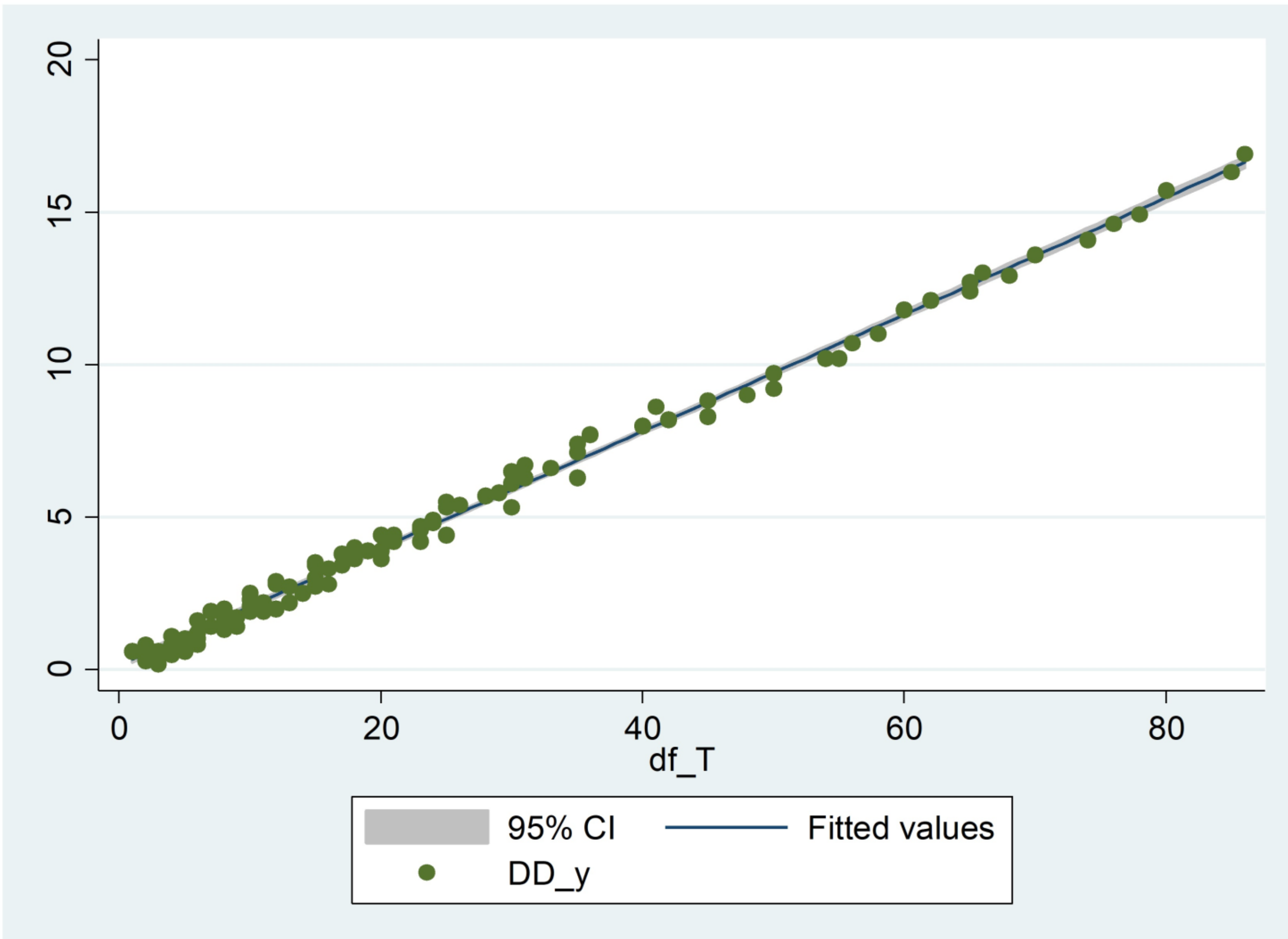
		Baseline Region														
		A	B	C	D	E	F	G	H	I	L	M	N	O	P	Q
	A	-														
	B	20	-													
	C	45	25	-												
	D	50	30	5	-											
	E	55	35	10	5	-										
	F	62	42	17	12	7	-									
Comparison Region	G	65	45	20	15	10	3	-								
	H	68	48	23	18	13	6	3	-							
	I	70	50	25	20	15	8	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
Wald chi2(1)        =    1398.59
Prob > chi2         =     0.0000
R-squared           =     0.9945
Adj R-squared       =     0.9945
Root MSE            =     0.3171
-----------------------------------------------------------------------------
-
             |   Observed   Bootstrap                         Normal-based
             |      Coef.   Std. Err.       z     P>|z|     [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 =

$$0.1525 + 0.1915 * 63.4 = 12.3$$

  (additional number of yearly patent applications per million of residents caused nationally by an intensity of € 63.4 Million worth of OP interventions in TO 1)

- 95% confidence interval of $\widehat{DDY}$ = [+11.46, 13.13]

- Interpretation of the CR-SEQDD findings:

CR-SEQDD estimate $\widehat{DDY}$ for the nationally-recorded intensity of OP interventions =12.3, when compared to the nationally–recorded raw change of the result indicator Y (ΔY =+12.7), indicates that the OP interventions were likely to be responsible for almost all of the of before-after-intervention change in the result indicator Y, with a minimal role played by spontaneous change in affecting such change.

Example II): Ideal data availability scenario, 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	-

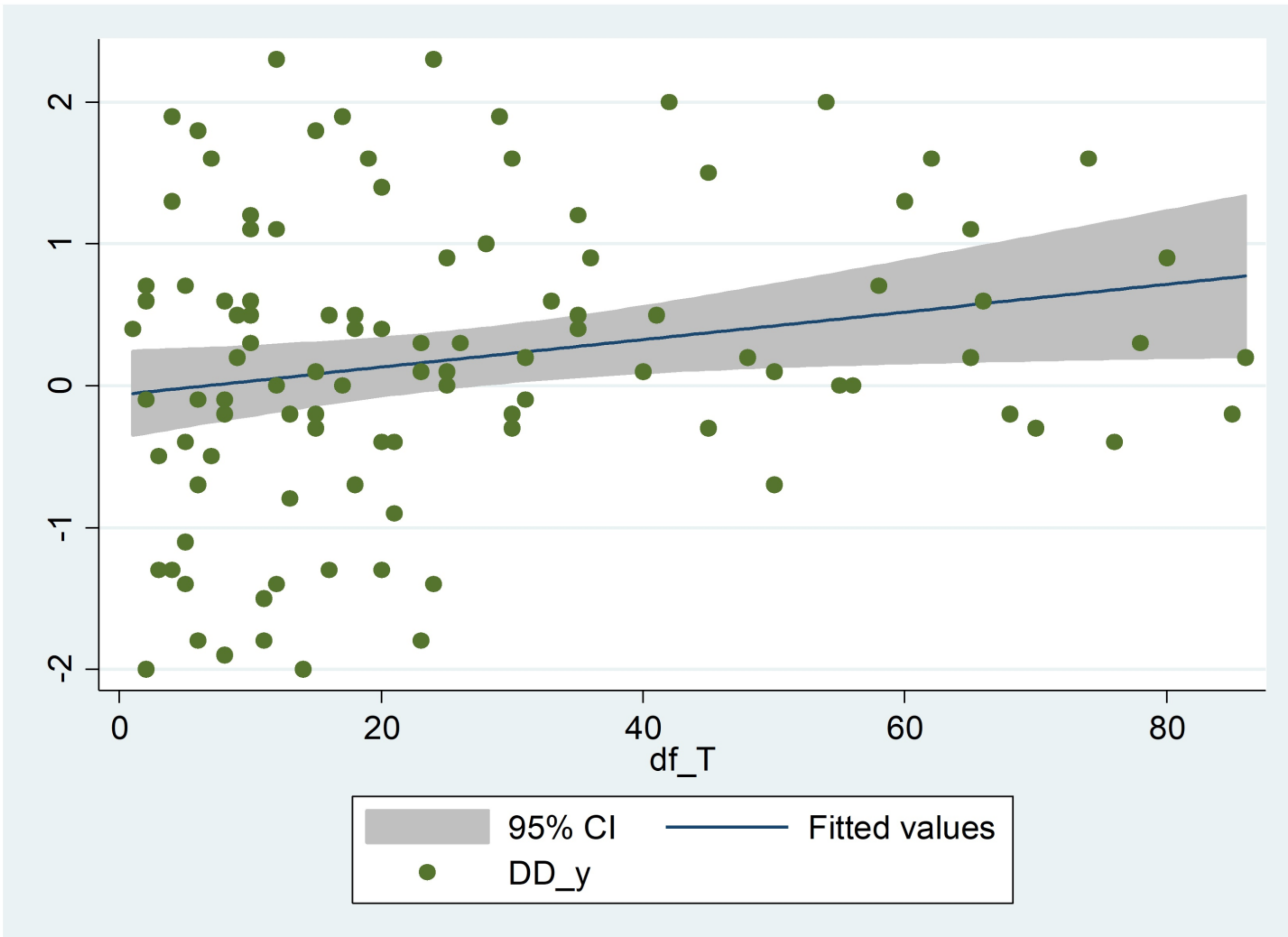
1= [No. Pat. Appl. / Mil. Residents] in terms of pairwise Difference-in-difference variation of Y (DDY) between Comparison and Baseline Regions

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

		Baseline Region														
Comparison Region		A	B	C	D	E	F	G	H	I	L	M	N	O	P	Q
	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)

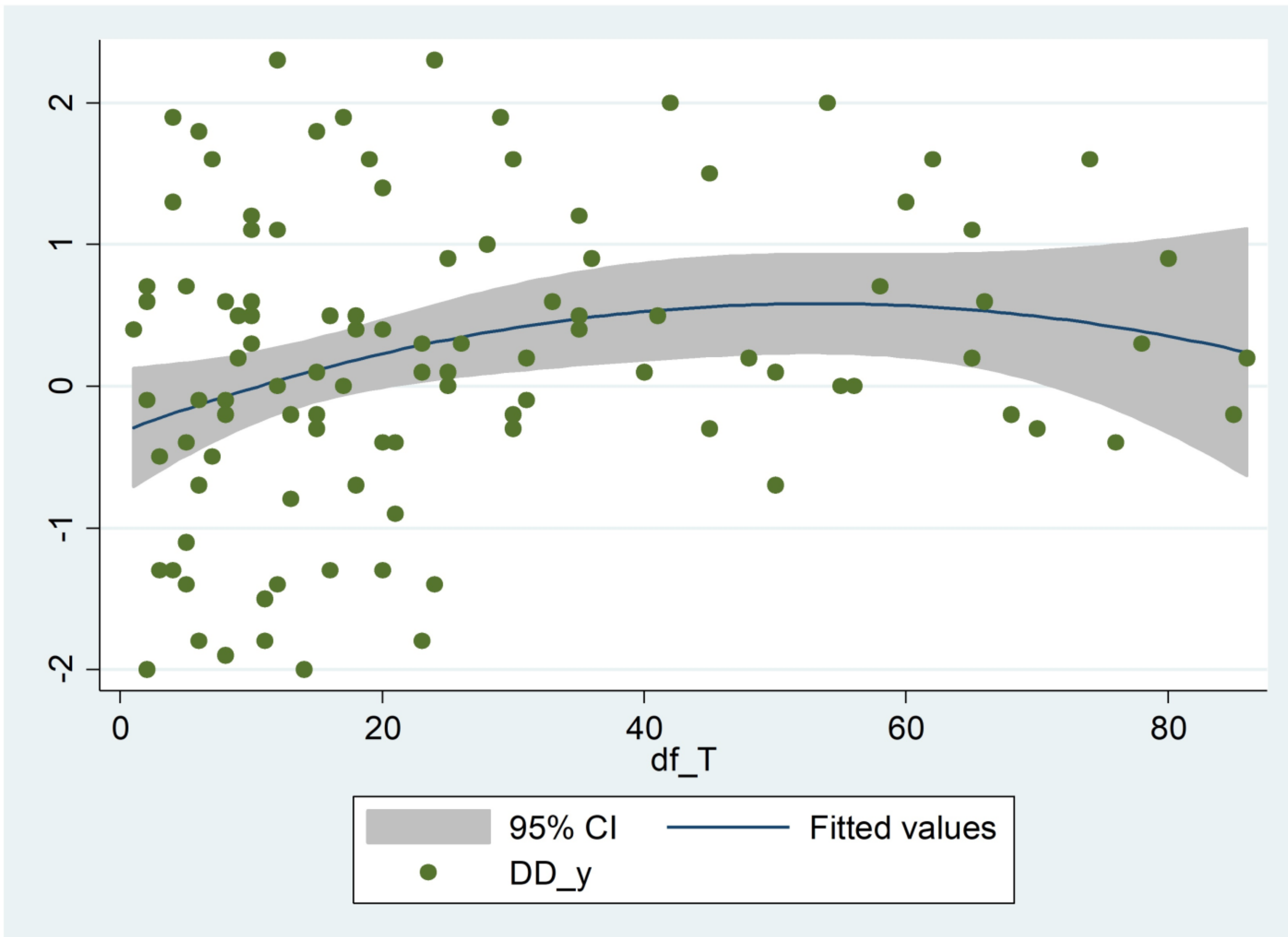


- Estimated parameters of the linear dose-response function:

Number of obs	= 105
Wald chi2(1)	= 1.27
Prob > chi2	= 0.2600
R-squared	= 0.0439
Adj R-squared	= 0.0346
Root MSE	= 1.0153

	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

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 quadratic dose-response function:

Number of obs	=	105
Wald chi2(2)	=	3.26
Prob > chi2	=	0.1963
R-squared	=	0.0670
Adj R-squared	=	0.0487
Root MSE	=	1.0079

	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 CR-SEQDD findings:

  - 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
  - Based on these parameters, the predicted value ($\widehat{DDY}$), estimated at the nationally-recorded intensity of the OP interventions (ΔT), is close to zero, with a point estimation that, for both functional forms, is equal to +0.55 (additional number of yearly patent applications per million of residents caused nationally by an intensity of €63.4 Million worth of OP interventions in TO 1)
  - Because of the very large standard errors of the coefficient estimates, the related 95% confidence interval of this ($\widehat{DDY}$) predicted value is extremely ample for both functional forms
  - 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. For this reason, CR-SEQDD estimates indicates that the nationally-recorded pre-post intervention ΔY=+63.4 is most likely due to spontaneous change, and that the causal contribution of the OP interventions is instead minimal