Non-Experimental Methods: Propensity Score Matching and Difference in Difference

Alessandro Maffioli
Office of Strategic Planning and Development Effectiveness
Mind the Gap: From Evidence to Policy Impact, Cuernavaca, June 15 2011

Structure of the presentation

I. Propensity Score Matching

- Basic Assumptions
- Implementation
  - Estimating the propensity score
  - Common support
  - Matching algorithm
  - Quality of the matching
  - Computing results

II. Difference in Difference

- Intuition
- Basic assumption
- DD y PSM

III. PSM y DD in practice: some example

Propensity Score Matching

Non-experimental evaluation

- As previously mentioned not always randomized data are available
- Individuals not assigned to treatment by a random process →we don't have an experimental control group
- It is then essential to understand and model the processes by which assignments to treatments are made
  - Self-selection (e.g., individual decision to apply)
  - Administrator selection (e.g., individuals assigned to treatment based on specific criteria)
  - Combination of self/administrator selection

The evaluation problem

The impact evaluation problem:

A = households that receive the program

B = households that do not receive the program

Y = outcome = % adoption cement floor

Not observable

ATT = ($\overline{Y}_A$| A participating) – ($\boxed{\overline{Y}_A\text{| A no Participating)}}$

($\overline{Y}_A$| A participating) – ($\overline{Y}_B$| B no Participating) = ATT + $\boxed{\text{(Difference A,B)}}$

Selection bias

I can use B as counterfactual only if (Difference A,B) = 0

Development Effectiveness

PSM and DD basic intuition

Both the PSM and DID try to correct the selection bias:

- **PSM** removes bias associated with **observables characteristics** that affect treatment assignment
- **DID** removes bias associated with **time-invariant characteristics** both **observable and unobservable** that affect treatment assignment
- Combining PSM and DID (usually) improves the estimations
- If the treatment assignment is affected by time-variant unobservable…….try IV!!!

Example

HH	TREATMENT	EDUCATION	INCOME
1	B	2	60
2	B	3	80
3	B	5	90
4	B	12	200
5	A	5	100
6	A	3	80
7	A	4	90
8	A	2	70

MATCH

INCOME A	INCOME COUNT	DIFF

ATT	

Example

HH	TREATMENT	EDUCATION	INCOME	MATCH	INCOME A	INCOME COUNT	DIFF
1	B	2	60				
2	B	3	80				
3	B	5	90				
4	B	12	200				
5	A	5	100	[3]			
6	A	3	80				
7	A	4	90				
8	A	2	70				

ATT	

Development Effectiveness

Example

HH	TREATMENT	EDUCATION	INCOME
1	B	2	60
2	B	3	80
3	B	5	90
4	B	12	200
5	A	5	100
6	A	3	80
7	A	4	90
8	A	2	70

MATCH
[3]
[2]

INCOME A	INCOME COUNT	DIFF

ATT	

IDB. Development Effectiveness

Example

HH	TREATMENT	EDUCATION	INCOME
1	B	2	60
2	B	3	80
3	B	5	90
4	B	12	200
5	A	5	100
6	A	3	80
7	A	4	90
8	A	2	70

MATCH
[3]
[2]
[2,3]

INCOME A	INCOME COUNT	DIFF

ATT	

Example

HH	TREATMENT	EDUCATION	INCOME
1	B	2	60
2	B	3	80
3	B	5	90
4	B	12	200
5	A	5	100
6	A	3	80
7	A	4	90
8	A	2	70

MATCH
[3]
[2]
[2,3]
[1]

INCOME A	INCOME COUNT	DIFF

ATT	

Development Effectiveness

Example

HH	TREATMENT	EDUCATION	INCOME	MATCH	INCOME A	INCOME COUNT	DIFF
1	B	2	60				
2	B	3	80				
3	B	5	90				
4	B	12	200				
5	A	5	100	[3]	100	90	10
6	A	3	80	[2]	80	80	0
7	A	4	90	[2,3]	90	(80+90)/2=85	5
8	A	2	70	[1]	70	60	10

ATT	**6.25**

Basic assumptions (1)

- **Conditional Independence Assumption (CIA) or unconfoundedness condition**: given a set of observable covariates *X,* which are not affected by treatment, potential outcomes are independent of the treatment assignment:
- **In practice:** the matching estimate the effect of treatment on treated assuming that *conditional on observable characteristics*, participation is independent of outcomes.
  – This removes bias associated with pre-treatment differences between treatment and comparison groups
  – Useful when data on pre-treatment, observed characteristics is rich
- **Limitation:** if treatment status is influenced by unobserved characteristics, the estimated impacts are biased

Development Effectiveness

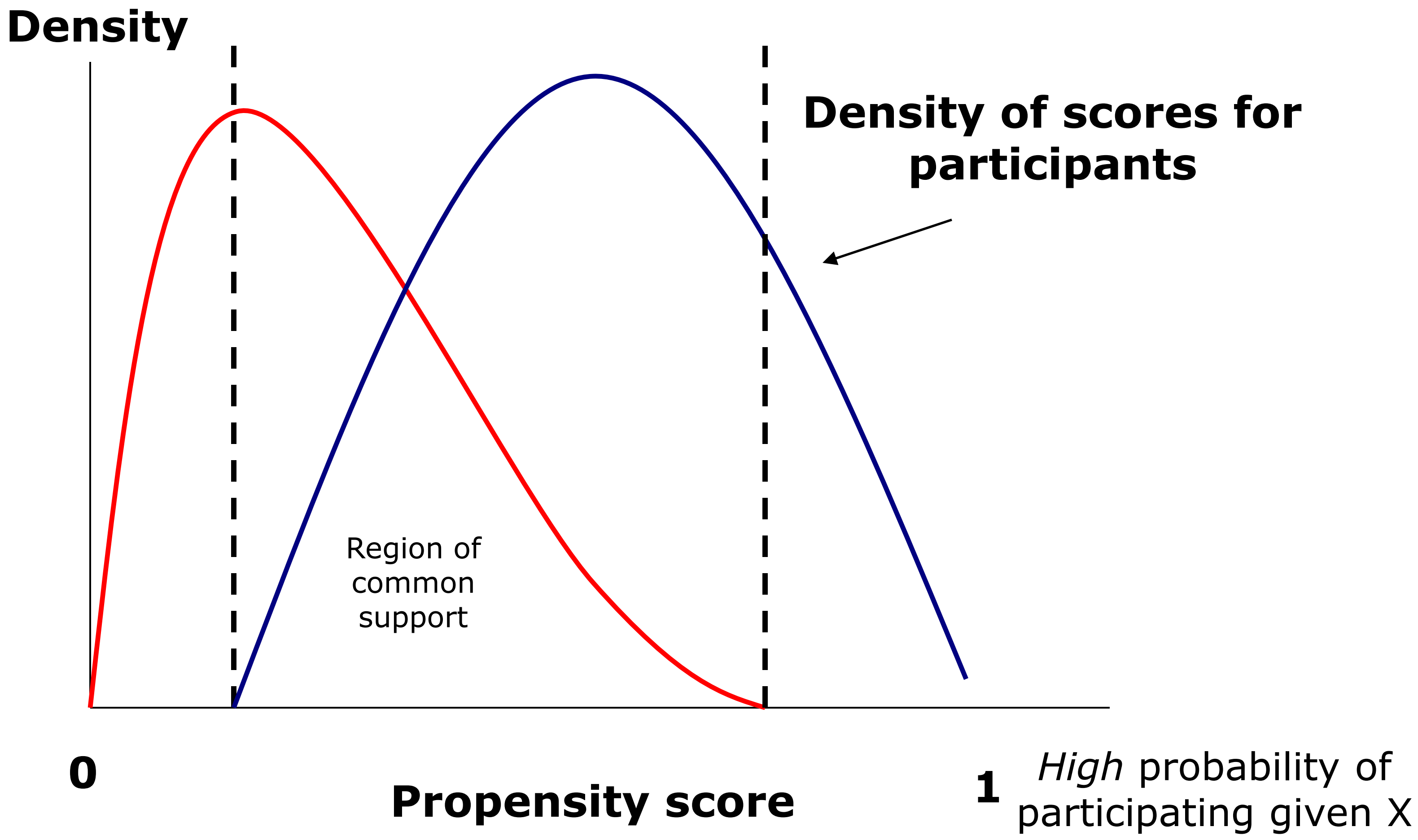
Basic assumptions (2)

- Propensity score matching: estimation of a “participation model” that reduces matching problem to a single dimension (propensity score)
- The propensity scores are then used to match treatment and comparison groups

We need to have some individuals in the control group that have similar characteristics to individuals in the treatment group

Overlap condition in PSM

Common support (overlap condition)



Implementation

Run Discrete Choice Model

- Dependent variable: Y=1, if participate; Y = 0, otherwise
- Choose appropriate conditioning variables
- Obtain propensity score: predicted probability

Match each participant to one or more nonparticipants on propensity score

- Identify the common support
- Choose the matching algorithm
- Check the quality of the matching

Estimate the impact based on new sample

Estimating the Propensity Score

1. Choose the model:
   - Logit or probit models (Multinomial for multi-treatment)
2. Choose the covariates:
   - Basic criteria: choose variables that make credible the CIA ➔ variables that affect simultaneously participation decision and outcome
   - To few or to many, how to choose?
     - Economic theory
     - Quality of the matching
     - Parsimonious specification and progressive adding
     - Goodness-of-fit of the model (careful!)

Notes: (i) the objective is matching, not estimating the coefficient

(ii) use the same sources for treated and control

Development Effectiveness

Choosing a matching algorithm

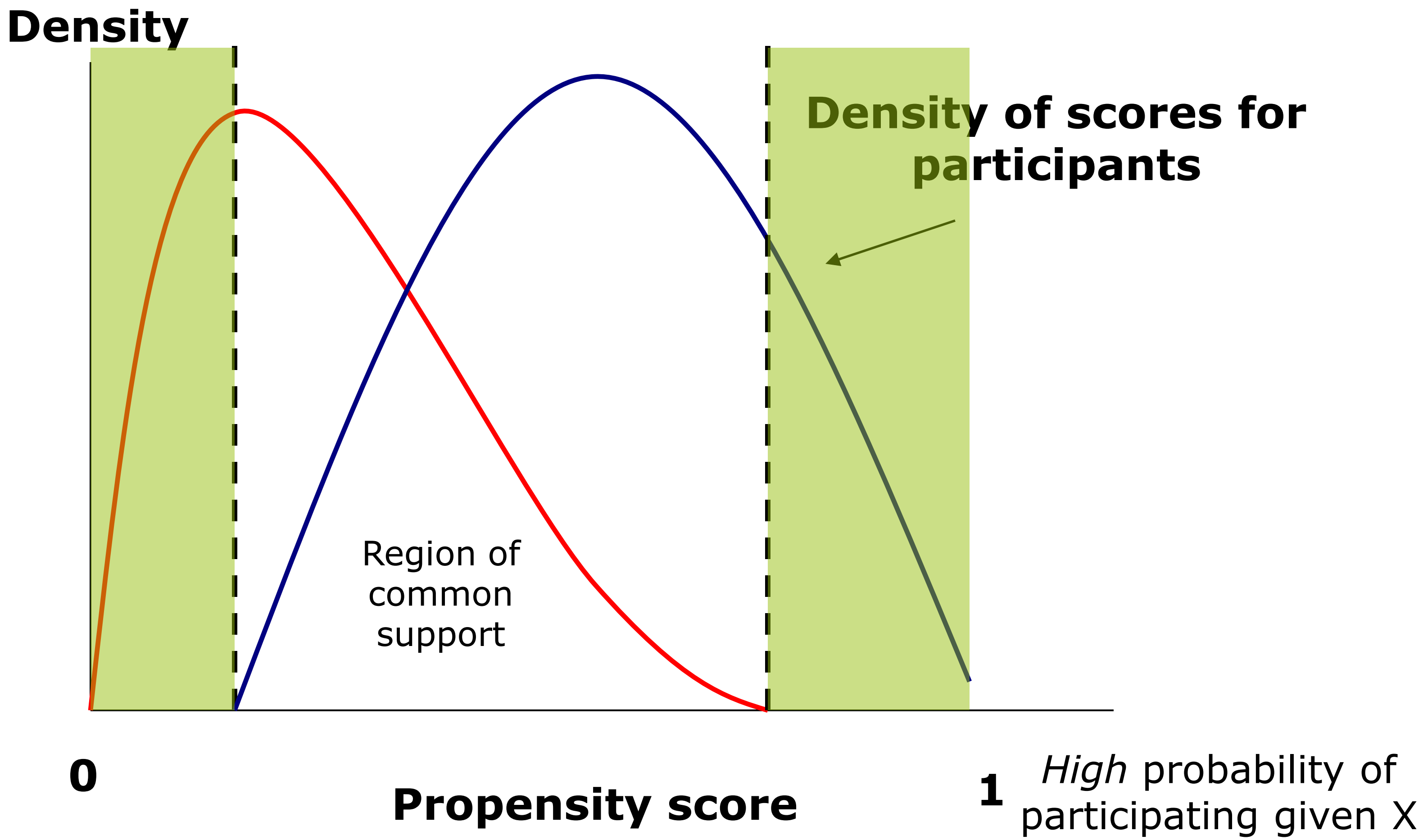
Model	Description	Alternatives	Key factor	BIAS	EFF
Nearest Neighbor	Choose the control(s) with the minimum mahala-nobis distance	• No replacement • Replacement • Oversampling	• Order of matching • Poorer matches • Poorer matches	✓ ✓ ✓ ✓	✕ ✕ ✕ ✕
Caliper and Radius	Choose the control(s) within a certain distance	• Caliper (NNnR) • Radius (NNR)	• Tolerance level • Radius definition	✓ ✓ ✓	✕ ✕ ✕
Stratification	Divide the common support in strata	• No. of strata	• Definition of the No. of strata	✓	✕
Kernel	Use weighted average of all the individual in the common support	• Kernel functions • Bandwitht	• Proper definition of the common support	✕	✓ ✓

Indentifying the common support

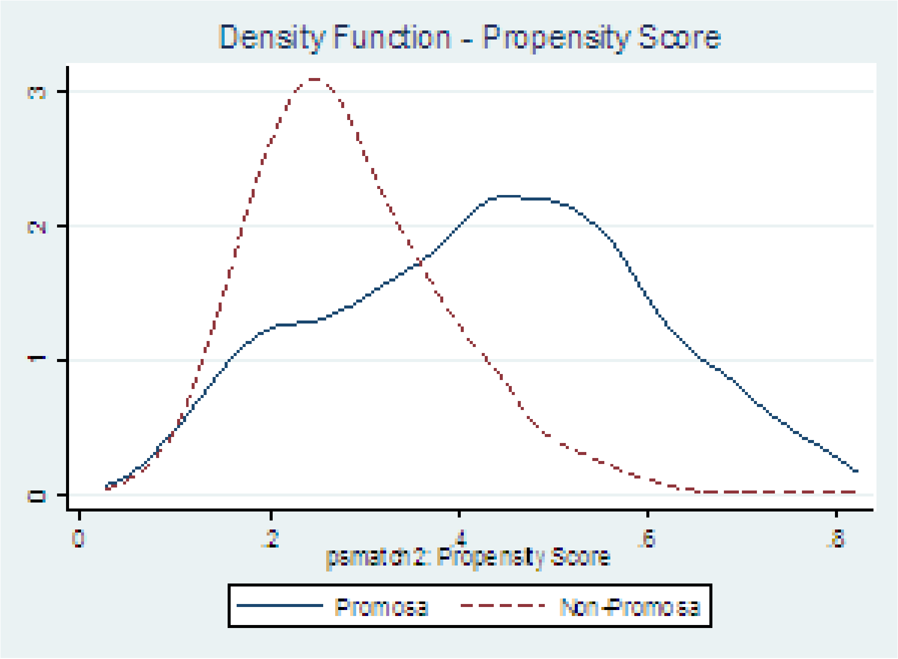
Different strategies can be used to identify the common support:

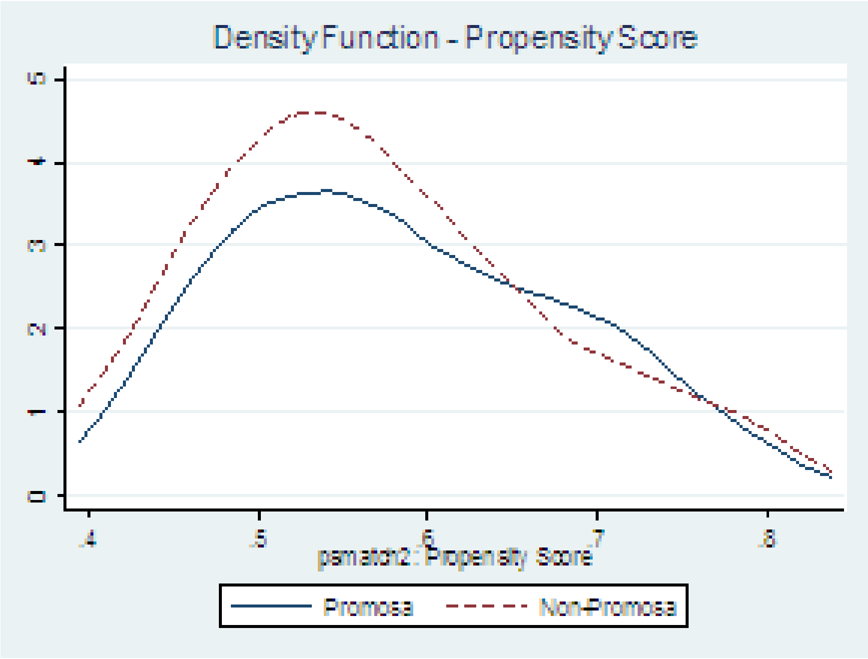
- Visual analysis: plot the density distribution of the propensity score for both groups and overlap the graphs
- Minimum and maximum criterion: delete all observations whose propensity score is smaller that the minimum and larger than the maximum of the opposite group
- Trimming: exclude all the observations in the areas where one of the two propensity score distribution is zero
- Statistical similarity of the p-score distributions: run a dissimilarity test on the two distributions

Minimum – maximum criterion



Testing equality of distributions





Kolmogorov-Smirnov equality of distribution test (ksmirnov)

Assessing the matching quality

Given the CIA assumption we need to check that the matching procedure has balanced the distribution of the relevant variables in the treatment and control groups

- Check for the standardized biased reduction before/after the matching (pstest with stata) $\Rightarrow$ at least 5% reduction

- Test equality of means in the treated and control groups before/after the matching (pstest with stata)

- Check for joint significance of the "participation model": after the matching the pseudo-R^2 should be fairly low

If quality of the matching is not satisfactory, the CIA failed

Summing up

When an evaluator uses a PSM methodology, you want to check:

- ✓ What data on observable characteristics are available
- ✓ Which variables are included in the "participation model" and why
- ✓ The identification of the common support
- ✓ Which matching algorithm is used and why
- ✓ The balancing of the of the relevant covariates (and of the PS distribution)

Main limitations of this method:

- ✓ Only reduces bias due to observables characteristic
- ✓ Internal validity limited the common support
- ✓ Highly intensive in data

Difference in Difference

Difference in Difference estimator

DD removes bias from the “Before-After” (BA) and the “Cross-section” CS estimator by using the double difference [BA CS].

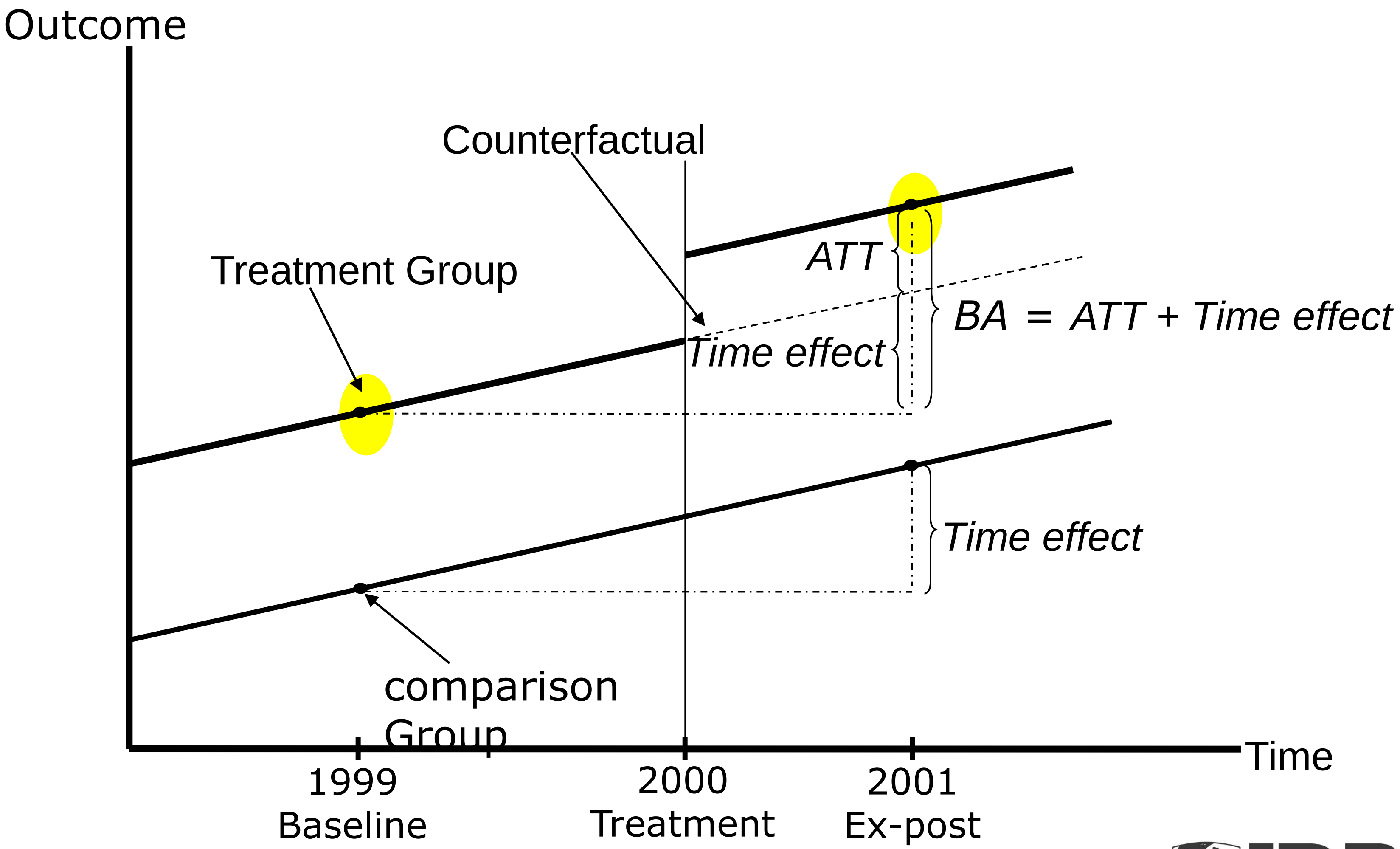
- the BA estimator has 2 components: impact+ time effect
- The CS estimator also has 2 components: impact+ the difference in levels due to group-specific factor

The DD estimator uses the BA of the comparison group to remove the time effects and the pre-treatment CS to remove the difference in levels

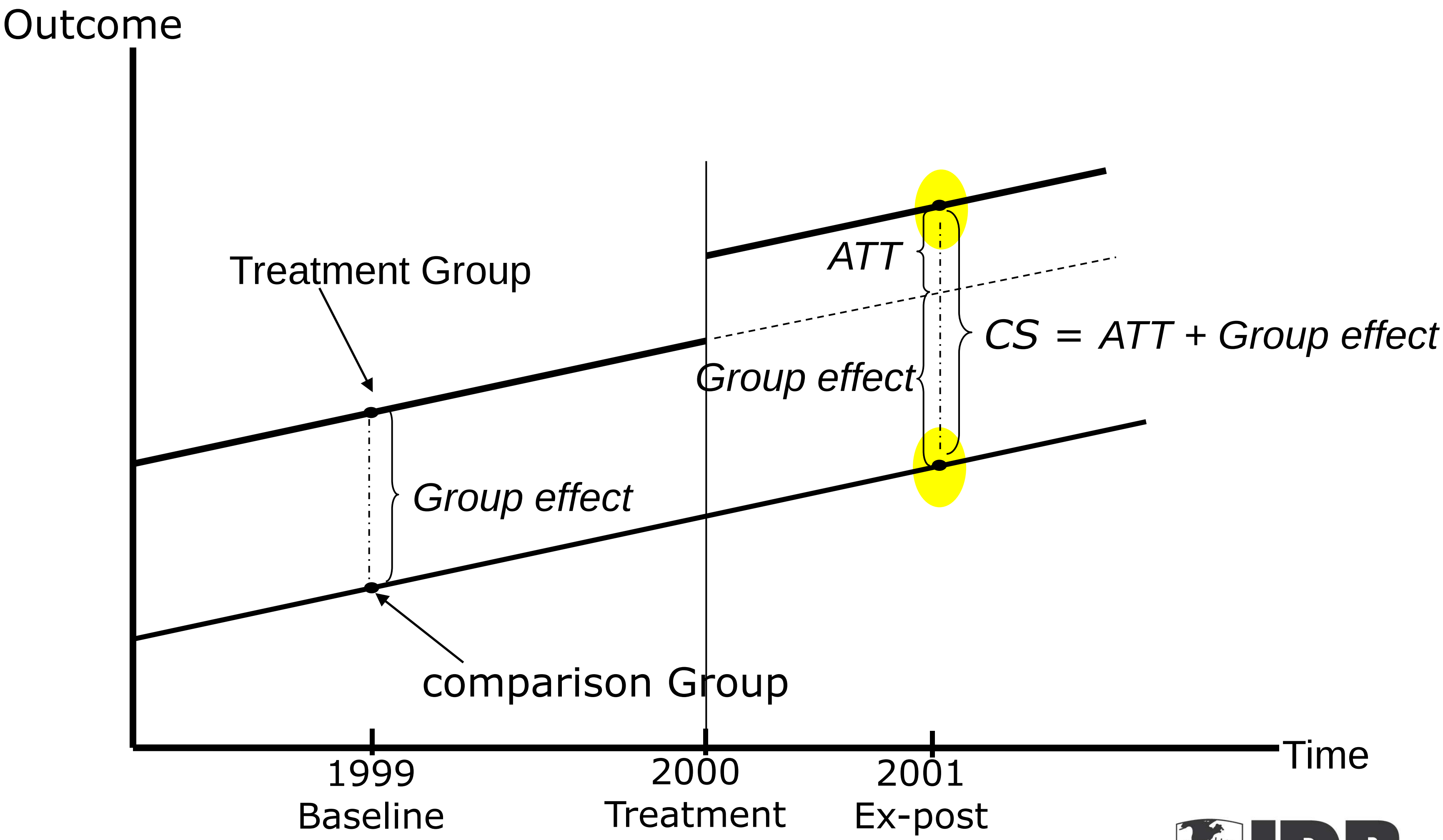
$$DD = [\, Y_{A,2001} - Y_{A,1999} \,] - [\, Y_{B,2001} - Y_{B,1999} \,]$$

$$DD = [\, Y_{A,2001} - Y_{B,2001} \,] - [\, Y_{A,1999} - Y_{B,1999} \,]$$

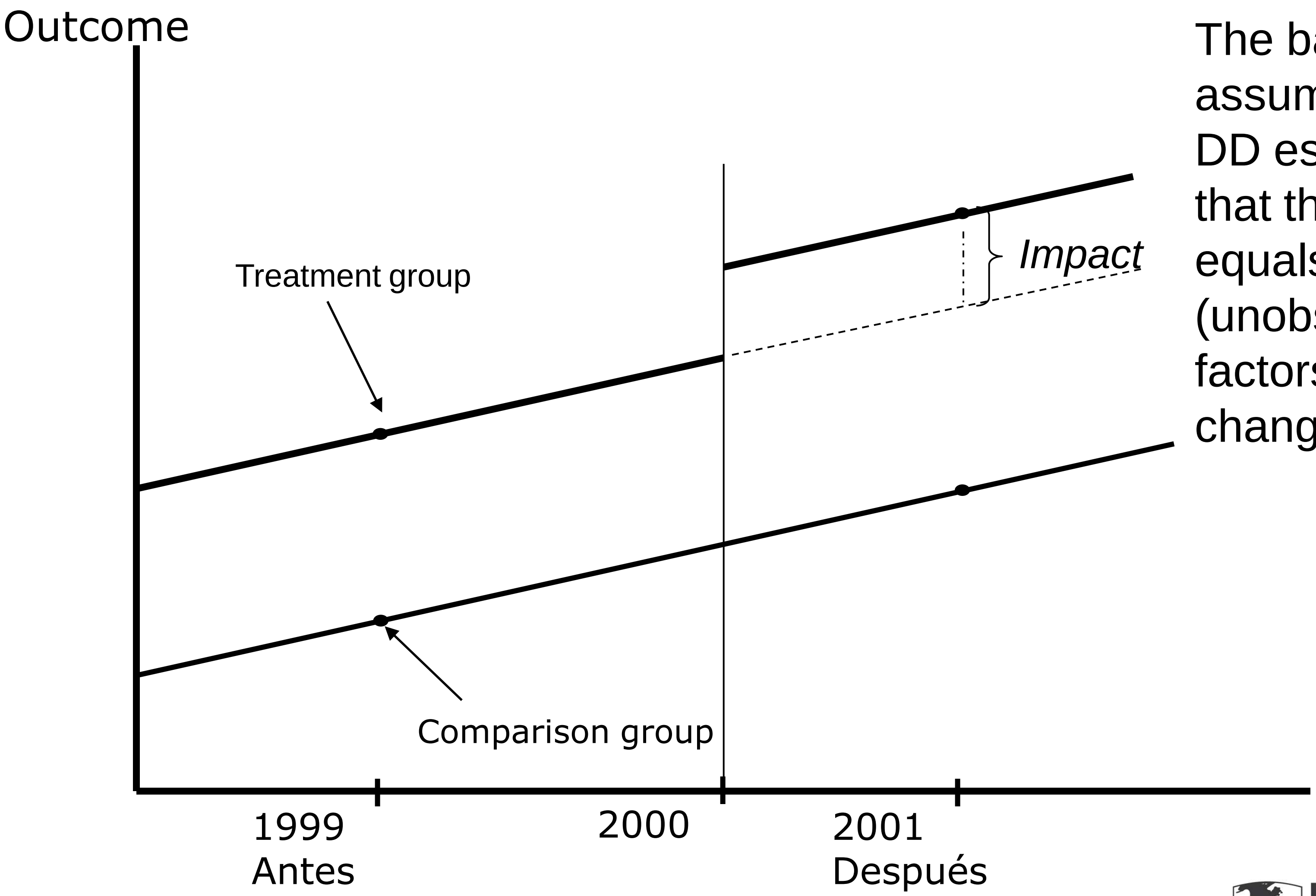
Graphical intuition: BA estimator



Graphical intuition: CS estimator



Graphical intuition: DD estimator



The basic assumption of the DD estimator is that the trends are equals (unobservable factors do not change over time)

DD estimator: basic assumption

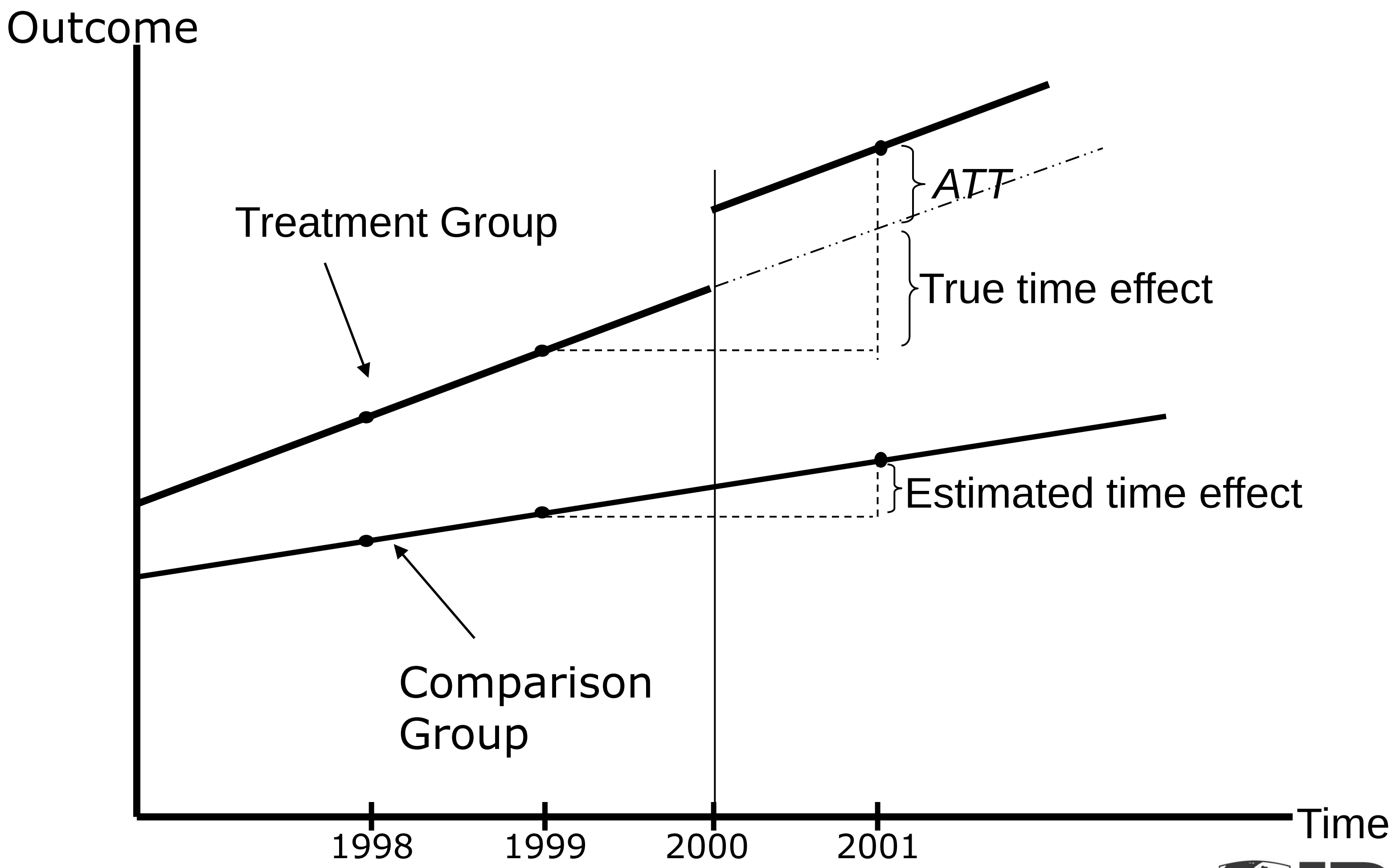
- Under what conditions does the DD estimator effectively estimate the parameter of interest? When the group B will be a good *contrafactual* for the group A?

- Key assumption: common trends
  - the time effect must be equal between treatment and comparison group
  - Unobservable group-specific factors must be time-constant

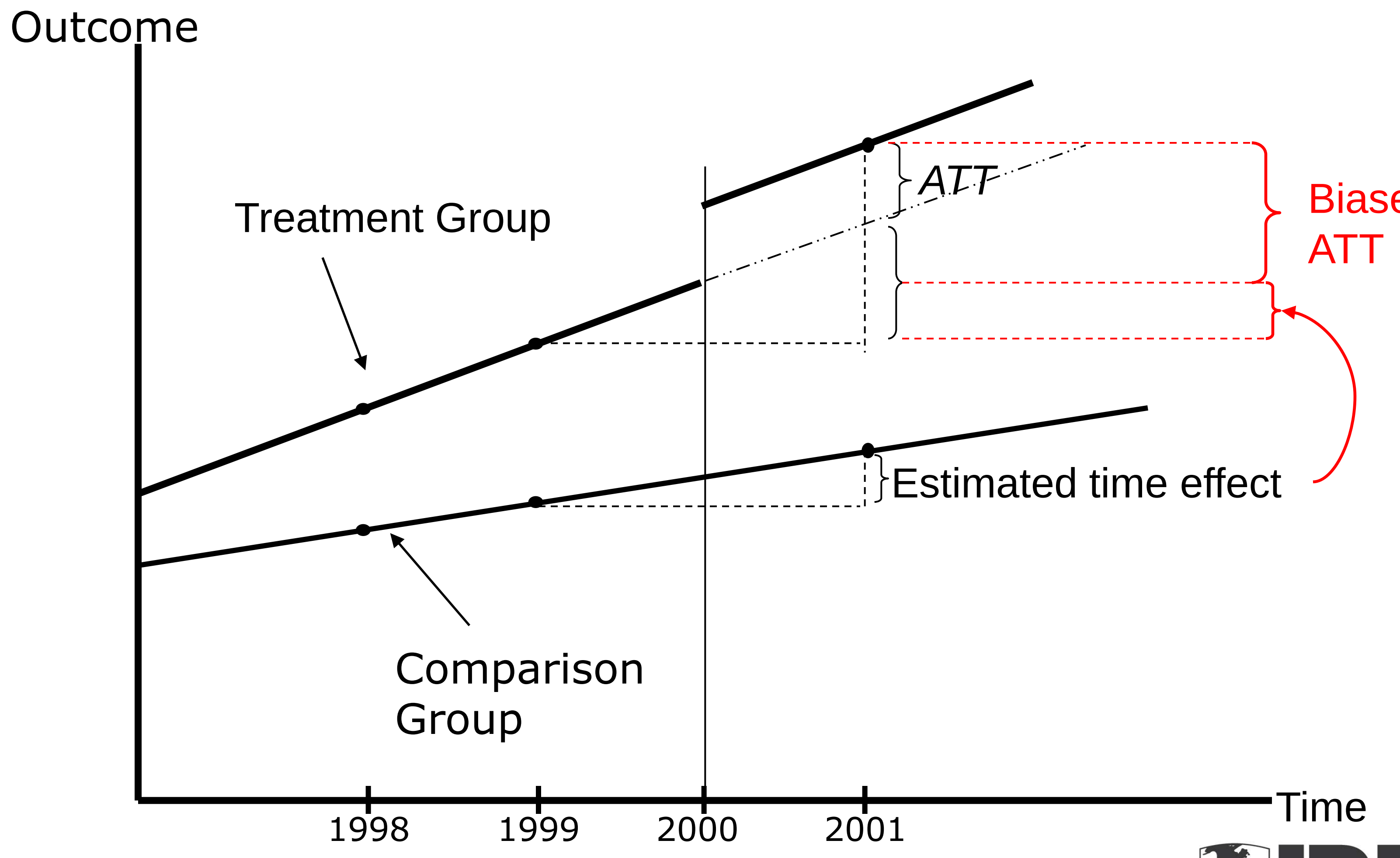
- If the groups have different trends, the DD estimator will then be biased
- This assumption is not testable. However, some evidence of its validity can (and should!) be provided $\Rightarrow$ parallel trend of outcomes before the treatment

Development Effectiveness

Example: failure of assumption



Example: failure of assumption



Combining PSM and DD

Pros of combining the two methods :

- DD improve the PSM because it controls for unobservable heterogeneity, if constant overtime
- PSM improves the DD because it could make the parallel trends assumption more credible

How to implement the PSM-DD

- Traditional PSM with DD: match on the basis of pre-treatment characteristics and compute the impact as the double difference CS BA. When possible, **match pre-treatment trends** of the outcome variable
- Fixed effect on the common support (weighted panel using p-scores)

Propensity Score Matching and Difference in Difference in Practice: Some Examples

Examples: evalutions of PDPs

- PREDEG (Cerdan et al. 2008):
  - Vouchers for farmers to increase technology adoption
  - Combination of PSM-DD

- FONCYT (Chudnovsky et al., 2006 and Ubfal Maffioli, 2011):
  - Grants for scientist to improve scientific productivity and cooperation
  - Difference in Difference (fixed effects)

- Public credit in Brazil, PCB (De Negri et al., 2011)
  - Loans for firms (SMEs) to increase export
  - PSM and DD (matching trends)

- COLCIENCIAS (De Negri et al., 2011)
  - Grants for SMEs(SMEs) to increase innovation and productivity
  - PSM and DD

Pre-treatment characteristics (PREDEG)

	Treated			Control			
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Mean difference
Characteristics of the producers							
Age	71	47.14	11.52	257	50.66	12.99	-3.52**
Gender	71	0.97	0.17	257	0.91	0.28	0.06*
Education	71	11.42	3.74	258	10.46	3.73	0.96**
Foreign	72	0.06	0.23	260	0.06	0.24	-0.01
Individual	71	0.80	0.40	257	0.89	0.32	-0.08*
Company	71	0.15	0.36	257	0.06	0.24	0.09**
Characteristics of the farms							
Size (Plants)							
Micro	72	0.08	0.28	260	0.29	0.45	-0.21***
Small	72	0.38	0.49	260	0.22	0.41	0.16***
Small-Medium	72	0.24	0.43	260	0.12	0.32	0.12***
Medium	72	0.15	0.36	260	0.04	0.20	0.11***
Large	72	0.06	0.23	260	0.03	0.17	0.02
Total land	71	38.44	51.98	258	57.80	204.69	-19.37
Employment							
Total Employment	71	5.69	5.58	258	5.36	15.08	0.33
Temporary employment	71	240.03	410.19	258	373.78	3026.21	-133.75
Skilled labor	71	0.34	0.61	258	0.97	5.89	-0.63
Skilled labor (%)	71	0.06	0.13	257	0.06	0.16	0.00
Residents	71	0.76	4.51	258	0.76	7.61	0.00
Machine and equipment							
Tractors	72	2.53	1.52	260	2.19	4.08	0.34
New tractors	72	0.13	0.33	260	0.12	0.44	0.00
FWD Tractors	72	0.29	0.59	260	0.33	2.06	-0.04
Other machinery	72	7.89	4.78	260	5.93	4.40	1.96***
Cold chamber	72	0.39	0.49	260	0.16	0.37	0.23***
Wire fence	71	0.14	0.35	258	0.18	0.38	-0.04

IDB Development Effectiveness

Pre-treatment characteristics (PREDEG)

	Treated			Control			
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Mean difference
Technology and Management							
Administrator	71	0.08	0.28	258	0.10	0.30	-0.02
Technical assistance	71	0.86	0.35	258	0.53	0.50	0.33***
Registers	71	0.77	0.42	258	0.47	0.50	0.31***
Undercover Sowing	71	0.04	0.20	258	0.00	0.06	0.04***
Irrigation systems	71	0.79	0.41	258	0.53	0.50	0.26***
Health tretment	71	0.01	0.12	258	0.04	0.19	-0.02
Other uses of land							
Vineyard	71	1.00	0.00	258	1.00	0.00	0.00
Market garden	71	0.27	0.45	258	0.42	0.49	-0.15**
Cereals	71	0.06	0.23	258	0.07	0.26	-0.02
Meadow	71	0.07	0.26	258	0.09	0.29	-0.02
Wood	71	0.14	0.35	258	0.22	0.42	-0.08*
Pasture	71	0.01	0.12	258	0.08	0.27	-0.06**
Livestock							
Cows	71	0.13	0.34	258	0.34	0.48	-0.22***
Sheeps	71	0.00	0.00	258	0.04	0.19	-0.04*
Porks	71	0.20	0.65	258	1.33	7.75	-1.14
Access to road							
Motorized access	71	0.31	0.47	258	0.36	0.48	-0.05
Non motorized access	71	0.07	0.26	258	0.05	0.23	0.02
Improved access	71	0.62	0.49	258	0.58	0.49	0.04
Permanent access	71	0.96	0.20	258	0.98	0.15	-0.02
Other infrastructure							
Phone	71	0.89	0.32	258	0.85	0.36	0.04
Electricity	72	0.97	0.17	260	0.96	0.20	0.01

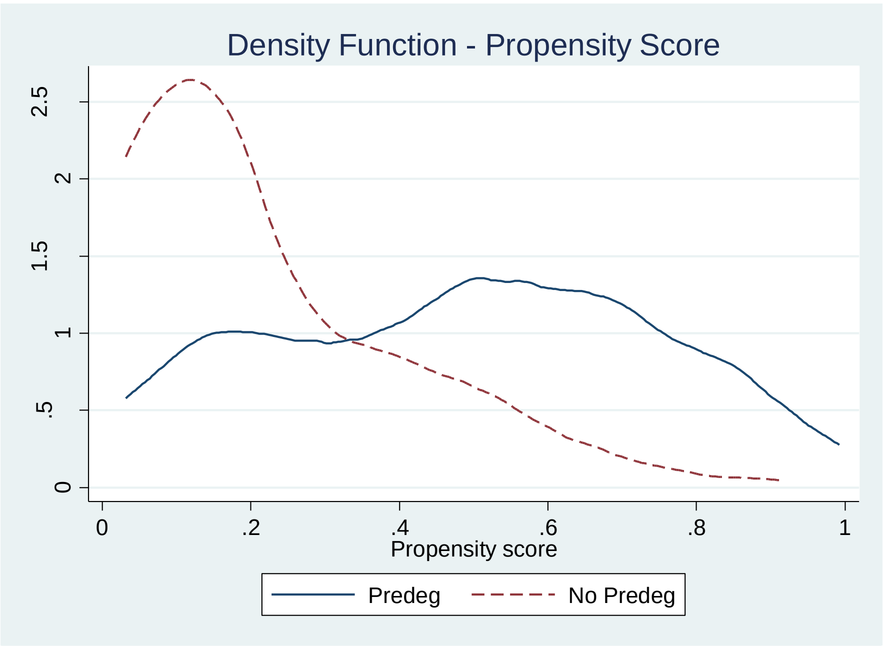
Participation model (PREDEG)

Variable	Coef.	Std. Err
Characteristics of the producers		
Age	-0.013	0.01
Gender	0.441	0.45
Education	0.051	0.03*
Foreign	-0.063	0.47
Individual	0.319	0.49
Company	0.471	0.56
Characteristics of the farms		
Size (Plants)		
Micro	-0.426	0.36
Small	0.886	0.31***
Small-Medium	0.830	0.34**
Medium	1.504	0.43***
Large	0.819	0.6
Location	0.230	0.34
Additional farm	-0.178	0.5
Employment		
Total Employment	0.053	0.03*
Skilled labour	-0.292	0.14**
Residents	0.004	0.05
Machine and equipment		
New tractors	-0.232	0.28
Cold chamber	0.365	0.24
Wire fence	-0.051	0.28
Technology and Management		
Administrator	-0.264	0.39
Technical assistance	0.662	0.26***
Registers	0.437	0.23*
Irrigation systems	0.570	0.23**
Health treatment	-0.204	0.62
Access to road and utilities		
Permanent access	-0.333	0.53
Phone	-0.492	0.34
Electricity	0.622	0.82
Constant	-3.252	1.33**
# obs.	325	
Pseudo R2	0.3315	

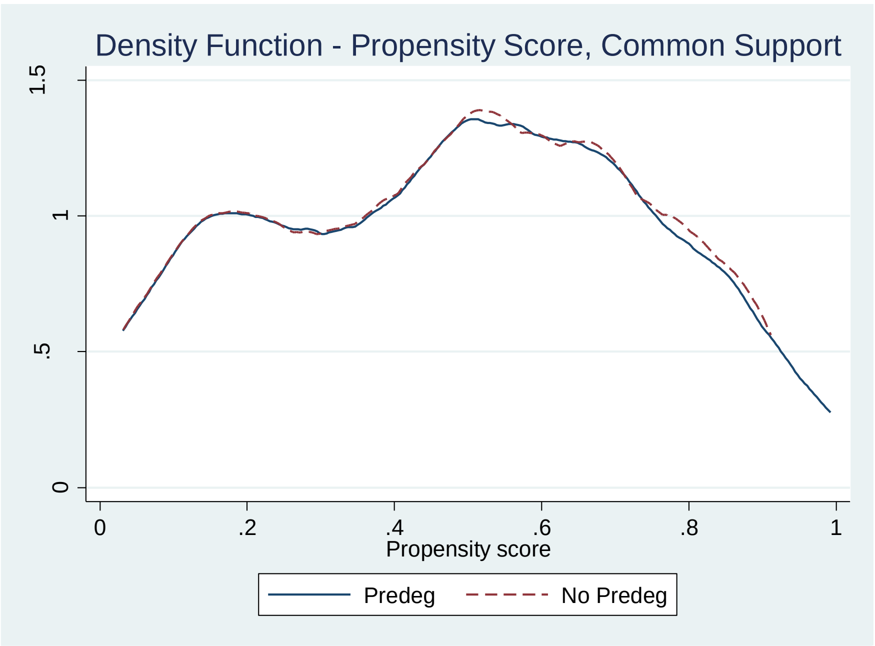
Development Effectiveness

Assessing the quality of the matching (PREDEG)

Propensity score distribution



Before Matching



After Matching

Assessing the quality of the matching (PREDEG)

Balancing of covariates

		Mean				t-test		Bal
Variable	**Sample**	**Treated**	**Control**	**% Bias**	**% reduct |bias|**	**t**	**p>|t|**	**Y/N**
Age	Unmatched	47.14	50.66	-28.7		-5.08	0.000	N
	Matched	47.14	46.24	7.4	74.4	0.46	0.646	Y
Gender	Unmatched	0.97	0.91	25		4.04	0.000	N
	Matched	0.97	0.99	-6.1	75.5	-0.58	0.563	Y
Education	Unmatched	11.42	10.46	25.8		4.72	0.000	N
	Matched	11.42	12.11	-18.5	28.2	-1.17	0.245	Y
Foreign	Unmatched	0.06	0.06	-2.5		-0.46	0.644	N
	Matched	0.04	0.07	-12	-370.8	-0.72	0.470	Y
Individual	Unmatched	0.80	0.89	-23.4		-4.59	0.000	N
	Matched	0.80	0.77	7.8	66.6	0.41	0.684	Y
Company	Unmatched	0.15	0.06	30.1		6.22	0.000	N
	Matched	0.15	0.18	-9.1	69.6	-0.44	0.657	Y

Development Effectiveness

Assessing the quality of the matching (PREDEG)

Balancing of covariates

Variable	Sample	Mean Treated	Mean Control	% Bias	% reduct |bias|	t-test t	t-test p>|t|	Bal Y/N
Micro	Unmatched	0.08	0.29	-54.6		-8.96	0.000	N
	Matched	0.08	0.04	11.3	79.4	1.03	0.305	Y
Small	Unmatched	0.38	0.22	35.5		6.86	0.000	N
	Matched	0.38	0.45	-15.7	55.9	-0.85	0.398	Y
Small-Medium	Unmatched	0.24	0.12	32.1		6.43	0.000	N
	Matched	0.24	0.23	3.7	88.3	0.2	0.844	Y
Medium	Unmatched	0.15	0.04	37.9		8.3	0.000	N
	Matched	0.15	0.10	19.3	49	1.01	0.316	Y
Large	Unmatched	0.06	0.03	12.2		2.44	0.015	N
	Matched	0.06	0.07	-6.9	43.2	-0.34	0.733	Y
Cold chambers	Unmatched	0.39	0.16	52.6		10.53	0.000	N
	Matched	0.38	0.37	3.3	93.8	0.17	0.863	Y
Registers	Unmatched	0.77	0.47	67.2		11.72	0.000	N
	Matched	0.77	0.77	0	100	0	1.000	Y
Technical assistance	Unmatched	0.86	0.53	76.2		12.74	0.000	N
	Matched	0.86	0.82	9.8	87.1	0.68	0.498	Y
Irrigation	Unmatched	0.79	0.53	56.5		9.79	0.000	N
	Matched	0.79	0.73	12.3	78.1	0.78	0.435	Y

Development Effectiveness

Robustness check (PROMSA)

Outcome	Nearest Neighbor (1)	Nearest Neighbor (5)	Caliper (0.001)	Caliper (0.0001)	Radius (0.001)	Normal Kernel	Epan. Kernel
Associability	0.6393	0.6707	0.6571	0.6794	0.6797	0.674	0.6813
	(.0443)***	(.0315)***	(.0477)***	(.051)***	(.0325)***	(.0289)***	(.0293)***
	[.0709]***	[.0812]***	[.0913]***	[.0876]***	[.0758]***	[.0785]***	[.0717]***
Sales location	0.1259	0.0129	0.1488	0.1528	0.013	-0.0197	-0.0127
	(.067)*	(.0425)***	(.0717)**	(.0775)**	(.044)***	(.0336)***	(.0348)***
	[.0801]	[.0822]	[.0838]*	[.1099]	[.0813]	[.0647]	[.0743]
Type of purchaser	0.1378	0.1081	0.1494	0.1346	0.1223	0.1136	0.111
	(.0463)***	(.0323)***	(.0489)***	(.0481)***	(.0341)***	(.0273)***	(.0282)***
	[.0623]**	[.0666]	[.0635]**	[.0486]***	[.0601]**	[.0485]**	[.0583]*
Access to credit	0.4122	0.3384	0.4328	0.4585	0.347	0.3573	0.3602
	(.0458)***	(.0379)***	(.0479)***	(.0517)***	(.0426)***	(.034)***	(.0351)***
	[.0864]***	[.0975]***	[.092]***	[.0793]***	[.0759]***	[.0742]***	[.0818]***
Access to formal credit	0.3441	0.3369	0.3782	0.4244	0.3471	0.3461	0.3464
	(.0385)***	(.0325)***	(.0409)***	(.0445)***	(.0346)***	(.0303)***	(.0307)***
	[.0808]***	[.1071]***	[.1091]***	[.0989]***	[.0986]***	[.1111]***	[.0958]***

Results: density of plantation (PREDEG)

Density of plantation	Apples			Peaches		
Predeg	105.9	-	-	50.3	-	-
	(42.7)**			(34.8)		
First year	-	62.3	-	-	5.8	-
		(38.6)			(27.3)	
Second year	-	156.7	-	-	48.0	-
		(56.8)***			(39.5)	
Third year	-	157.9	-	-	68.6	-
		(50.8)***			(42.2)	
Fourth year	-	112.1	-	-	109.5	-
		(43.2)**			(56.2)*	
Fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Time dummies	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Common support	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	481	481	481	515	515	515
Number of producers	110	110	110	137	137	137

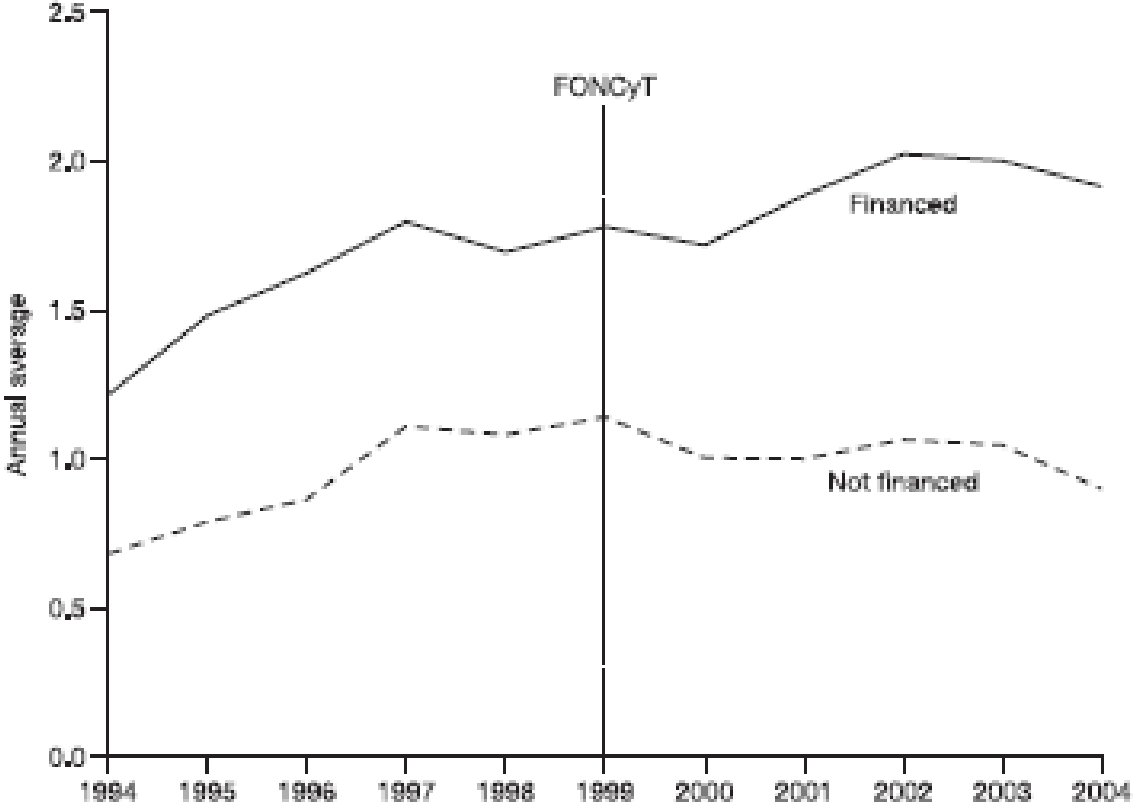
Robust standard errors clustered at producer level in parentheses

* significant at 10%, ** significant at 5%, *** significant at 1%

Development Effectiveness

Parallel trends assumption (FONCYT)

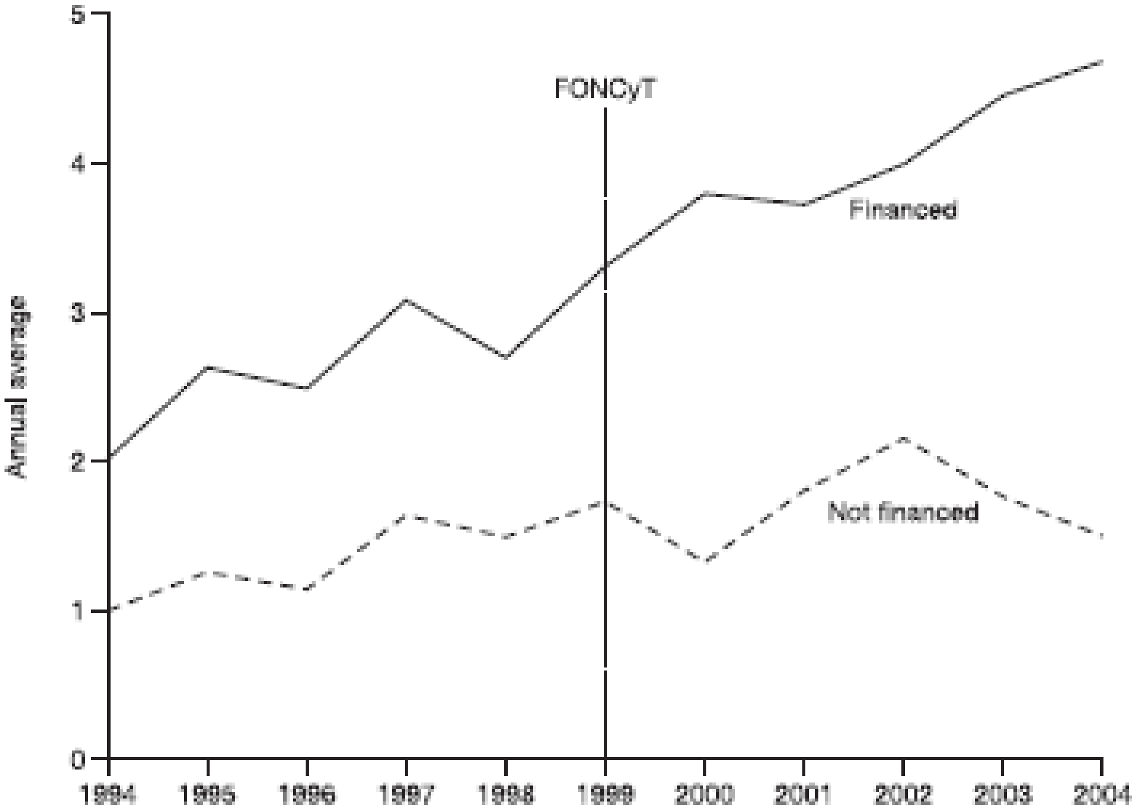
Number of publications



Development Effectiveness

Parallel trends assumption (FONCYT)

Quality of publications



Note: The quality index is the impact factor.

Results: scientific productivity (FONCYT)

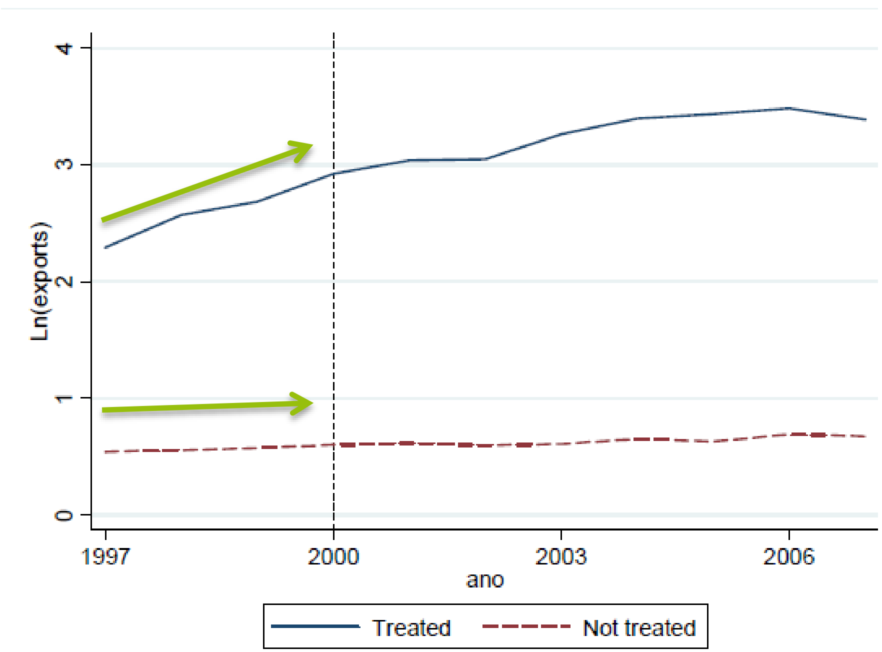
	Publications				Impact Index			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Foncyt	.195	2.436	.214	2.442	.911	4.219	.726	5.548
	(.086) **	(.450) ***	(.111)*	(.573)***	(.207)***	(1.18)***	(.252)***	(1.53)***
	[.106]*	[.487]***	[.136]	[.701]***	[.255]***	[1.29]***	[.314]**	[1.65]***
Age		.042		0.035		0.136		0.164
		(.011)***		(.015)**		(.033)***		(.041)***
		[.013]***		[.016]**		[.038]***		[.049]***
Foncyt* Age in 2005		-.038		-.035		-.071		-.085
		(.007) ***		(.009)***		(.019)***		(.025)***
		[.008]***		[.010]***		[.021]***		[.026]***
Foncyt* Doctorate		.013		-.108		.763		.369
		(.112)		(.141)		(.275)***		(.341)
		[.137]		[.172]		[.312]**		[.380]
Foncyt* Gender		-.175		-.319		-.031		-.658
		(.118)		(.148)**		(.310)		(.385)*
		[.158]		[.202]		[.373]		[.450]
Sample	3549	3549	2308	2308	3548	3548	2309	2309
R-squared	0.627	0.631	0.661	0.664	0.533	0.536	0.504	0.509

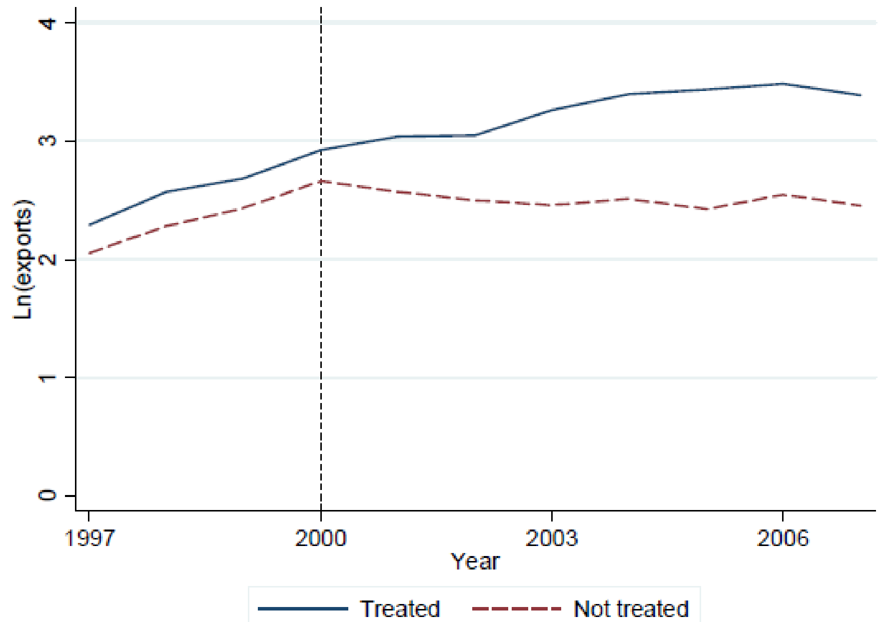
Notes: all regressions include researcher fixed effects and time dummies. Heteroskedasticity robust standard errors are shown in parentheses. Standard errors clustered at the researcher level are shown in brackets. Results in Columns (3), (4), (7), and (8) use the sample restricted to common support. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.

Development Effectiveness

Matching trends (PCB)

- PSM can be used to make the DD more credible by selecting a comparison group with similar trends ex–ante. De Negri et al. combine the methods to evaluation the impact of public credit lines in Brazil.





Results: scientific productivity (PCB)

Table 4: impact on exports (full sample)

	(1)	(2)	(3)
BNDES	0.4765***	0.3880***	0.3896***
	(0.095)	(0.080)	(0.080)
lage		0.0449	0.0434
		(0.038)	(0.039)
lskill		0.0347***	0.0338***
		(0.012)	(0.012)
lwage		0.0395***	0.0399***
		(0.010)	(0.010)
patentes		-0.0082	-0.0075
		(0.017)	(0.017)
finep		1.0418***	1.0307***
		(0.368)	(0.368)
premio		5.8490***	5.8482***
		(0.064)	(0.064)
limp		0.0717***	0.0717***
		(0.004)	(0.004)
Constant	0.7106***	-0.1515	-0.1544
	(0.007)	(0.115)	(0.115)
Fixed effects	✓	✓	✓
Time dummies	✓	✓	✓
Industry-year interactions	×	×	✓
R^2	0.02	0.297	0.30
Obs.	492480	492480	492480
No. of firms	49248	49248	49248

Cluster-robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Development Effectiveness

Robustness of the results (COLCIENCIAS)

- Crespi, Maffioli y Melendez use a “placebo” test to check the validity of these results

	Ln(employment)	**Ln(labor prod)**	**Number of products**	**Market share**
Colciencias	0.159**	0.149**	0.124*	0.011
	[0.077]	[0.064]	[0.068]	[0.012]
Fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Time dummies	*Yes*	*Yes*	*Yes*	*Yes*
Common support	*Yes*	*Yes*	*Yes*	*Yes*
Observations	36,468	36,468	36,473	36,473
Number of firms	2,997	2,997	2,997	2,997

Using leads to test results (COLCIENCIAS)

- In this case the test shows that the impact on employment could still be biased.

	Ln(employment)	**Ln(labor prod)**	**Number of products**	**Market share**
Colciencias +1	0.163*	-0.031	0.132	0.015
	[0.093]	[0.073]	[0.085]	[0.013]
Colciencias+2	0.120*	0.022	0.093	0.015
	[0.068]	[0.063]	[0.086]	[0.013]
Fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Time dummies	*Yes*	*Yes*	*Yes*	*Yes*
Common support	*Yes*	*Yes*	*Yes*	*Yes*
Observations	39,409	39,409	39,415	39,415
Number of firms	2,997	2,997	2,997	2,997

IDB
Development Effectiveness

Resources

- PSM: Heinrich, Maffioli Vazquez (2010), A Premier for Applying Propensity-Score Matching
- DID: Blundell & Costa Dias (2002), Alternative Approaches to Evaluation in Empirical Microeconomics
- Angrist & Pischke (2009), Mostly Harmless Econometrics
- Gertler et al. (2010), Impact Evaluation in Practice
- IDB methodological guidelines (how to evaluate projects in…agriculture, innovation, tourism, cluster development…)
- SPD experts

Development Effectiveness