

Valutare gli effetti delle politiche

Luca Mo Costabella

ASVAPP – Associazione per lo Sviluppo della Valutazione e l'Analisi delle Politiche Pubbliche

Valutazione degli effetti, o controfattuale

2

Studio della relazione *causale* tra un evento (trattamento) e una condizione (outcome)

Nelle politiche pubbliche, studio dei cambiamenti prodotti dall'esposizione alla politica pubblica su certe condizioni della collettività

Due dimensioni di riferimento

3

Il **trattamento**

Condizione DI cui si stima l'effetto, identifica l'esposizione o meno alla politica

T = variabile trattamento

(spesso, ma non sempre, $T=\{0,1\}$)

Due dimensioni di riferimento

4

L'outcome

**Condizione SU cui si stima l'effetto, da tradurre in
grandezza osservabile e misurabile**

Y = outcome, o variabile risultato

Esempio: la tachipirina

5

Trattamento

T = prendere la tachipirina

Outcome

Y = temperatura

T ha un effetto su Y?

Cosa cambia per Y se T=0 o T=1?

Esempio: la formazione antinfortunistica

6

Trattamento

T = partecipare al corso

Outcome

**Y = avere avuto un incidente nel cantiere
nei 12 mesi dopo**

T ha un effetto su Y?

Cosa cambia per Y se T=0 o T=1?

L'effetto sui trattati è (o... sarebbe) una differenza

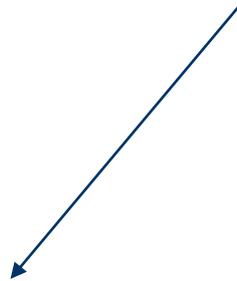
7

- A- Qual è probabilità di infortunio dopo il corso? (% infortuni tra i formati)
B- Quale sarebbe stata senza il corso? (% infortuni attesa)

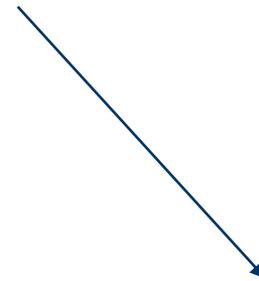
Effetto del corso = $A - B$

A = **fattuale**

B = **controfattuale**



Si osserva



Va **stimato**

Come si stima il controfattuale?

8

con un **gruppo di controllo**
(control group design)

oppure **osservando il passato**
(pre-post design)

Control group design

9

Ho un gruppo di “trattati”, di cui osservo la situazione dopo il “trattamento”

E un gruppo di “non trattati”, o “controlli”, la cui situazione potrebbe dirci come sarebbero stati i primi se non fossero stati “trattati”

Cosa è successo ai beneficiari dell'intervento?
E ai non beneficiari?

Tasso di infortuni tra i
partecipanti al corso

Tasso di infortuni tra i non
partecipanti

La differenza può rappresentare l'effetto dell'intervento

Ma questo valore stima in modo credibile il controfattuale?

Le minacce alla validità del *comparison group design*

11

Perché i partecipanti hanno partecipato al corso e gli altri no?

Ci deve essere qualche meccanismo, osservabile o (più spesso) non osservabile, che ha determinato la partecipazione

Qualcosa che rende **sistematicamente diversi** partecipanti e non partecipanti

Quindi non è detto che la stima del controfattuale sia credibile:

Se i due gruppi sono diversi, avrebbero potuto mostrare condizioni diverse anche se nessuno avesse partecipato ai corsi

Come migliorare il *comparison group design*

12

Le differenze di partenza tra i due gruppi sono la causa del cosiddetto *selection bias*

Eliminare (o ridurre) le differenze di partenza è la strada da seguire per ridurre questa distorsione

- possiamo farlo a posteriori, cercando di pareggiare i due gruppi posti a confronto
- possiamo farlo a priori, studiando meccanismi di selezione che ci permettano di ridurre fin da principio le differenze di partenza

Gli esperimenti randomizzati

13

L'idea è semplice: dato un bacino di potenziali utenti, **tiro a caso** per decidere chi partecipa

Così facendo, non ho motivo di supporre che ci siano differenze di partenza tra i due gruppi: il lancio di moneta crea gruppi simili

Quindi i non partecipanti rappresentano in modo credibile il controfattuale

Certo questa strada non è molto apprezzata dal decisore pubblico...

In ogni caso, pensare per tempo alla valutazione aiuta poi il lavoro

Esempio: TSUNAMI (2016-19)

14

Traineeships as a Springboard out of Unemployment for those Affected by Mental Illness

Regione Piemonte

Agenzia Piemonte Lavoro

Mestieri e ForCoop

Fondazione Adecco per le Pari Opportunità

CREARE

ASVAPP



Project awarded under the 2015 Call for Proposals
on social policy innovations supporting reforms in social services

L'inserimento lavorativo dei disabili psichici

La sfida

Le loro chance di occupazione sono la metà degli altri disabili.
come aiutare i disoccupati a trovare lavoro?

Ipotesi di intervento: il supported employment

Un job coach

- offre un servizio intensivo e molto mirato
- prepara (in qualche mese) la persona al lavoro
- la inserisce nel mercato competitivo con un tirocinio
- monitora la situazione
- supporta l'inserimento dopo il tirocinio

Durata del trattamento: max un anno

16



Il disegno di valutazione

17



Gli effetti su tirocini e occupazione

	Controlli	Gruppo sperimentale	Differenza
Avviati al lavoro in 12 mesi	0.25	0.23	-0.02
Avviati a tirocinio in 12 mesi	0.08	0.20	+0.12**
N	735	851	

Gli effetti su tirocini e occupazione

	Controlli	Gruppo sperimentale	Differenza
Avviati al lavoro in 12 mesi	0.25	0.23	-0.02
Avviati a tirocinio in 12 mesi	0.08	0.20	+0.12**
N	735	851	

Ma molti non sono stati trattati davvero...

Solo il 51% ha partecipato!

Di questi, solo uno su tre è arrivato al tirocinio!

Esperimento non “integro” -> Risultati meno affidabili, da correggere (con fatica)

E se non c'è selezione casuale?

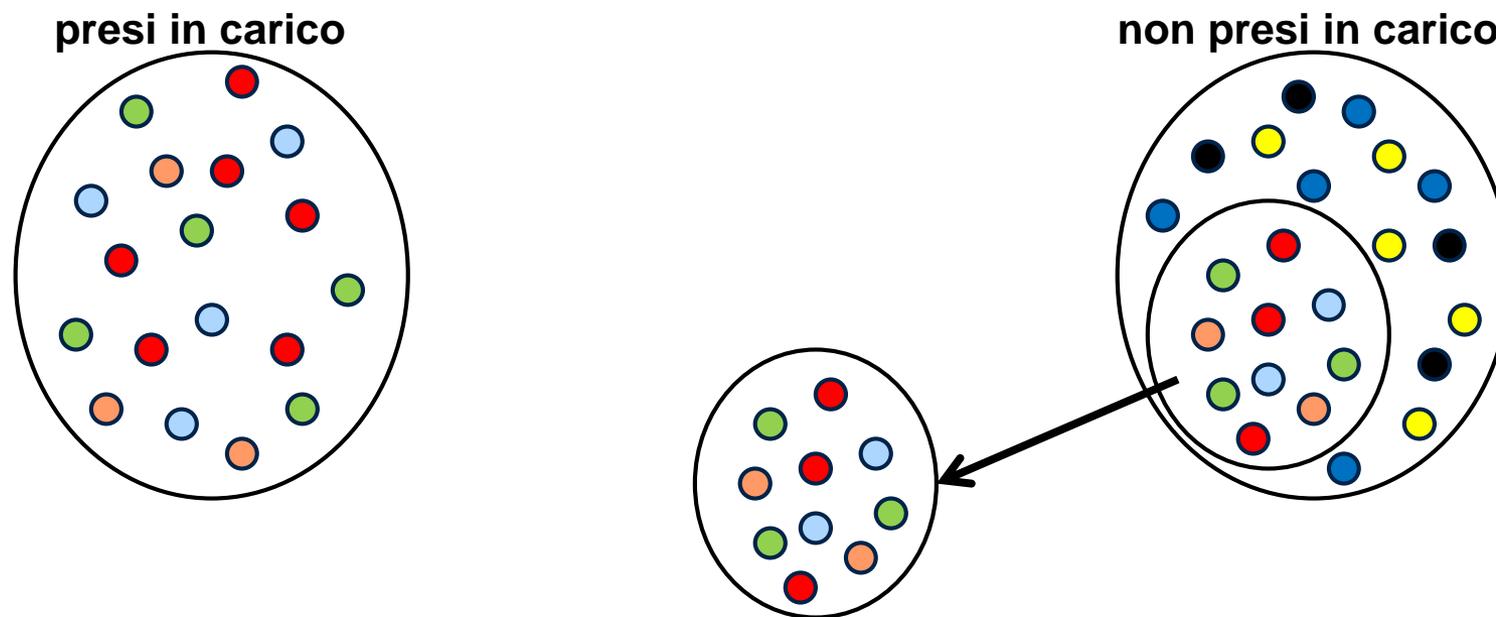
Come provare a risolvere il problema

22

Siamo in grado di confrontare soggetti **simili**?

La logica del *matching*

23



I presi in carico vanno confrontati con un sottoinsieme
di non presi in carico

un sottoinsieme formato da quelli ad essi più simili

Un esempio: il Contratto di Ricollocazione della Sardegna

24

Politica (nazionale) «innovativa» per il reinserimento lavorativo

Integrazione politiche attive e passive

Servizi erogati da soggetti accreditati pubblici e privati

Retribuzione a risultato

Sperimentazione regionale in Sardegna: il CRiS

25

Target: 2.708 beneficiari della prima concessione
della mobilità in deroga (anno 2014)

1.421 invitati -> **1.241 partecipanti**
1.287 mai invitati

Divisi in quattro «classi di occupabilità»

Quattro classi di occupabilità

26

classe di occupabilità	Classe4 p≤20%	Classe3 20%<p≤27%	Classe2 27%<p≤35%	Classe1 p>35%
femmine	23%	39%	30%	29%
stranieri	0%	1%	1%	0%
25 anni o meno	0%	0%	2%	1%
26-35 anni	1%	14%	28%	24%
36-45 anni	2%	22%	47%	41%
46-55 anni	15%	53%	18%	28%
56 anni o più	82%	11%	5%	6%
obbligo scolastico o titolo inferiore	81%	75%	50%	62%
diploma o qualifica professionale	18%	23%	43%	32%
laurea o titolo superiore	1%	2%	7%	6%
hanno lavorato nei 6 mesi precedenti	4%	4%	8%	70%
hanno lavorato nei 24 mesi precedenti	5%	7%	14%	96%

Quali sono gli effetti del CRiS?

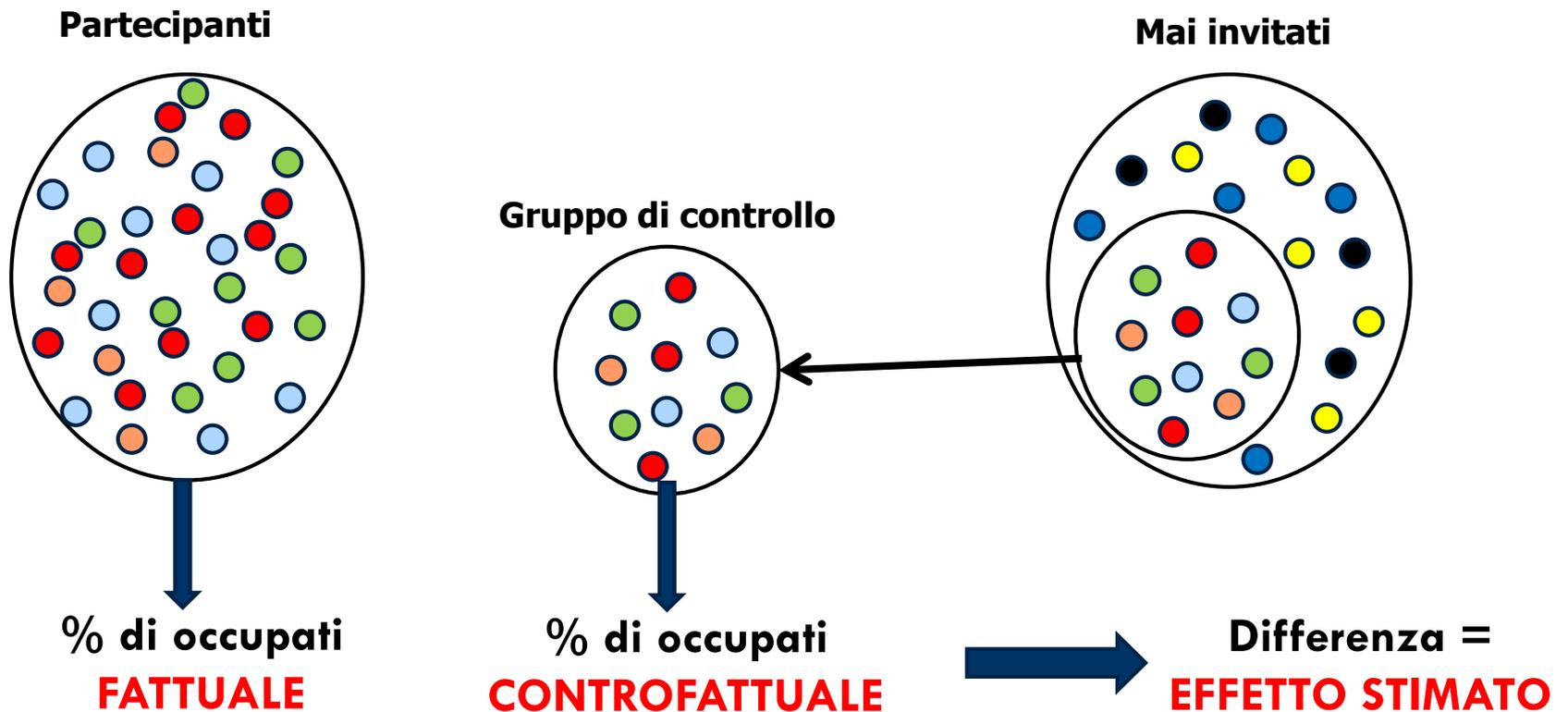
27

**Qual è la condizione lavorativa dei partecipanti
DOPO la partecipazione al CRiS?**

Quale **sarebbe stata la condizione lavorativa dei partecipanti
senza la partecipazione al CRiS?**

La logica del *matching*

28



Condizione necessaria per un buon matching di soggetti simili

29

**In teoria
la somiglianza deve riguardare
tutte le caratteristiche**

**Nella pratica
deve riguardare
un ricco set (tendente all'esaustivo)
di caratteristiche**

Il Caso del CRiS: osserviamo per tutti queste caratteristiche (dal SIL Sardegna):

30

- sesso;
- età;
- nazionalità;
- provincia di residenza;
- titolo di studio;
- storia lavorativa precedente:
- condizione lavorativa (occupato/non occupato) osservata mese per mese nei due anni precedenti l'inizio del CRiS;
- contratto di lavoro prevalente nei due anni precedenti;
- settore di lavoro prevalente nei due anni precedenti;
- qualifica di lavoro prevalente nei due anni precedenti;
- numero di lavori svolti nei due anni precedenti;
- percezione di sostegni al reddito all'inizio del CRiS.

Come procediamo

31

Definiamo un sistema di misura della distanza/somiglianza

es: distanza tra *propensity score*

Poi abbiniamo e confrontiamo soggetti simili

es: *nearest neighbour matching*

es: *radius matching*

es: *kernel matching*

Il *propensity score*

32

Il *propensity score* è la probabilità, stimata sulla base delle *observables*, che un'unità venga assegnata al trattamento

$$P(T = 1|X) = \frac{e^{(X\beta)}}{1+e^{(X\beta)}}$$

Una sorta di «sintesi intelligente» delle *observables*

Il risultato della stima del pscore è questo:

33

Nome		T	score
Pippo	maschio, 45 anni, diplomato, disocc. da 6 mesi, [...]	1	.395
Pluto	maschio, 19 anni, obbligo, disocc. da 20 mesi, [...]	1	.798
Topolino	maschio, 36 anni, laureato, disocc. da 9 mesi, [...]	1	.174
Minni	femmina, 33 anni, laureata, disocc. da 6 mesi, [...]	0	.710
Basettoni	maschio, 55 anni, diplomato, disocc. da 2 mesi, [...]	0	.143
Eta Beta	maschio, 51 anni, obbligo, disocc. da 9 mesi, [...]	0	.382

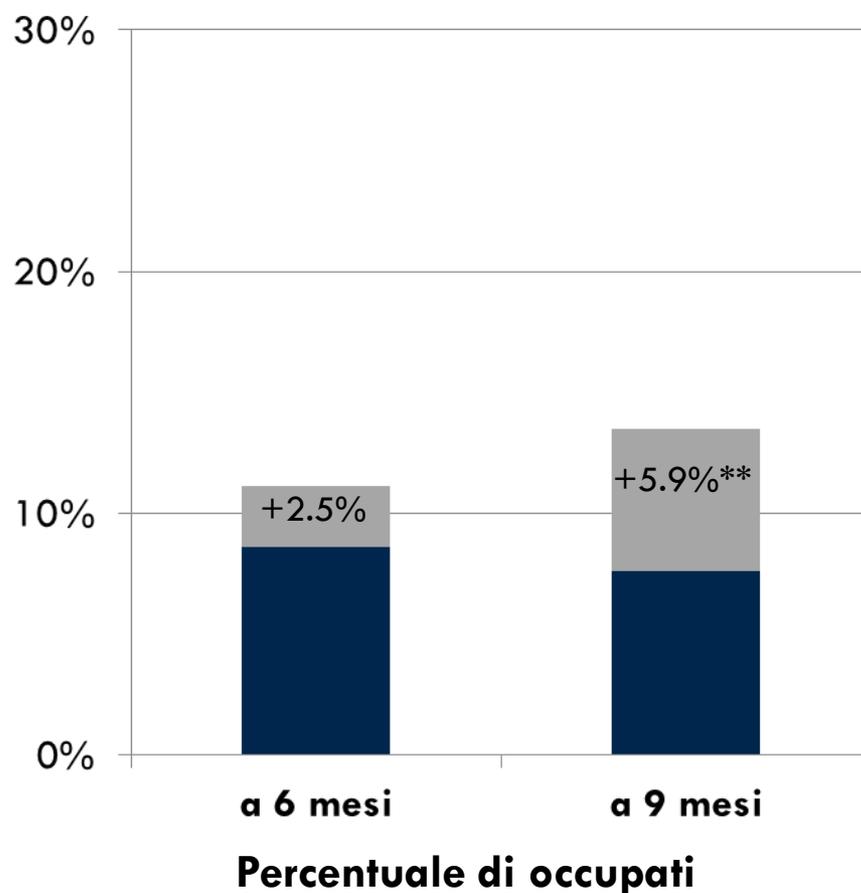
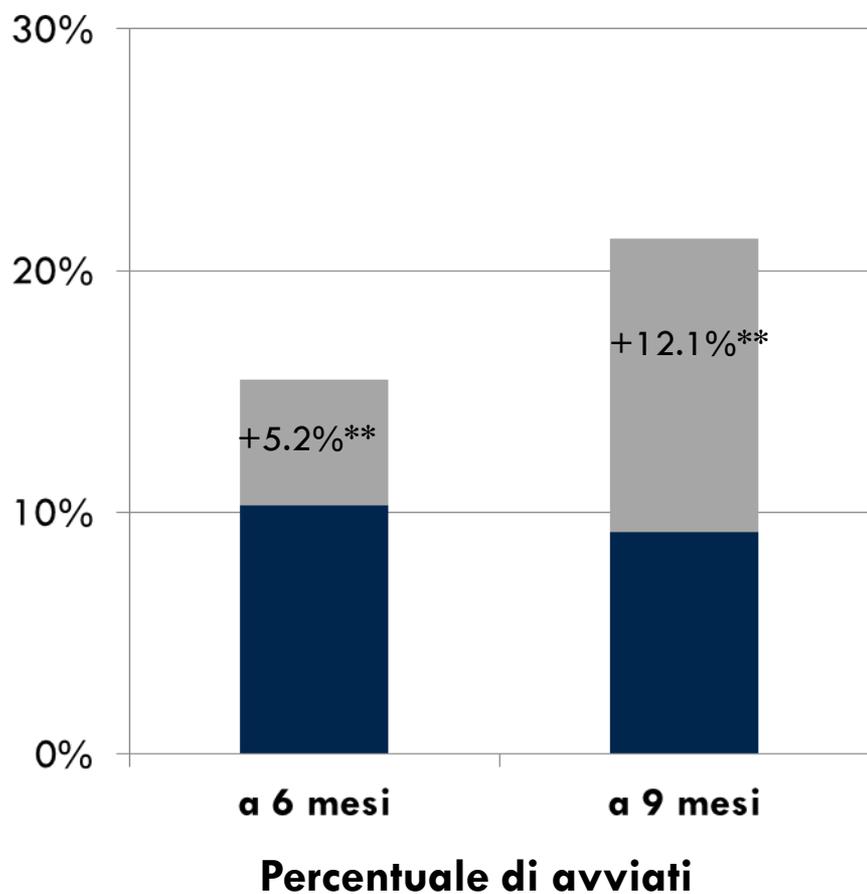
Ora abbiniamo soggetti simili

Il caso del CRiS: la composizione dei gruppi dopo l'abbinamento con Kernel matching

35	Partecipanti	Controlli	Controlli abbinati
Femmine	33%	23%	31%
Maschi	67%	77%	69%
Italiani	99%	99%	99%
Stranieri	1%	1%	1%
25 anni o meno	0%	1%	0%
26-35 anni	12%	17%	10%
36-45 anni	29%	26%	30%
46-55 anni	32%	20%	35%
56 anni o più	26%	36%	25%
Titolo dell'obbligo o inferiore	65%	72%	63%
Diploma quadriennale	31%	24%	32%
Laurea	4%	3%	5%
Hanno lavorato nei 6 mesi precedenti	19%	13%	16%
Non hanno lavorato nei 6 mesi precedenti	81%	87%	84%
Hanno lavorato nei 24 mesi precedenti	35%	30%	34%
Non hanno lavorato nei 24 mesi precedenti	65%	70%	66%
N	1.100	572	

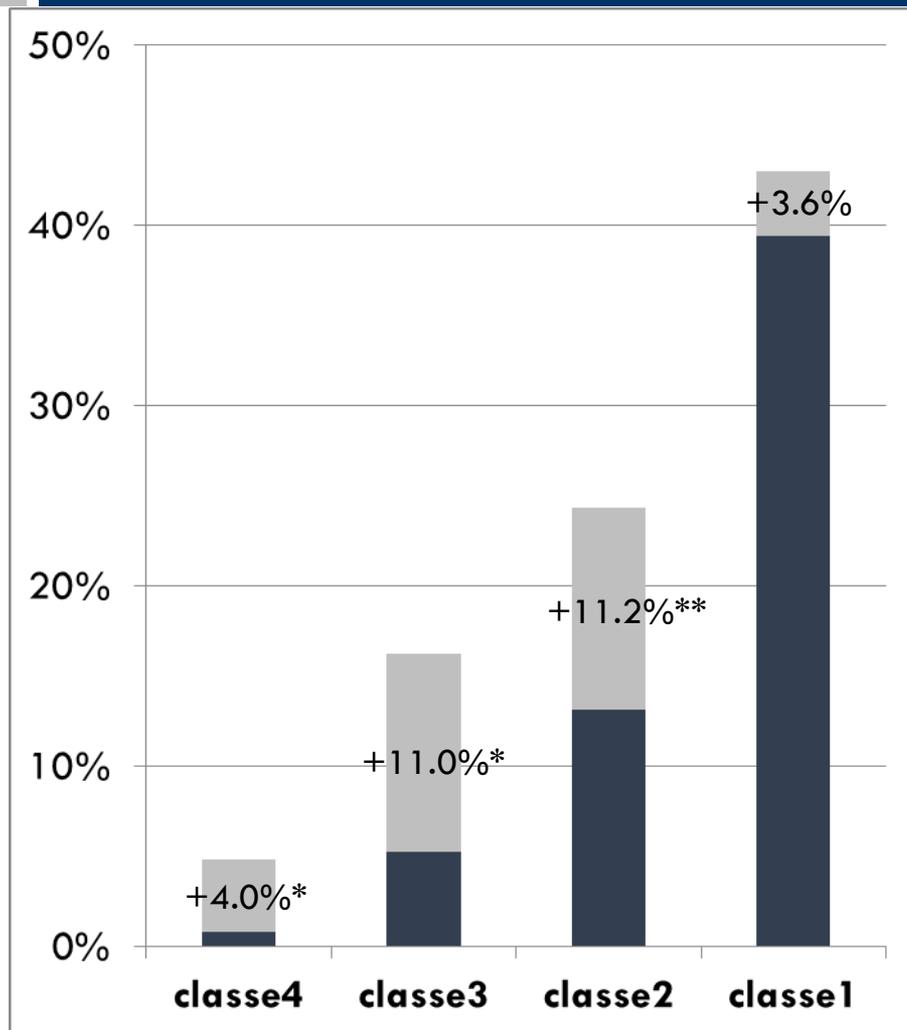
La stima degli effetti: confrontiamo gli esiti di trattati e controlli abbinati

36

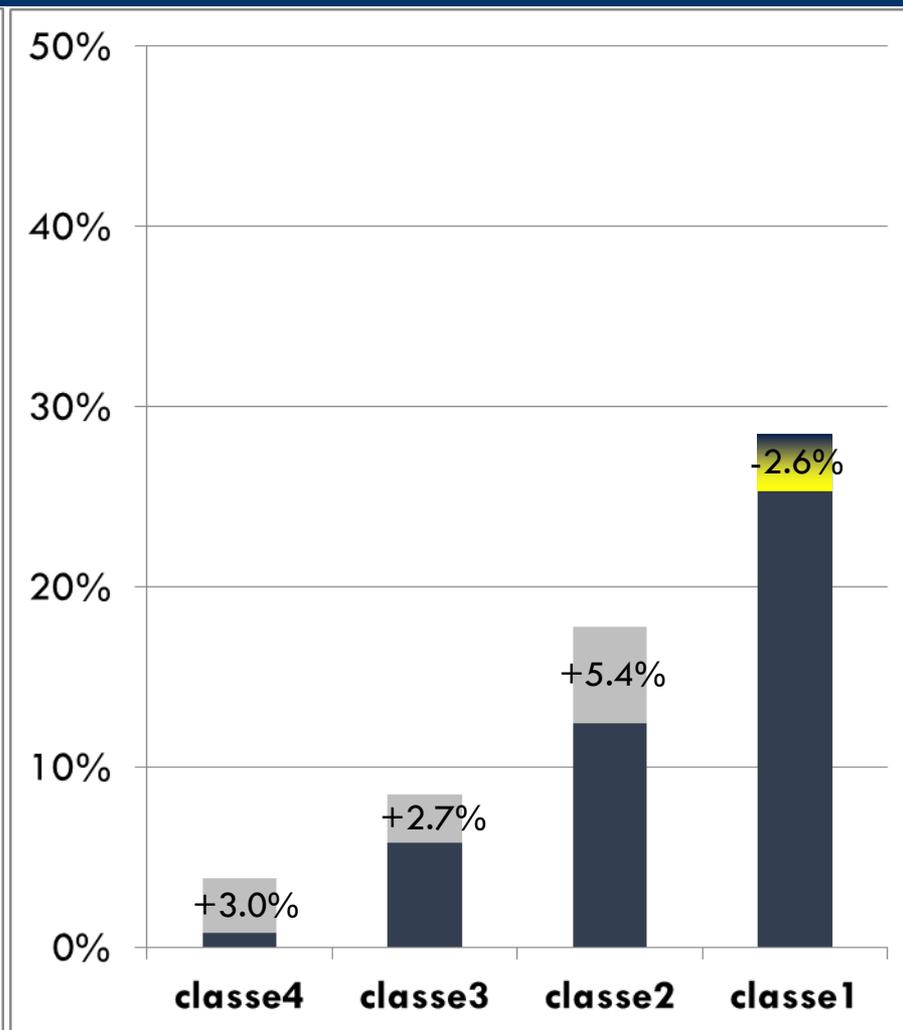


La stima degli effetti di breve periodo, per classe

37



Percentuale di avviati entro 9 mesi



Percentuale di occupati dopo 9 mesi

Quando i gruppi sono troppo diversi, non c'è abbinamento che tenga

38

Trattati tutti maschi, non trattati tutte donne

Trattati tutti 60enni, non trattati tutti 25enni

Trattati tutte grandi imprese, non trattati tutte PI

Ma vi sono dei casi particolari

39

I supporti dei due gruppi non si sovrappongono, ma si «toccano»

Vi sono poi dei casi particolari

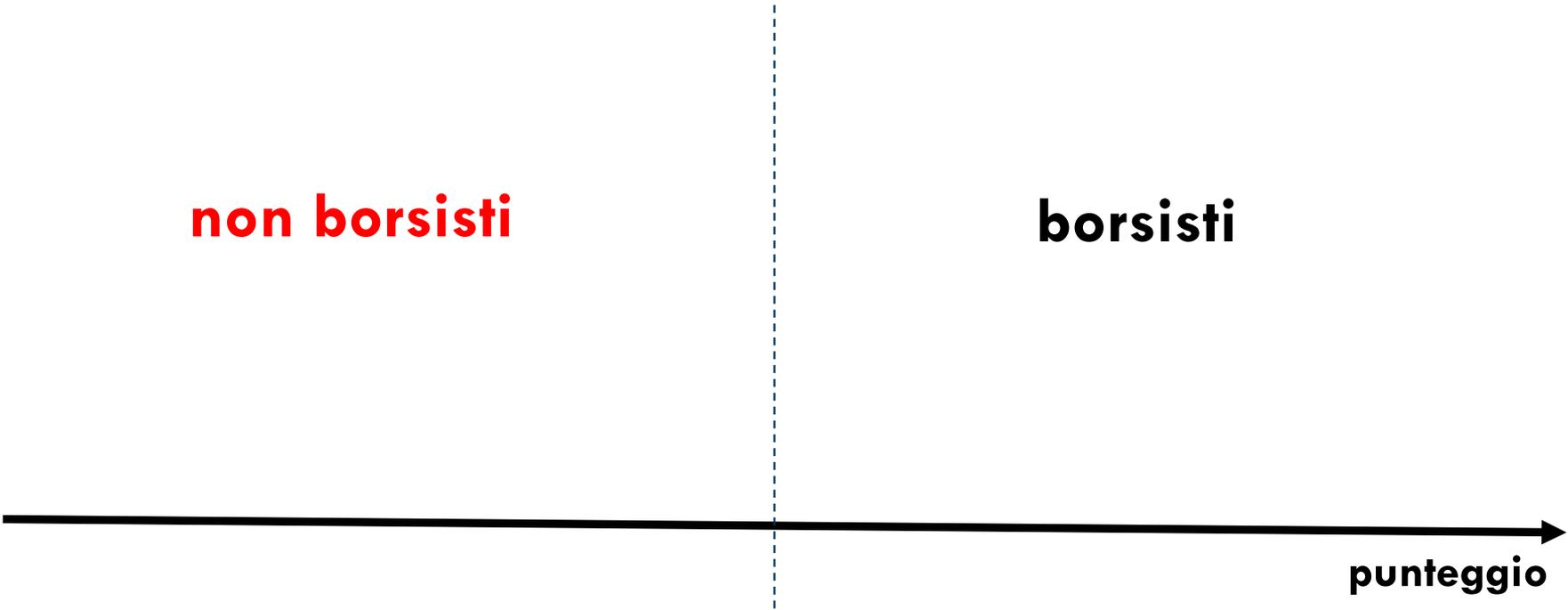
40

Assegnazione borse universitarie basata su graduatoria di merito e bisogno

non borsisti

borsisti

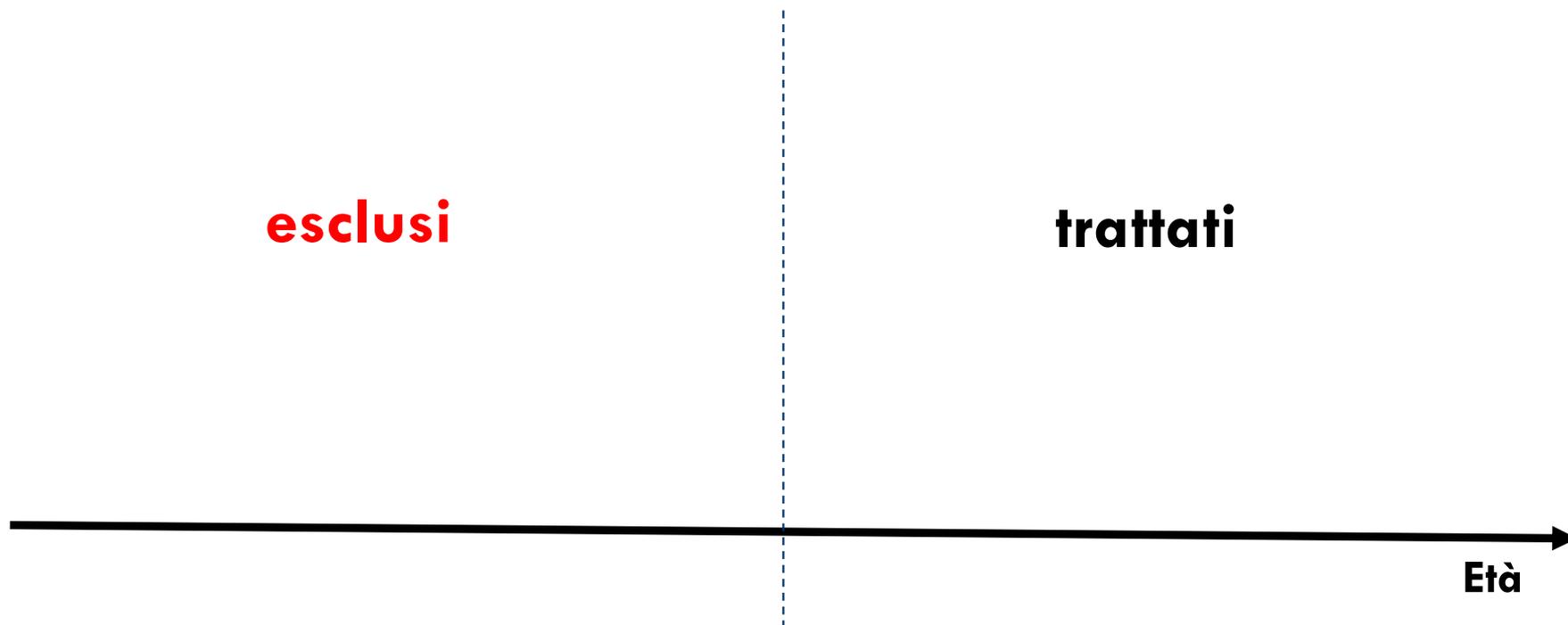
punteggio



Vi sono poi dei casi particolari

41

Riduzione dei ticket sanitari per gli over 65



Vi sono poi dei casi particolari

42

Politiche del lavoro con accesso prioritario ai disoccupati di lunga durata

esclusi

trattati

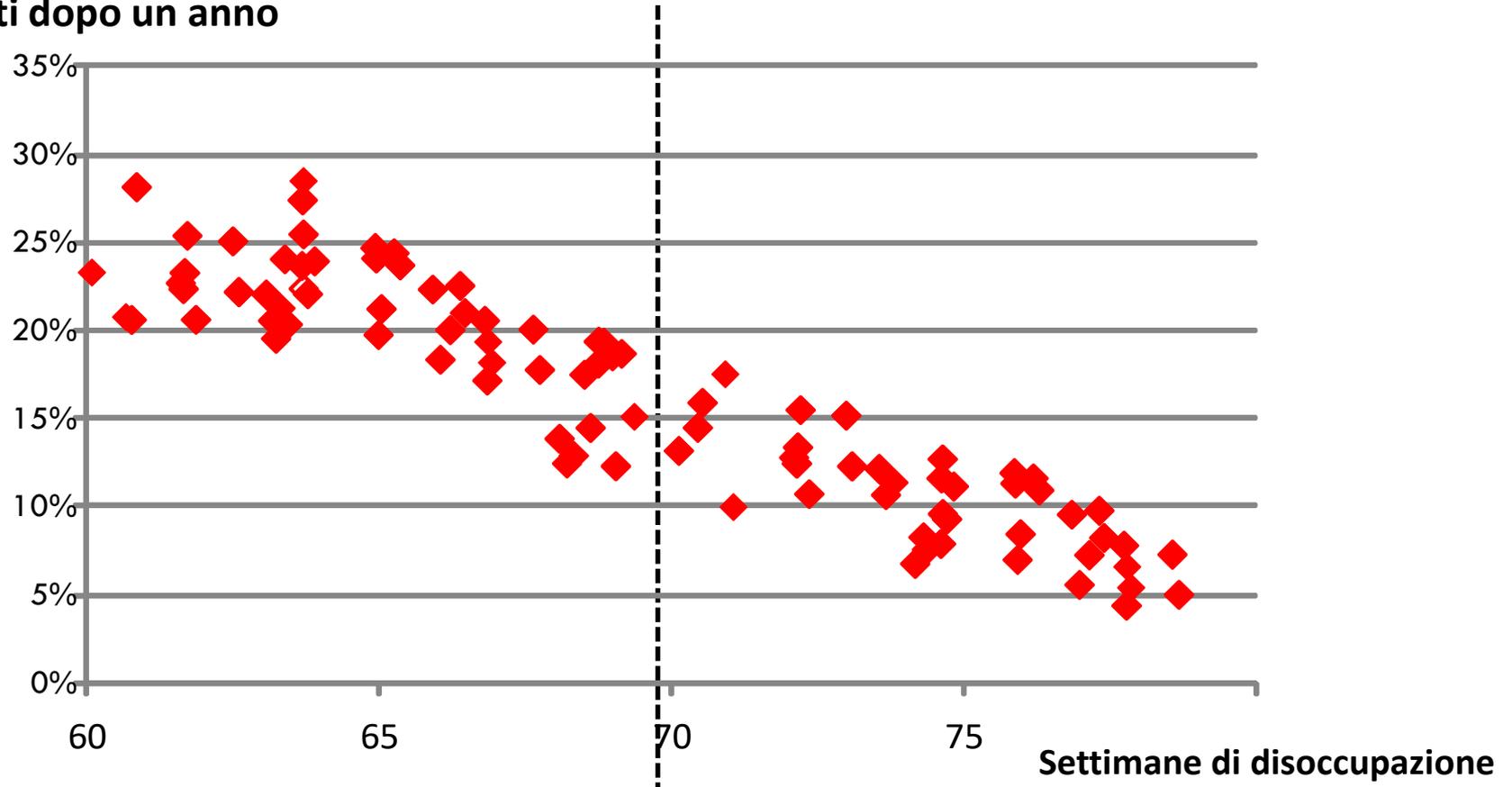
Anzianità di disoccupazione

A horizontal axis with an arrow pointing to the right, labeled 'Anzianità di disoccupazione'. A vertical dashed line is drawn across the axis, dividing the space into two regions. The region to the left of the dashed line is labeled 'esclusi' in red text. The region to the right of the dashed line is labeled 'trattati' in black text.

Cosa succederebbe senza trattamento?

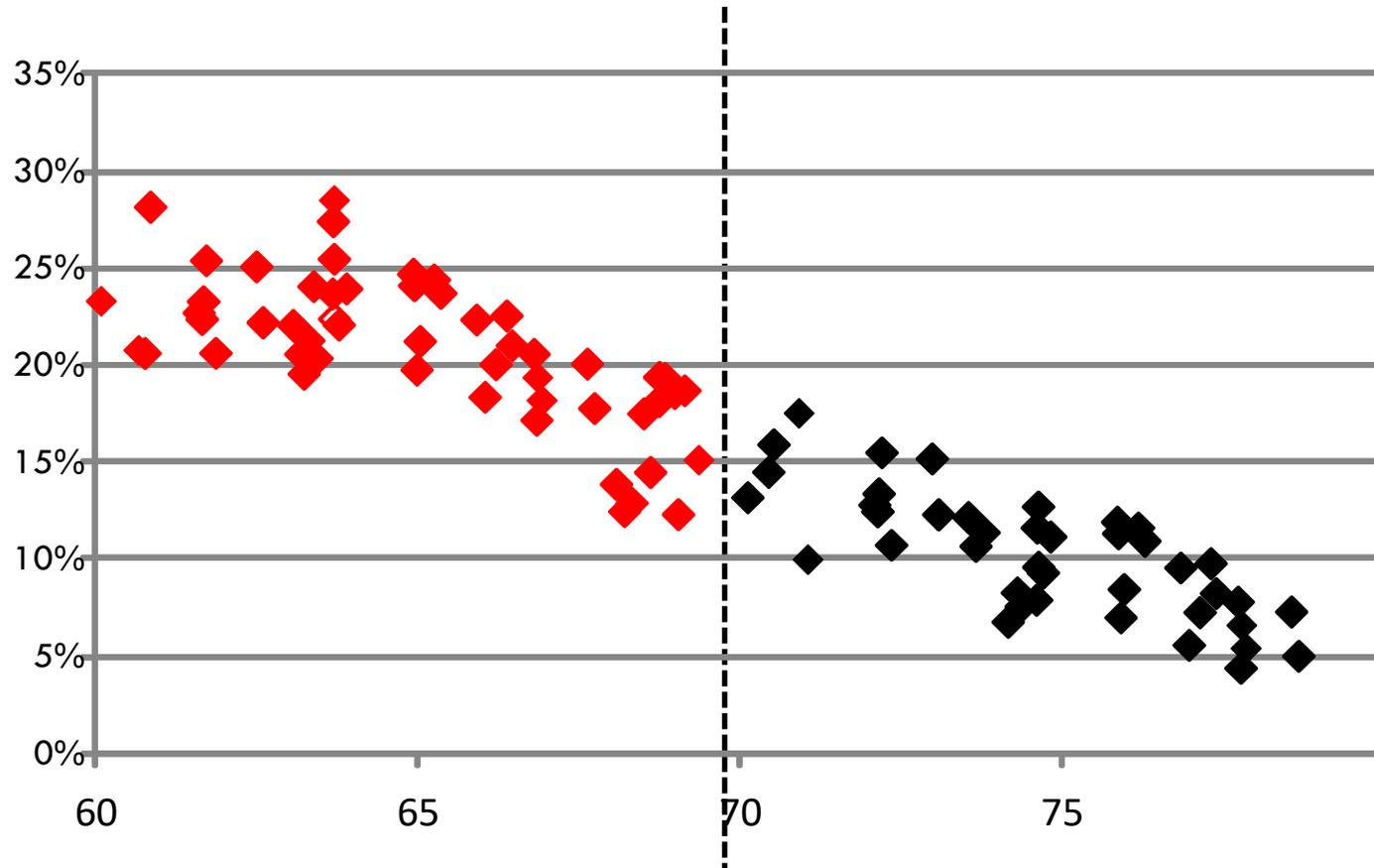
43

Occupati dopo un anno



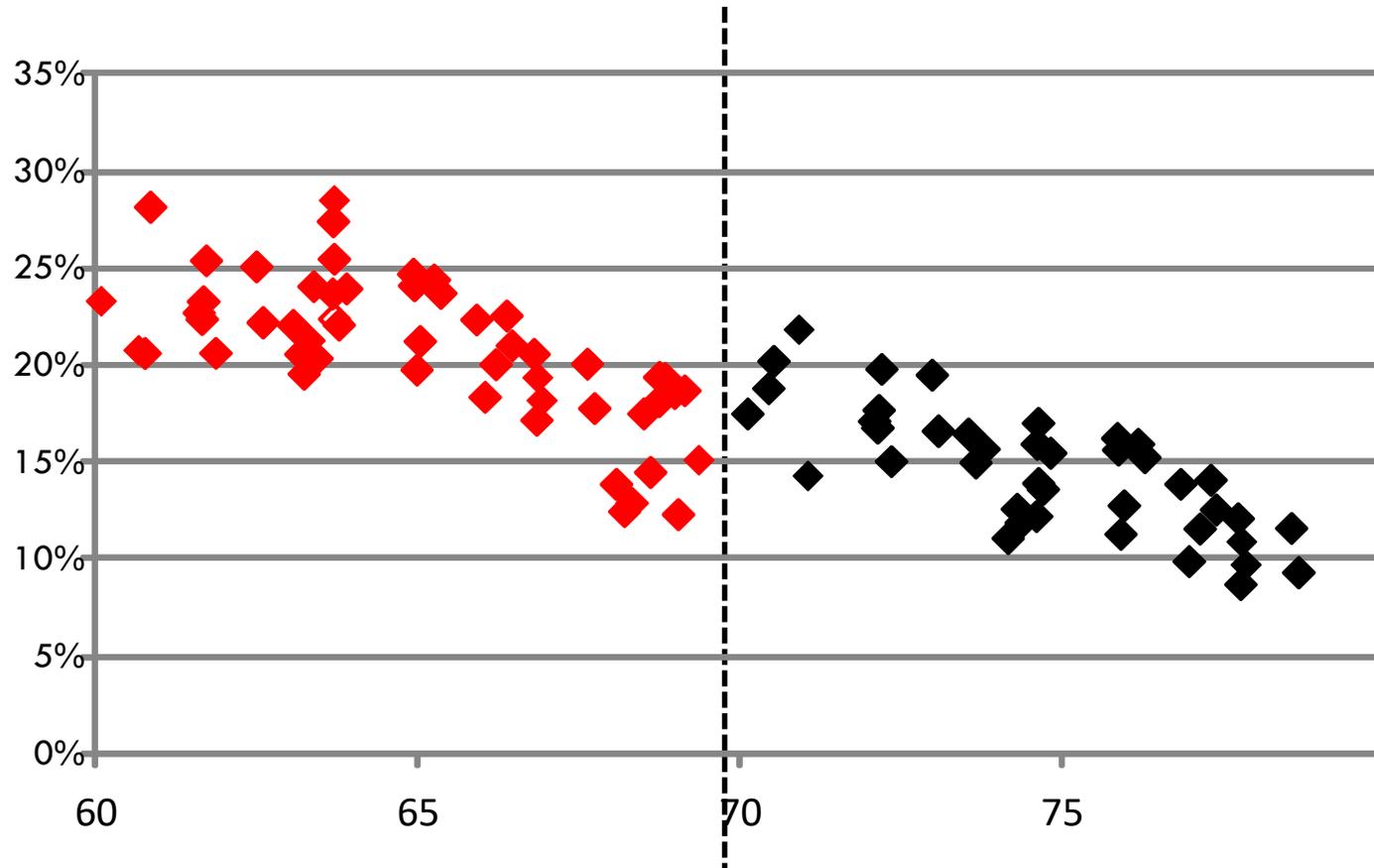
E con il trattamento?

44



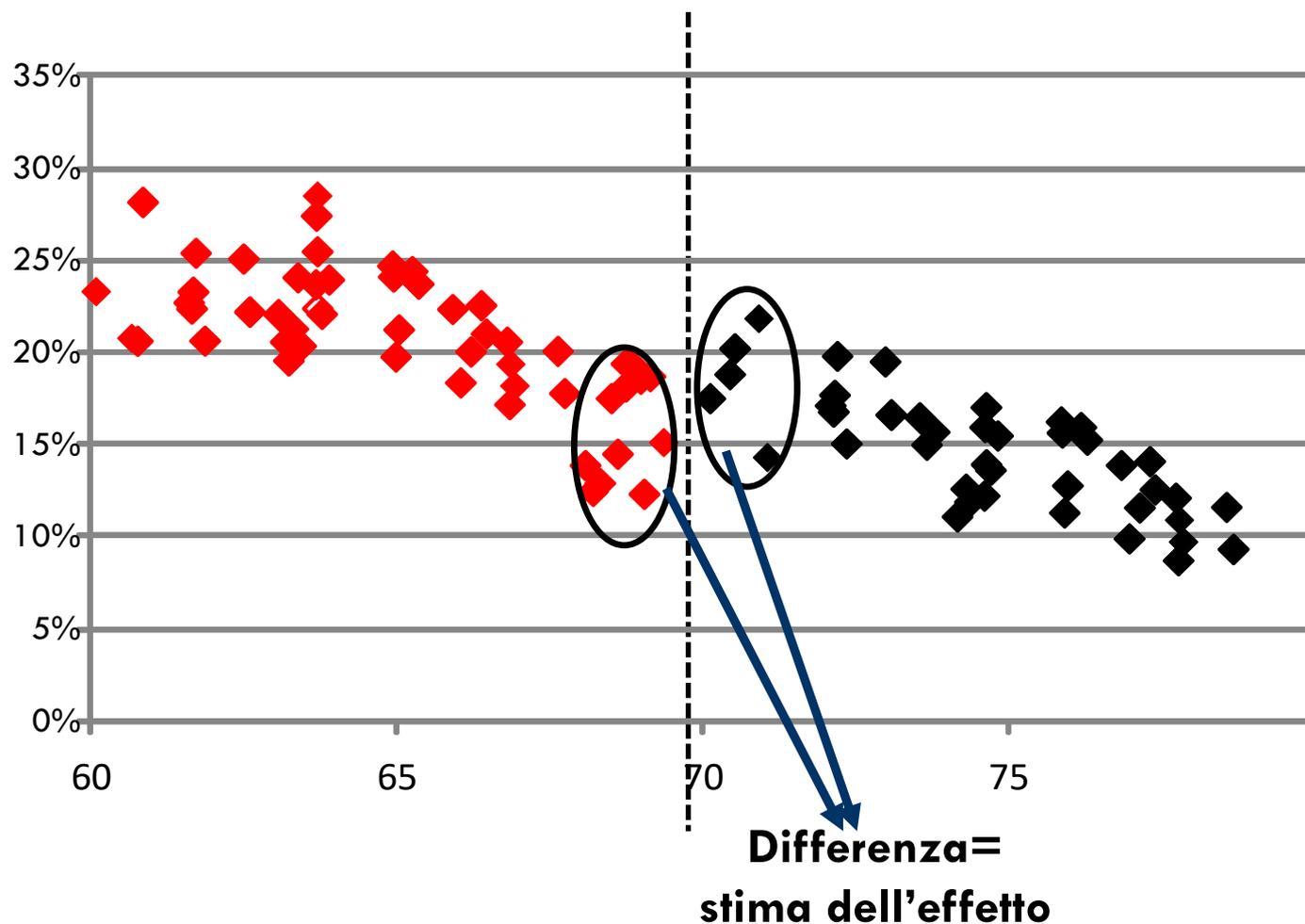
E con il trattamento?

45



Il confronto intorno al punto di discontinuità (o *Regression Discontinuity Design – RDD*)

46



Torniamo al Contratto di Ricollocazione della Sardegna

47

Una cosa che non abbiamo detto prima

La retribuzione del soggetto accreditato

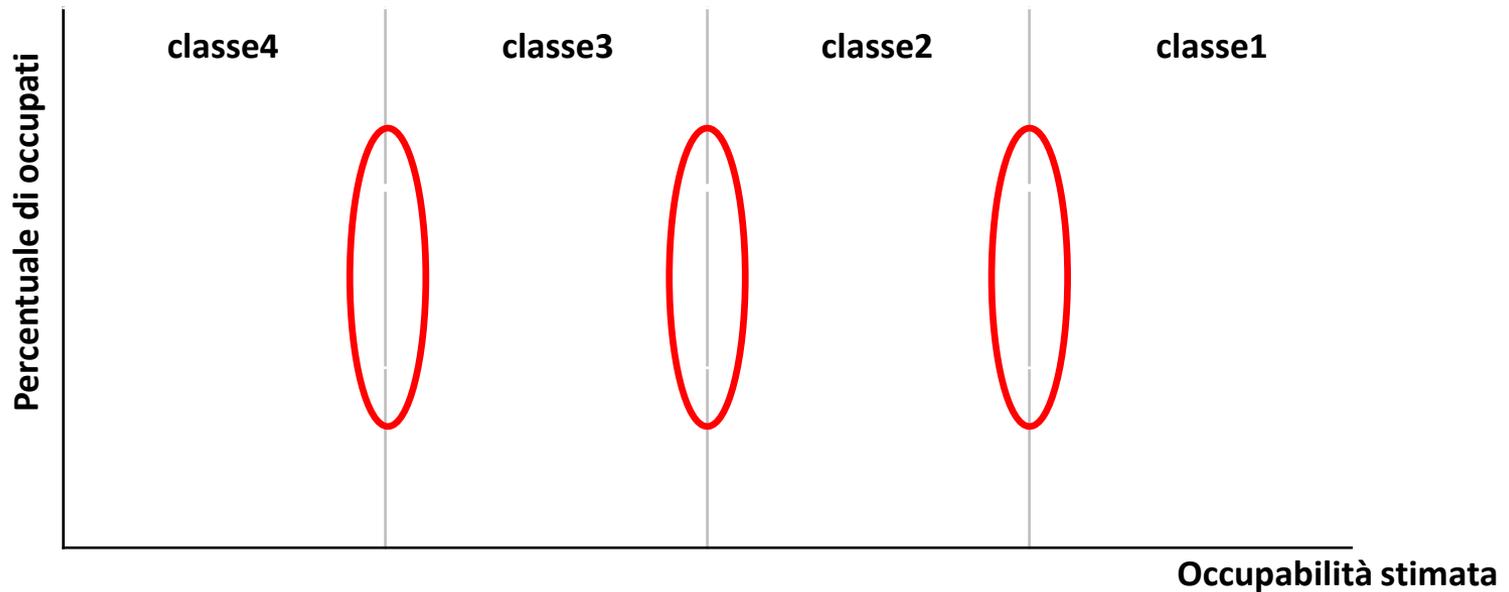
dipende dal grado di occupabilità del lavoratore

Come stimare l'efficacia dei compensi differenziali?

48

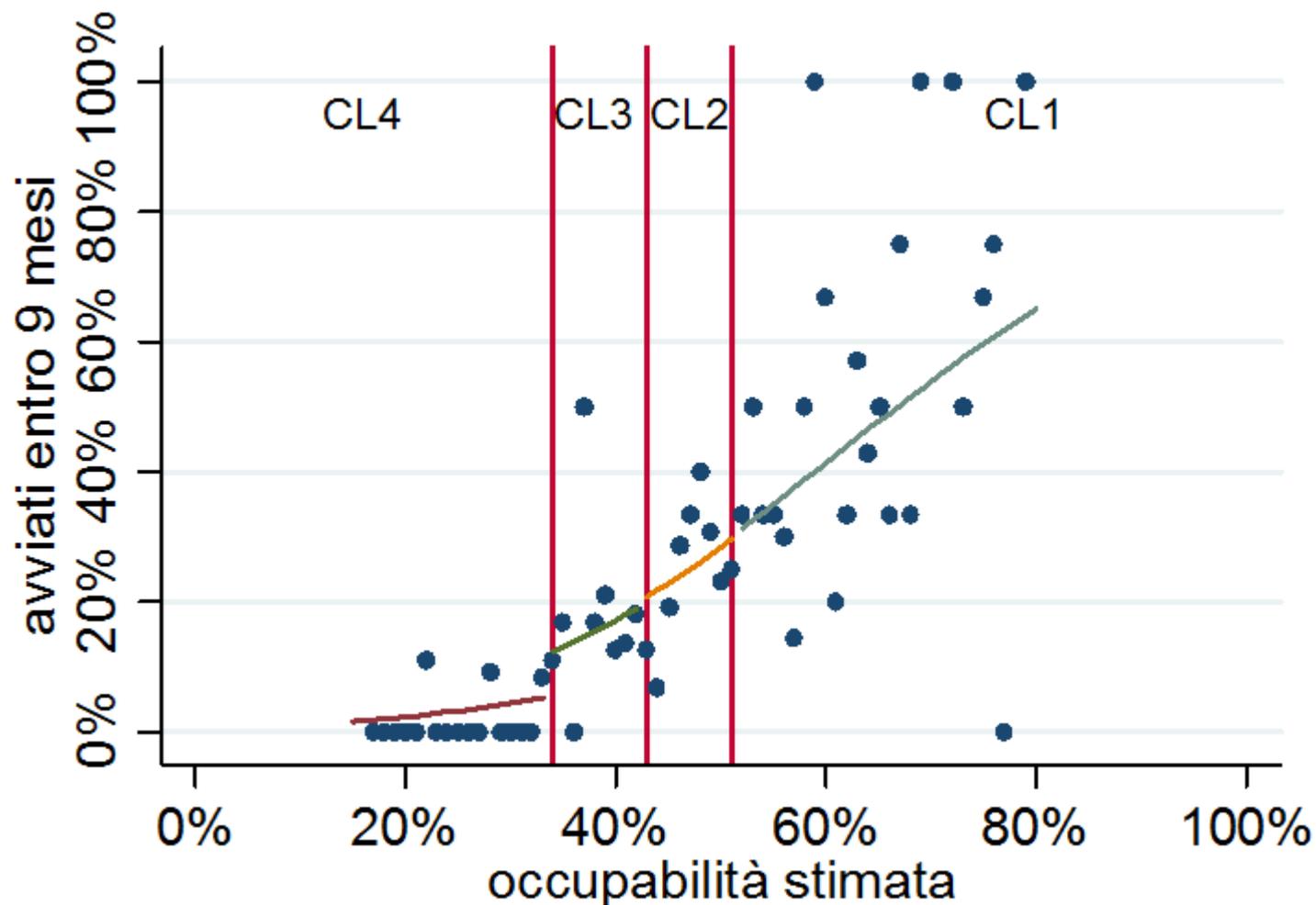
Selezione basata su «graduatoria»

Esempio di *sharp RDD*



Le stime degli effetti: i confronti intorno alla soglia

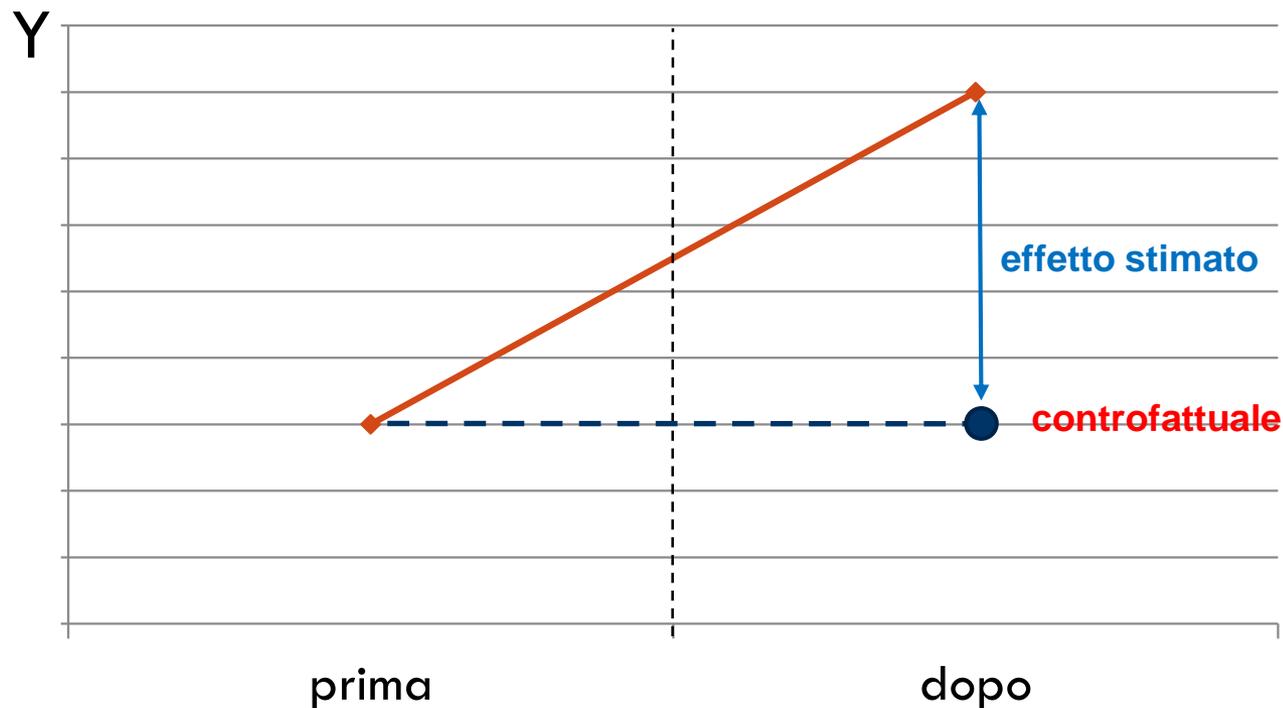
49



Pre-post design

50

L'idea generale è di sfruttare la storia passata di Y per fornire una stima del controfattuale



Basta un confronto prima-dopo?

51

Un confronto prima-dopo è una buona stima dell'effetto se possiamo affermare che in assenza della politica nulla sarebbe cambiato

Spesso è falso. Tutto può cambiare nel tempo

condizioni lavorative, redditi dei lavoratori
assunzioni, fatturati delle imprese

condizioni di salute

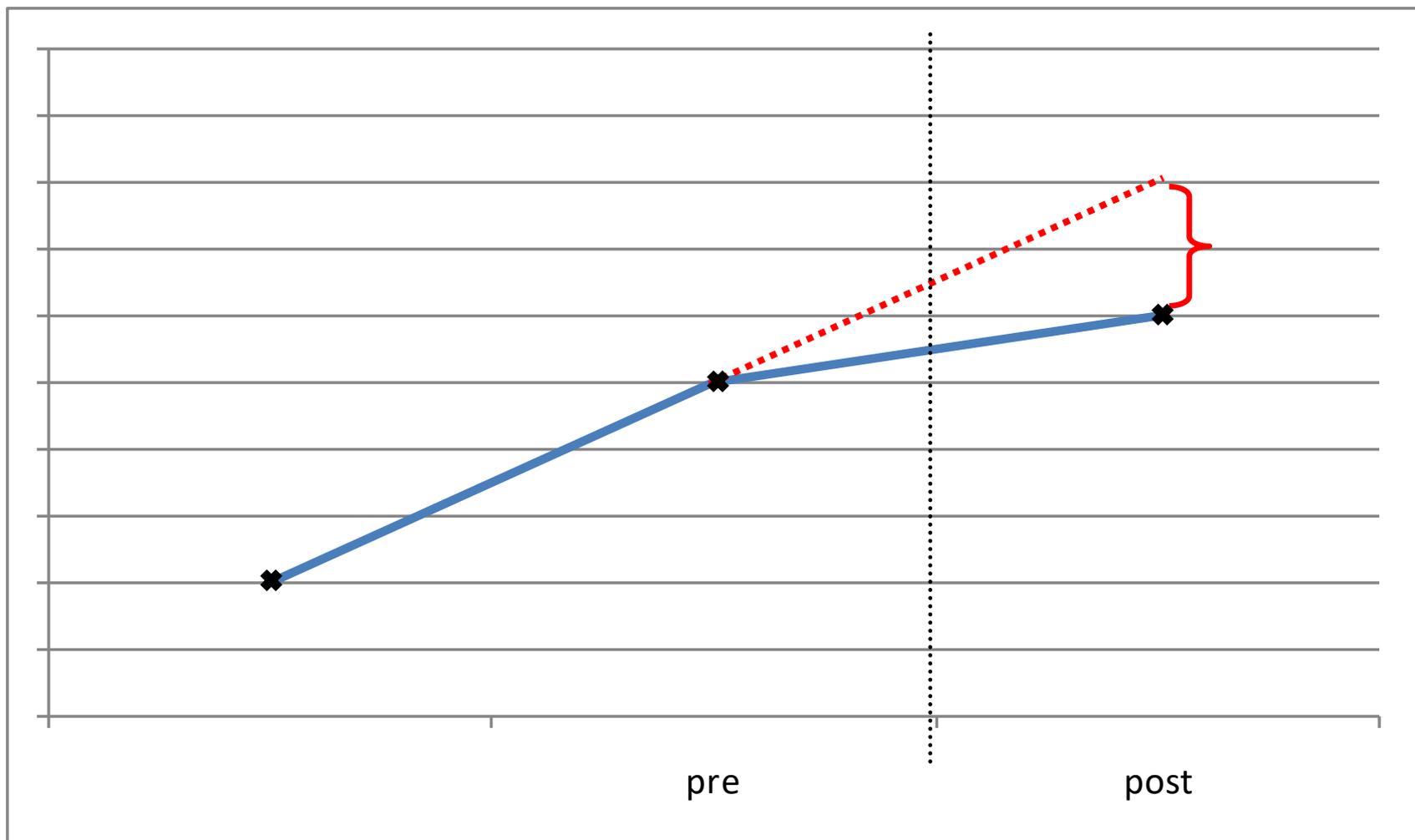
comportamenti

....

L'analisi del passato non può essere così semplicistica

Se abbiamo dati più ricchi....

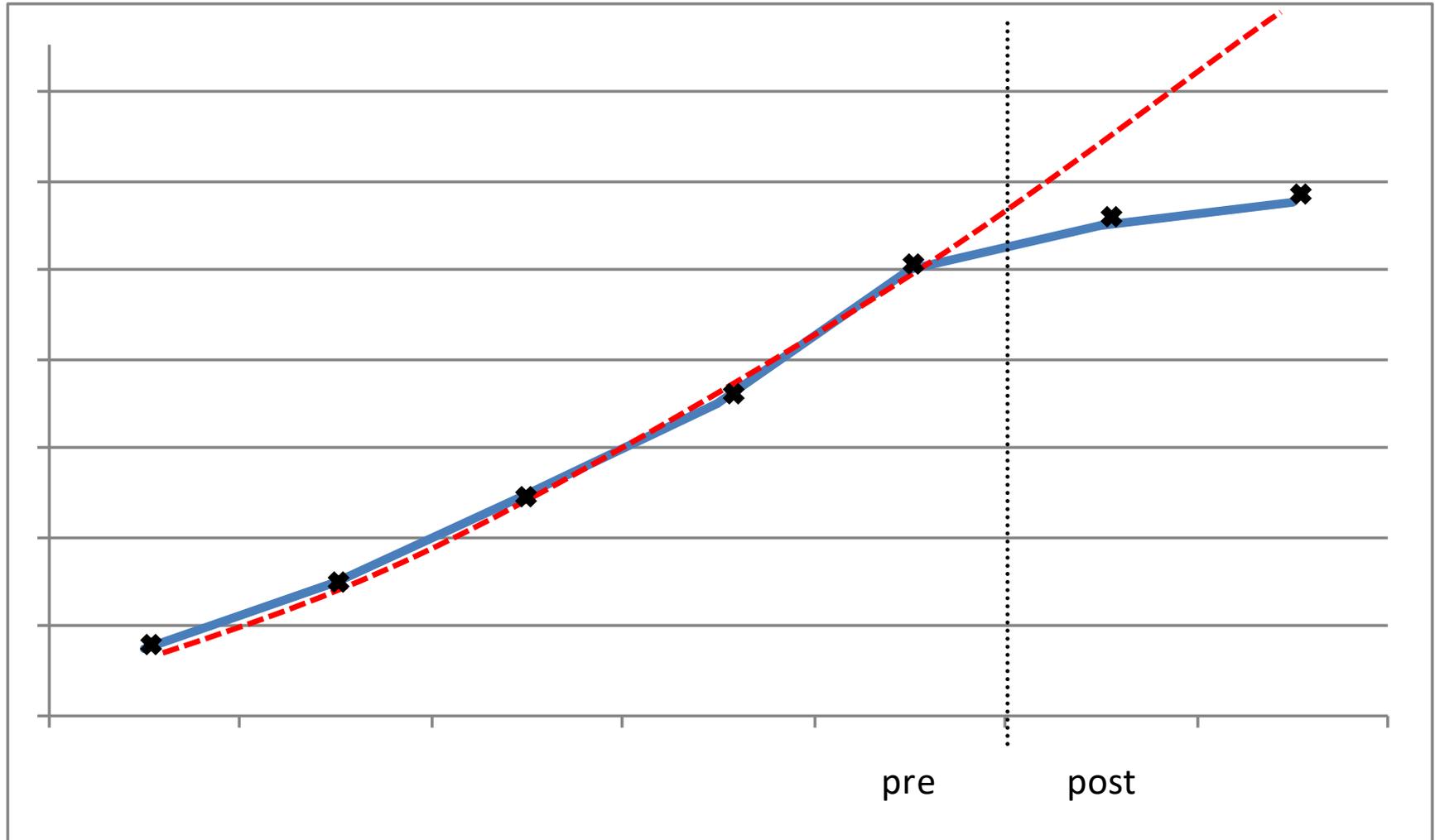
52



Con qualche osservazione in più...

Analisi delle serie storiche interrotte

53



Un (altro) esempio dalle PAL

54

**Incentivi per la stabilizzazione dei
lavoratori precari**

Provincia di Torino, 2007

Gli incentivi per la stabilizzazione

55

L'incentivo (di circa 4.500 euro*) è erogato alle imprese che assumono a tempo indeterminato un proprio lavoratore temporaneo (dipendente o atipico)

Imprese piemontesi, lavoratori torinesi

Gli incentivi per la stabilizzazione

56

Gli incentivi sono erogati in tre tranches:

- gennaio 2007 (4.700.000 euro)
- settembre 2007 (3.700.000 euro)
- gennaio 2008 (1.700.000 euro)

caratterizzate da lievi differenze nei criteri di erogazione:

- precedenza a certe categorie di lavoratori “deboli”
- retroattività (!!!!)

L'assalto agli incentivi

57

Gli incentivi vanno letteralmente a ruba

Ma è un segno di successo?

Valutazione degli effetti

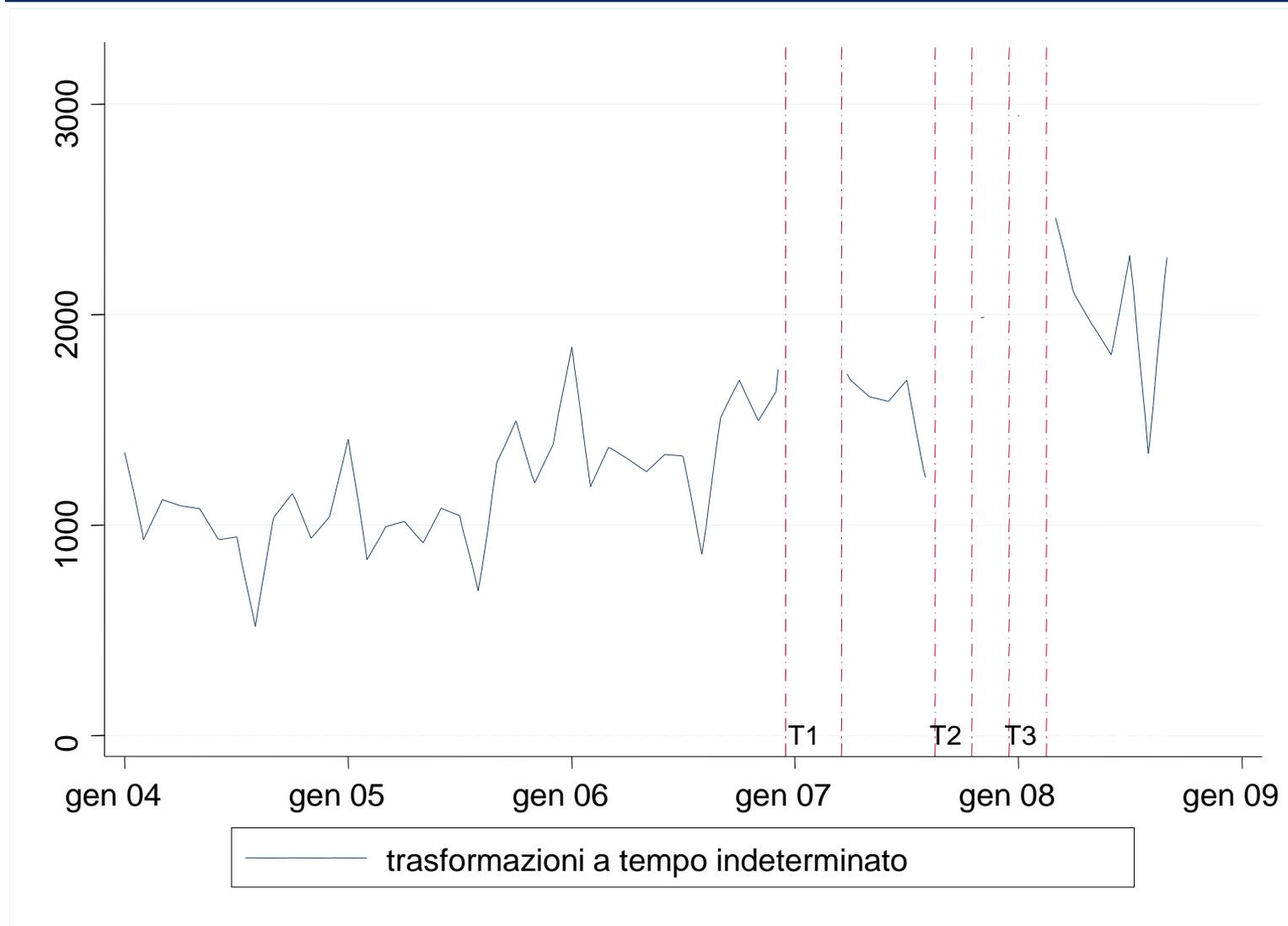
58

Poter ricevere un incentivo spinge le aziende a trasformare i contratti?

Con il SILP costruiamo la serie storica delle trasformazioni a TI nei periodi senza incentivi

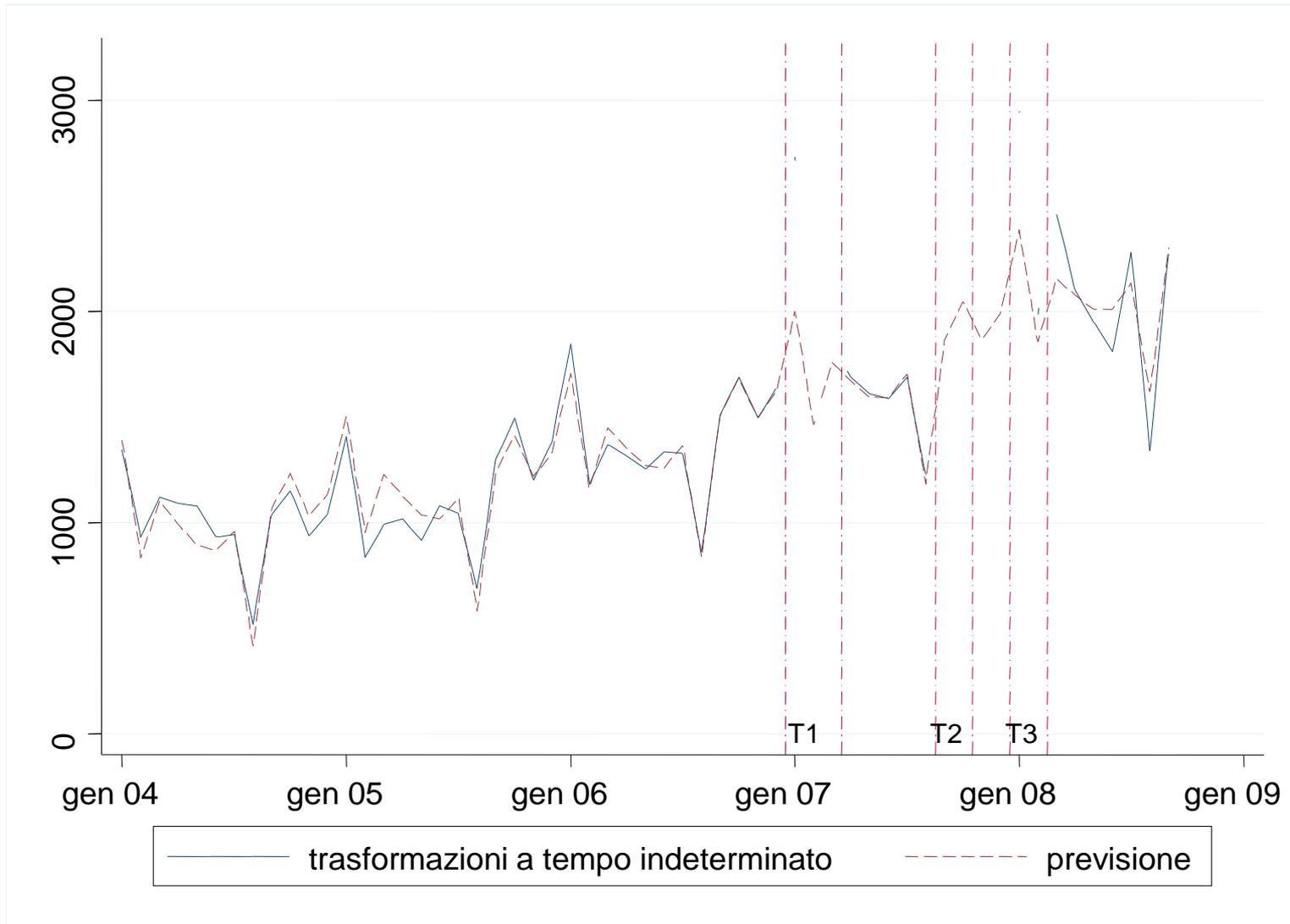
La serie storica mensile...

59



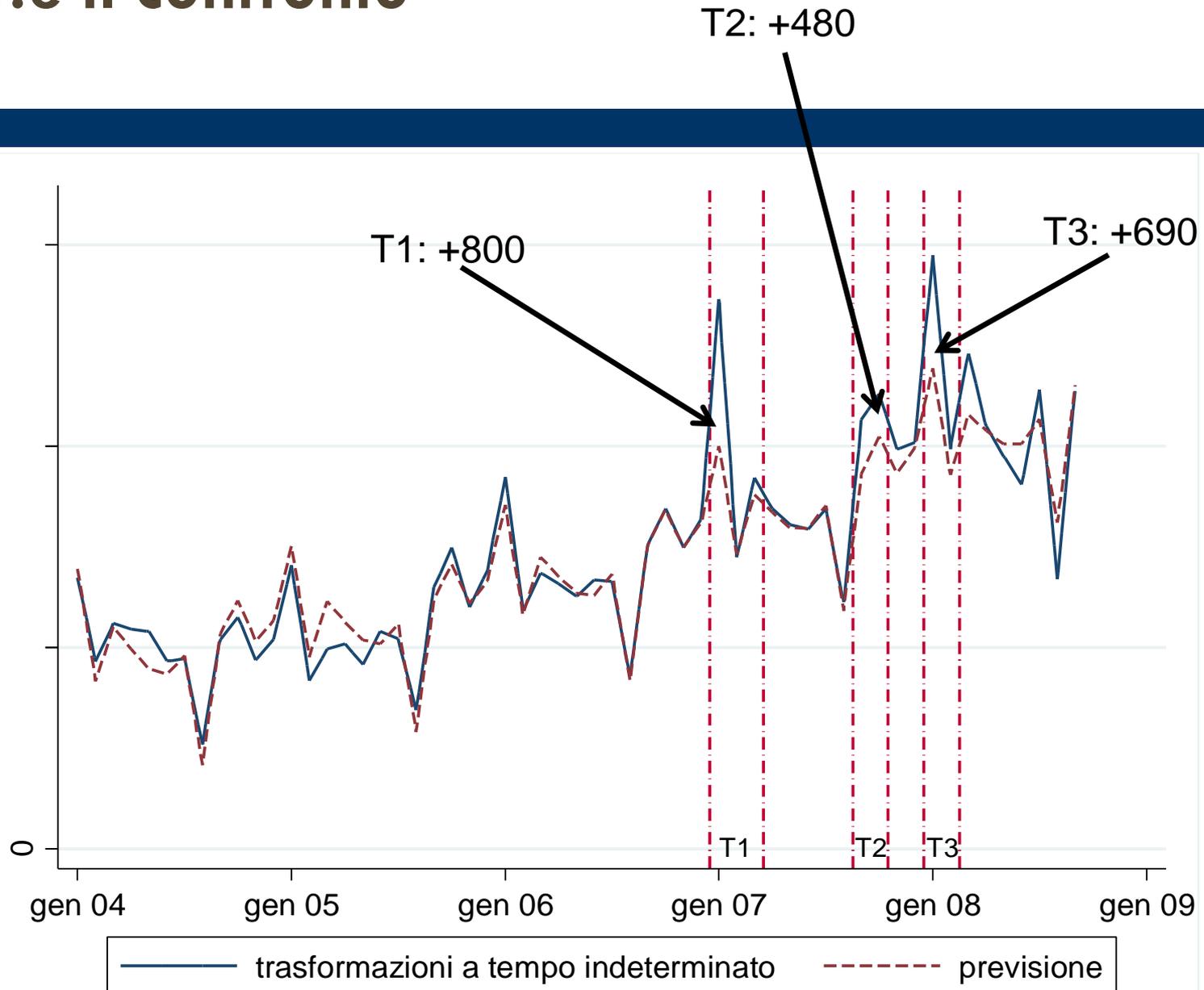
...la sua previsione...

60



...e il confronto

61



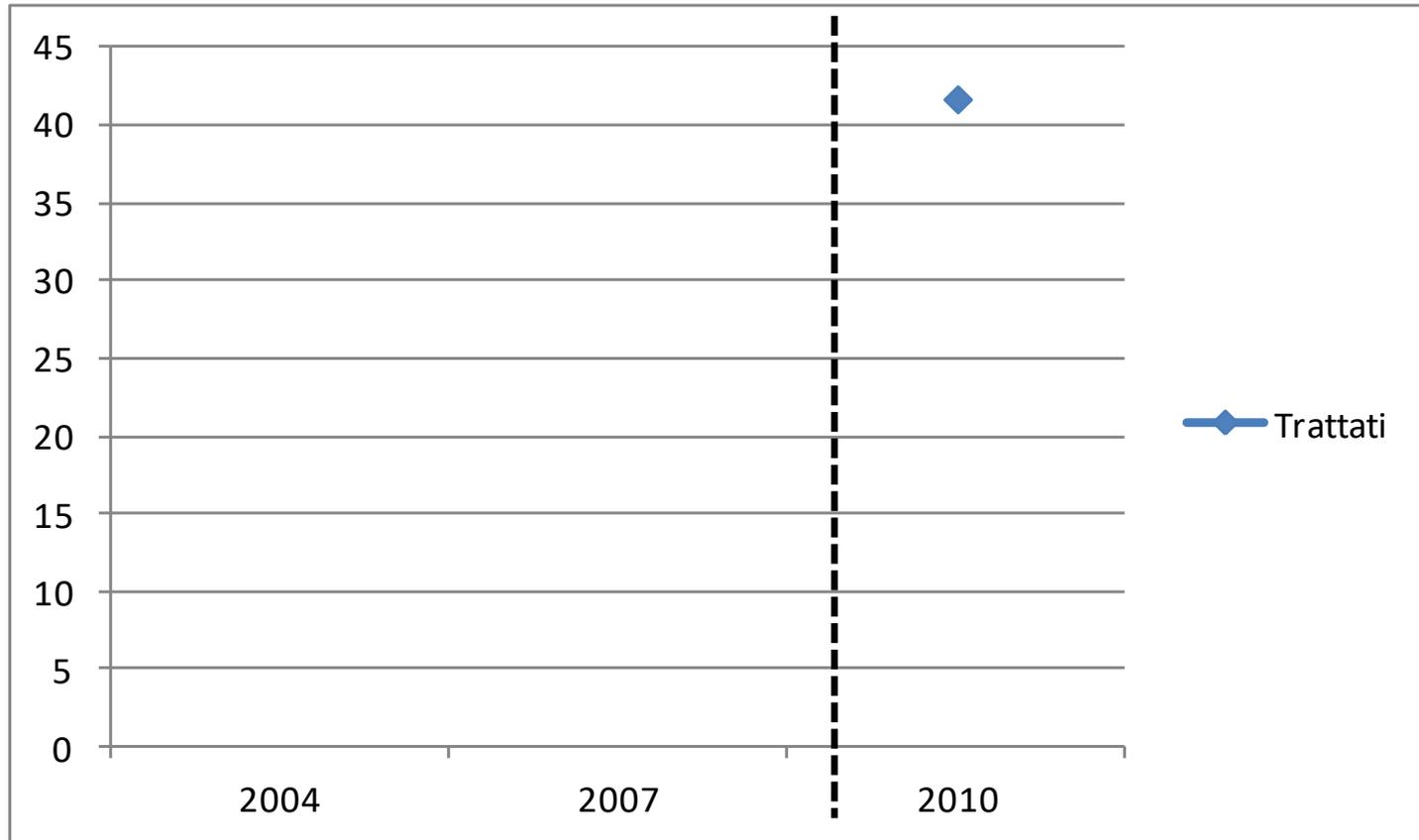
Combinare pre-post e gruppi di controllo?

Si può

Si chiama **differenza nelle differenze**

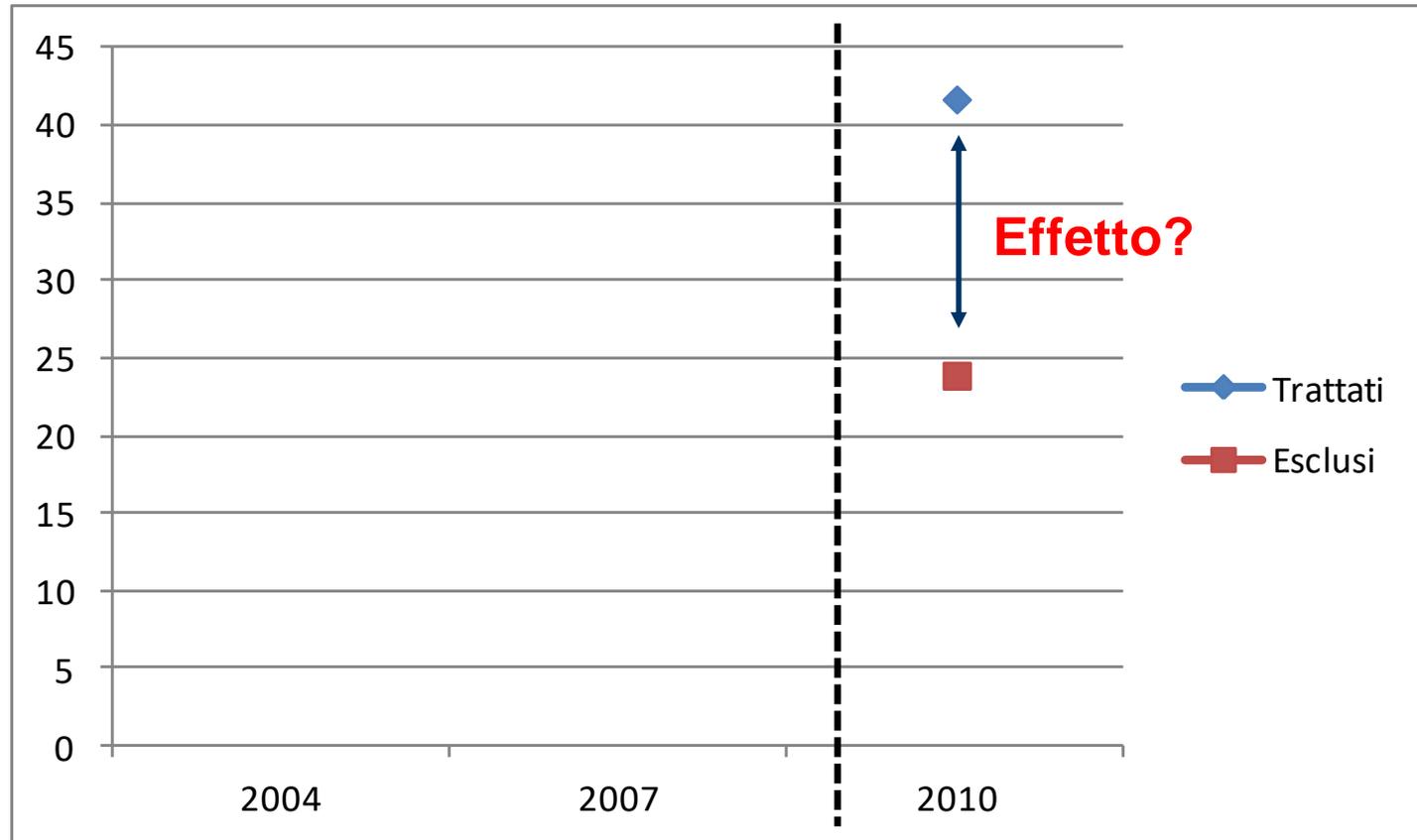
Esempio di DinD

63



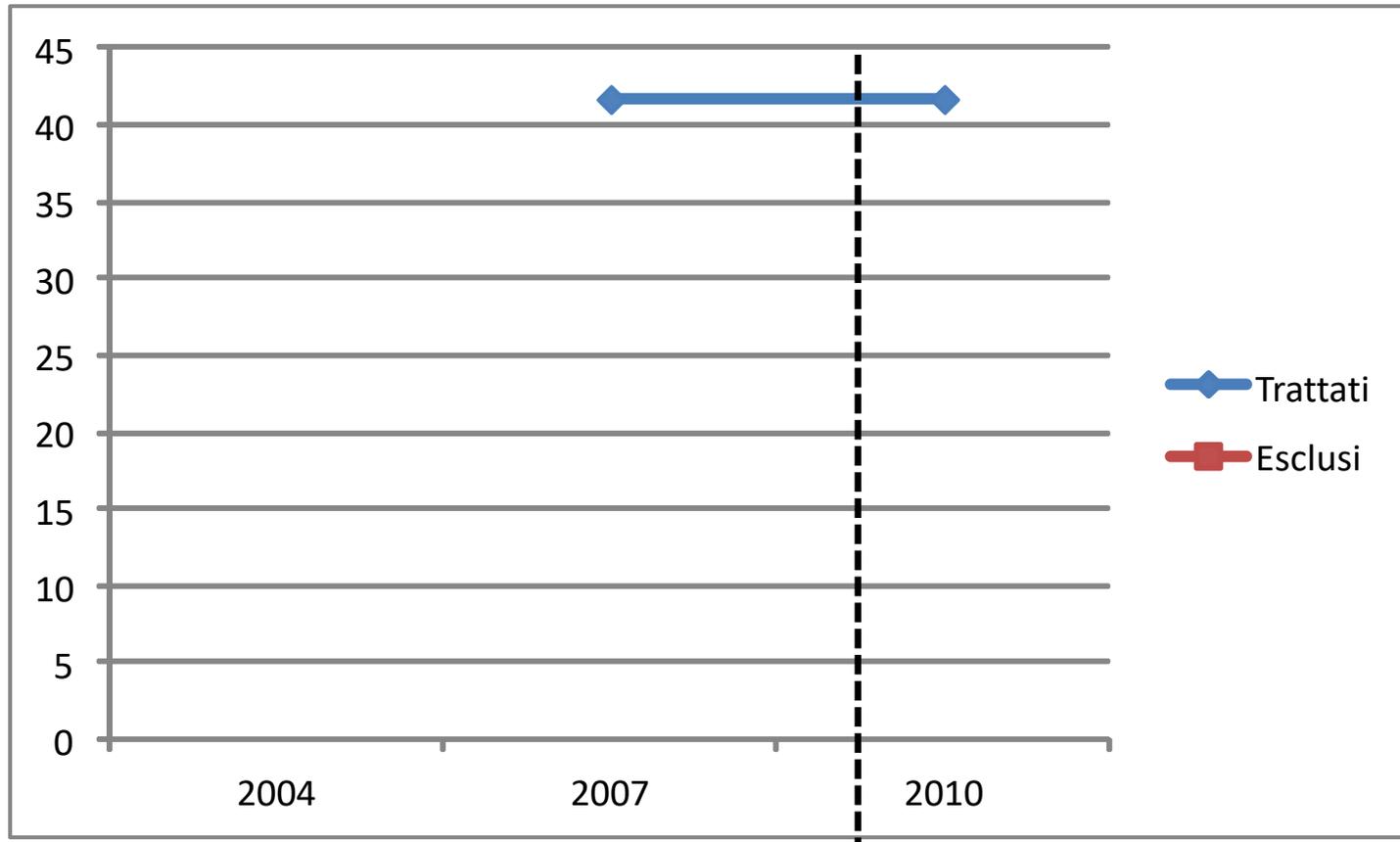
Esempio di DiD

64



Esempio di DinD

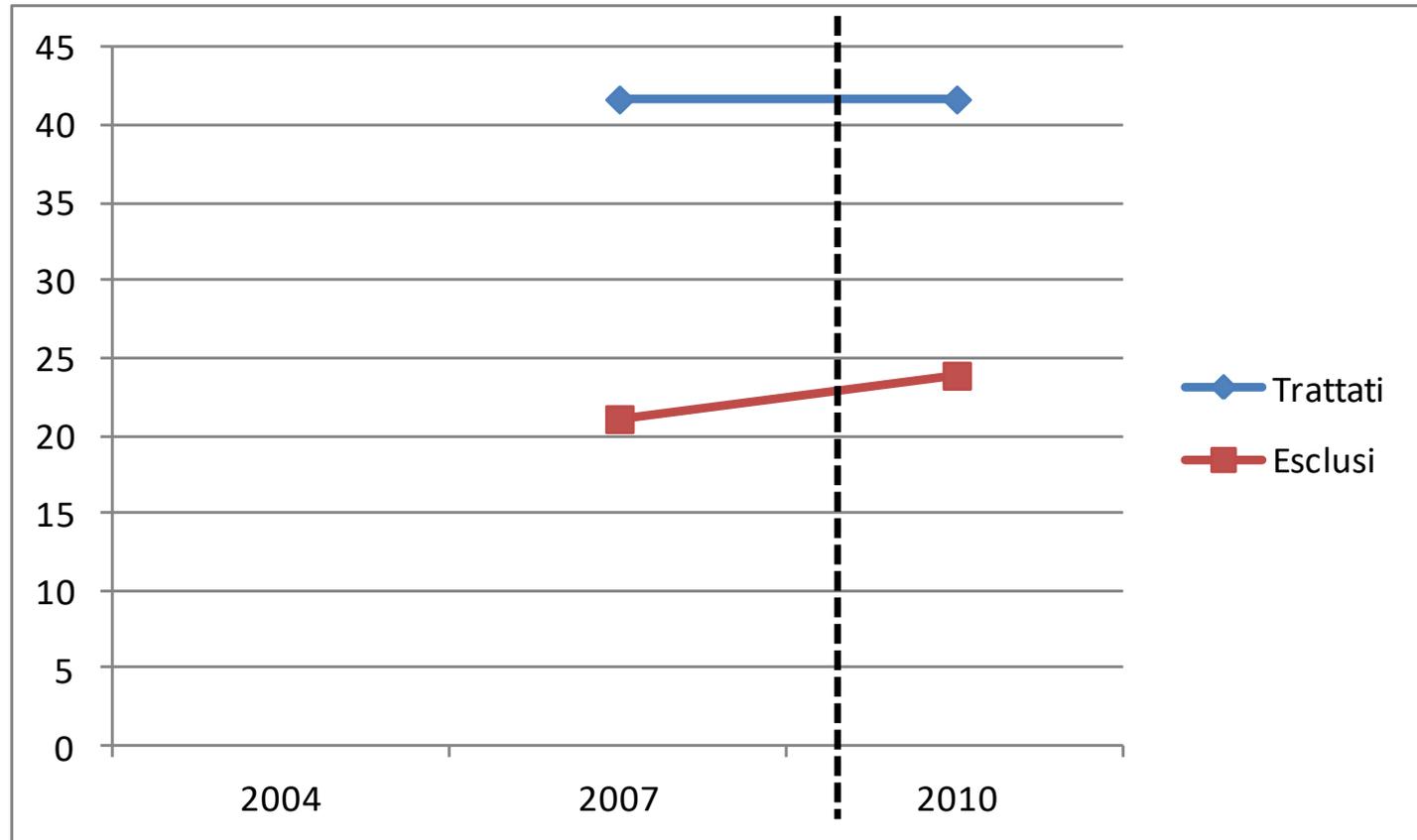
65



Esempio di DinD

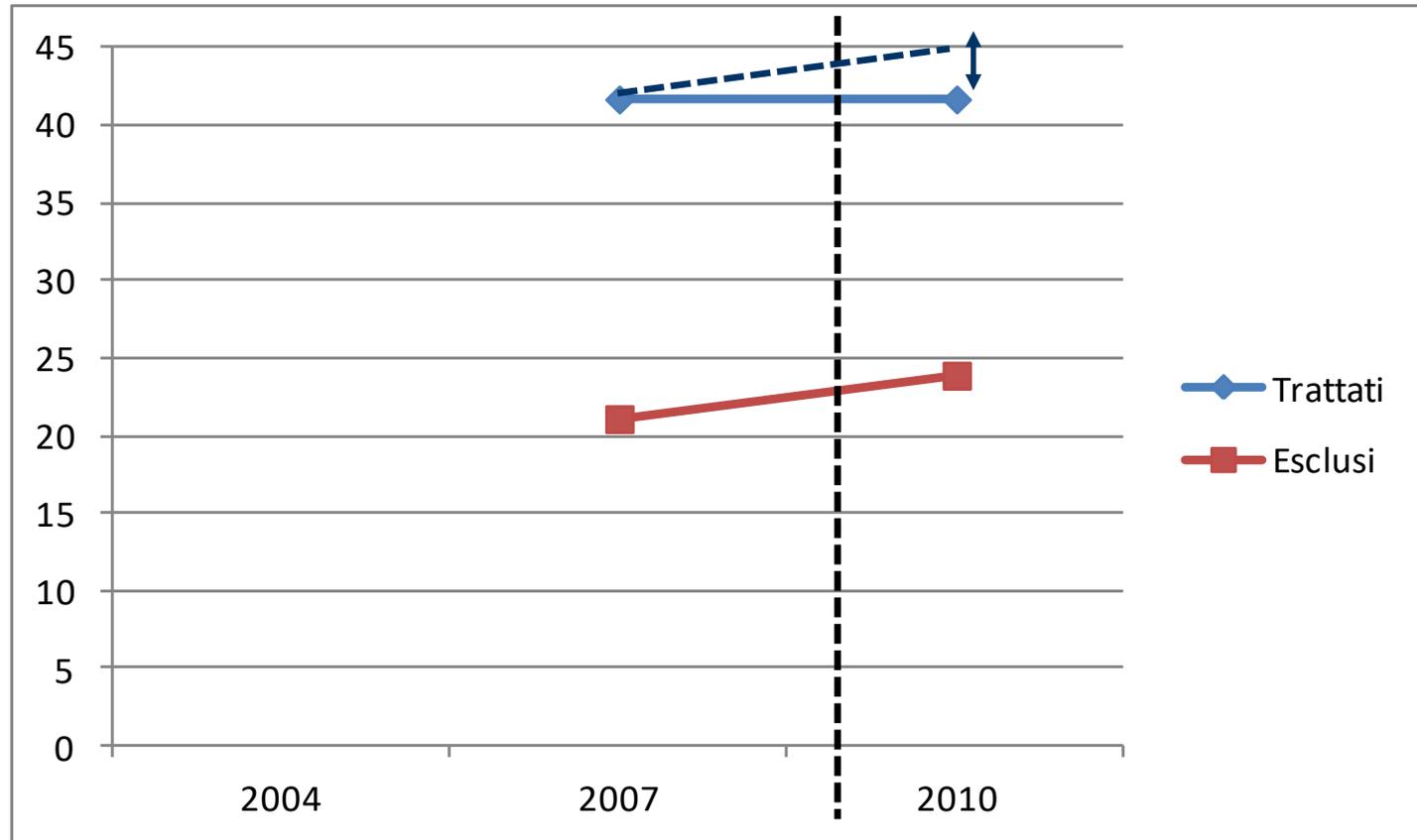
66

Osserviamo cosa succedeva prima



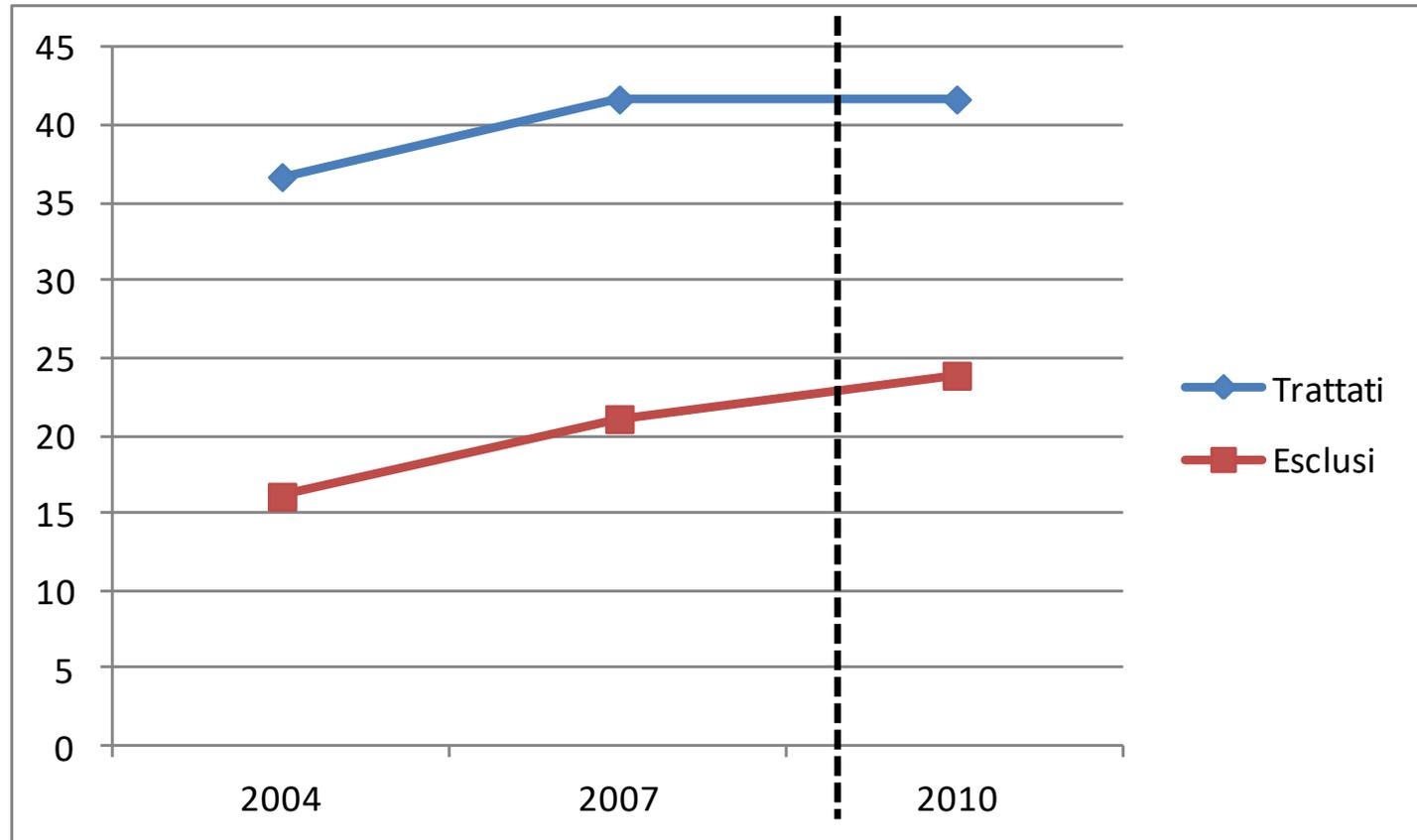
Una stima migliore?

67



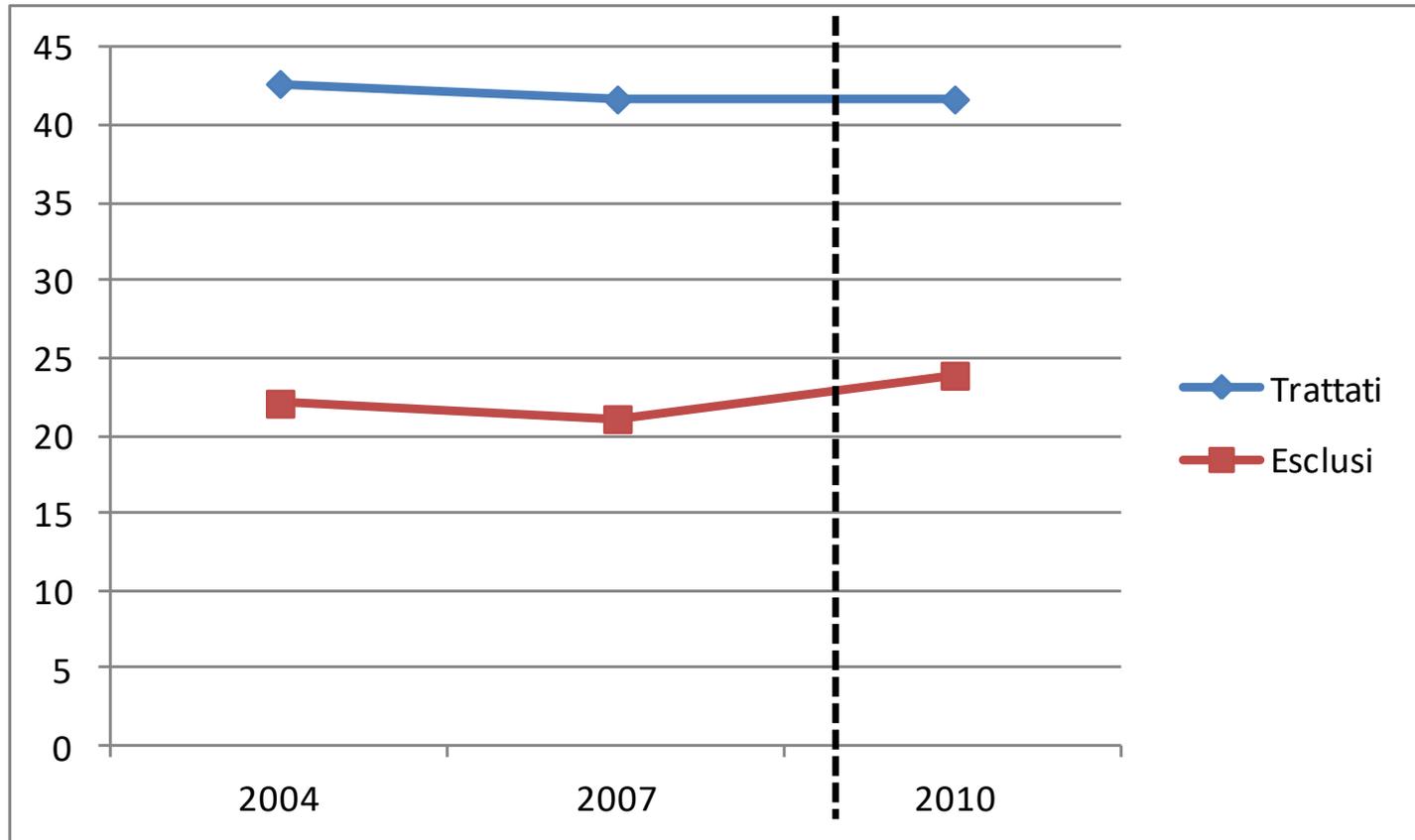
La stima è credibile solo se...

68



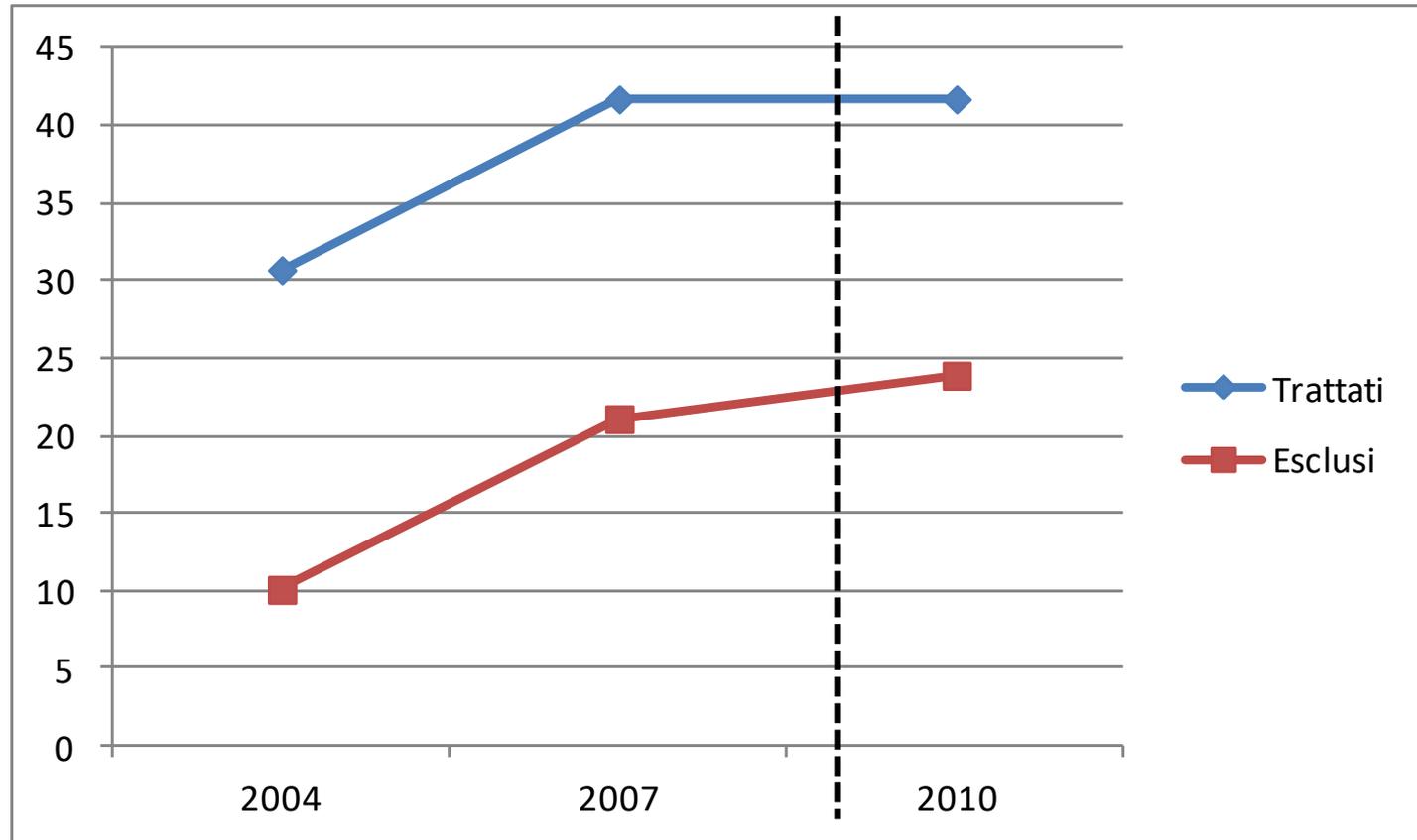
La stima è credibile solo se...

69

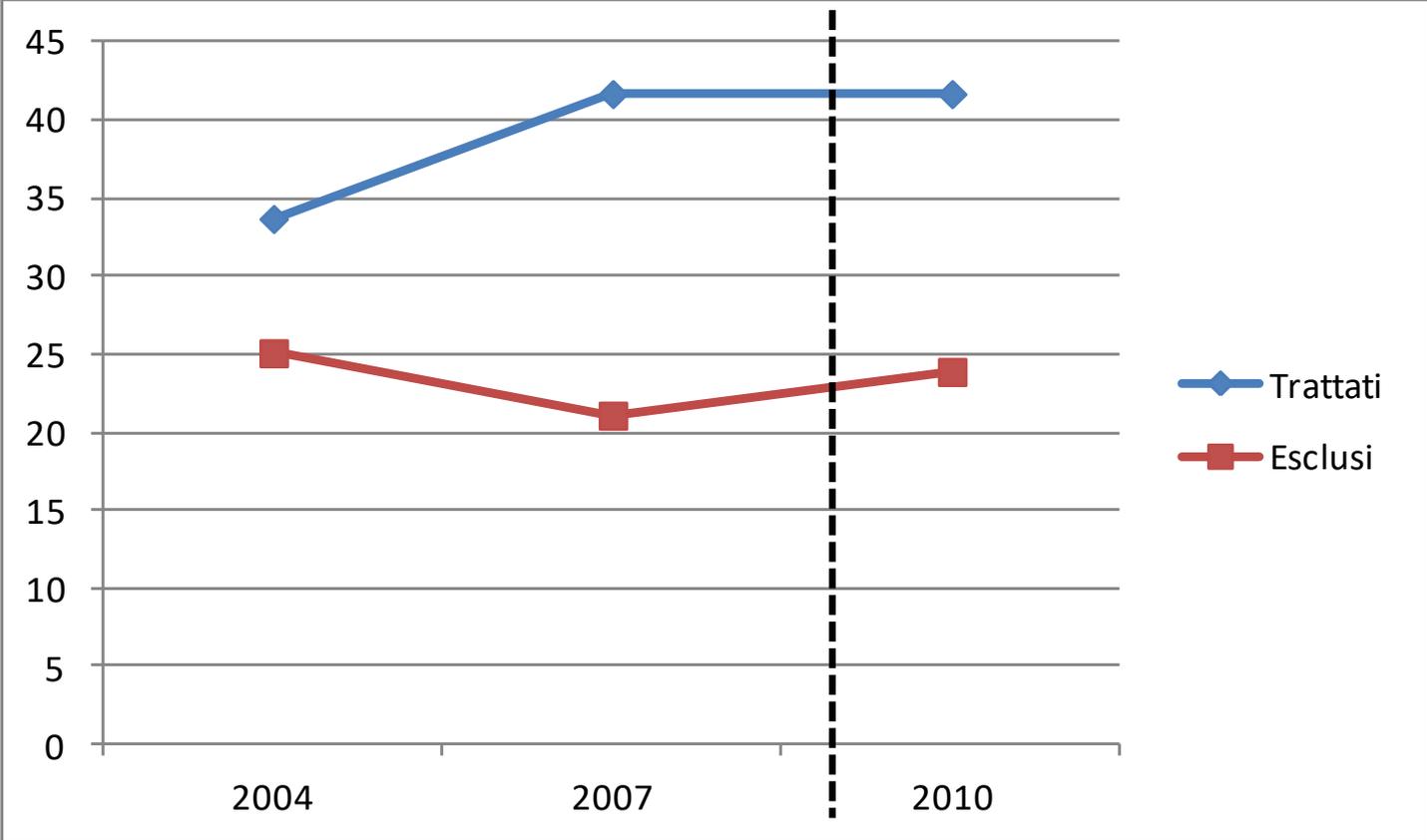


La stima è credibile solo se...

70



Ma non se...



Torniamo agli incentivi per la stabilizzazione

72

+800

+480

+690

= quasi 2000 trasformazioni in più del previsto

rispetto a quello che avremmo previsto di osservare in assenza degli incentivi

per ogni 5.000 euro investiti sono state incentivate le imprese ad assumere a TI un lavoratore

Il controfattuale è ricostruito in modo credibile?

73

I dati sul periodo in assenza di incentivi sono utilizzati per prevedere “cosa sarebbe successo in assenza di incentivi” negli altri periodi

La differenza tra “cosa è successo” e “cosa sarebbe successo” stima l'effetto

Correttamente?

Sì, se nei periodi T1, T2, T3 non è successo qualcos'altro a creare discontinuità

Alcuni “fattori di rischio”

74

A inizio 2007 si aprono le porte dell'UE alla Romania. Segue una massiccia regolarizzazione di rapporti di lavoro di cittadini rumeni.

(ma questo pesa solo sulle assunzioni, non sulle trasformazioni)

A inizio 2007 cambiano le regole sulle comunicazioni obbligatorie degli avviamenti. E delle trasformazioni.

Cosa ha comportato?

Possiamo usare un gruppo di controllo

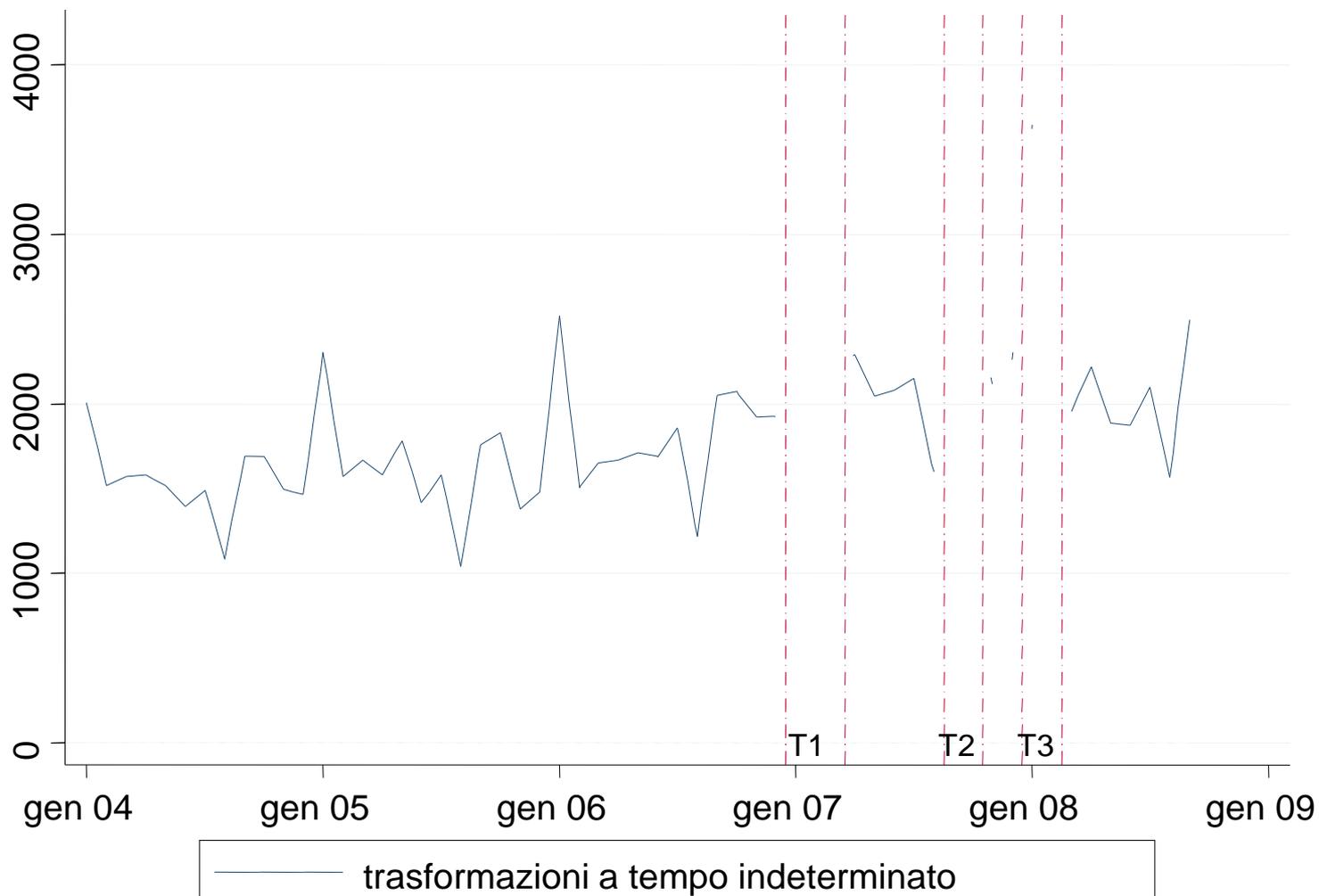
75

Composto da chi?

Nota: per usare una strategia DiD la differenza trattati controlli deve essere solo il trattamento, non altri eventi concomitanti

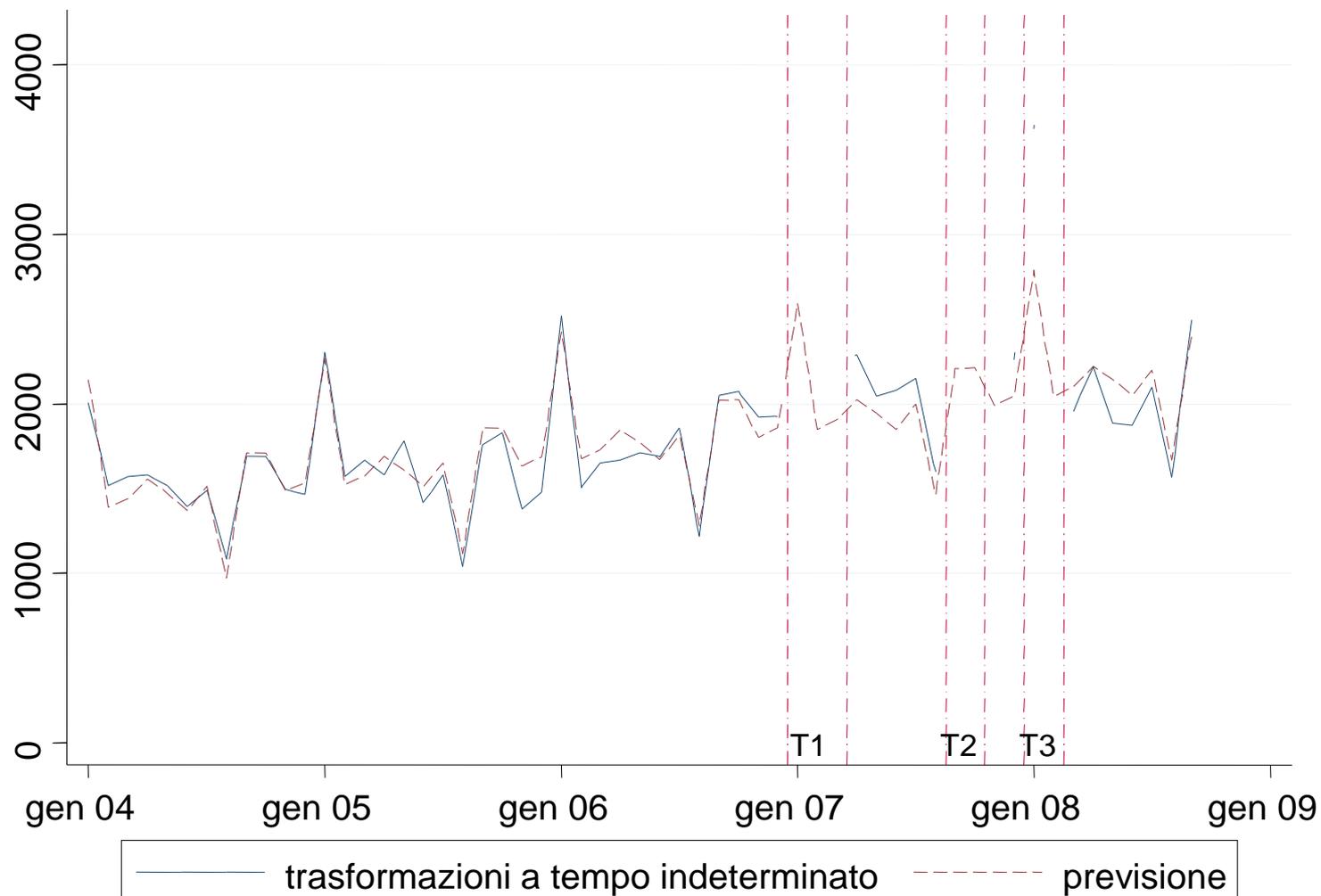
Le altre 7 province piemontesi (senza incentivi)

76



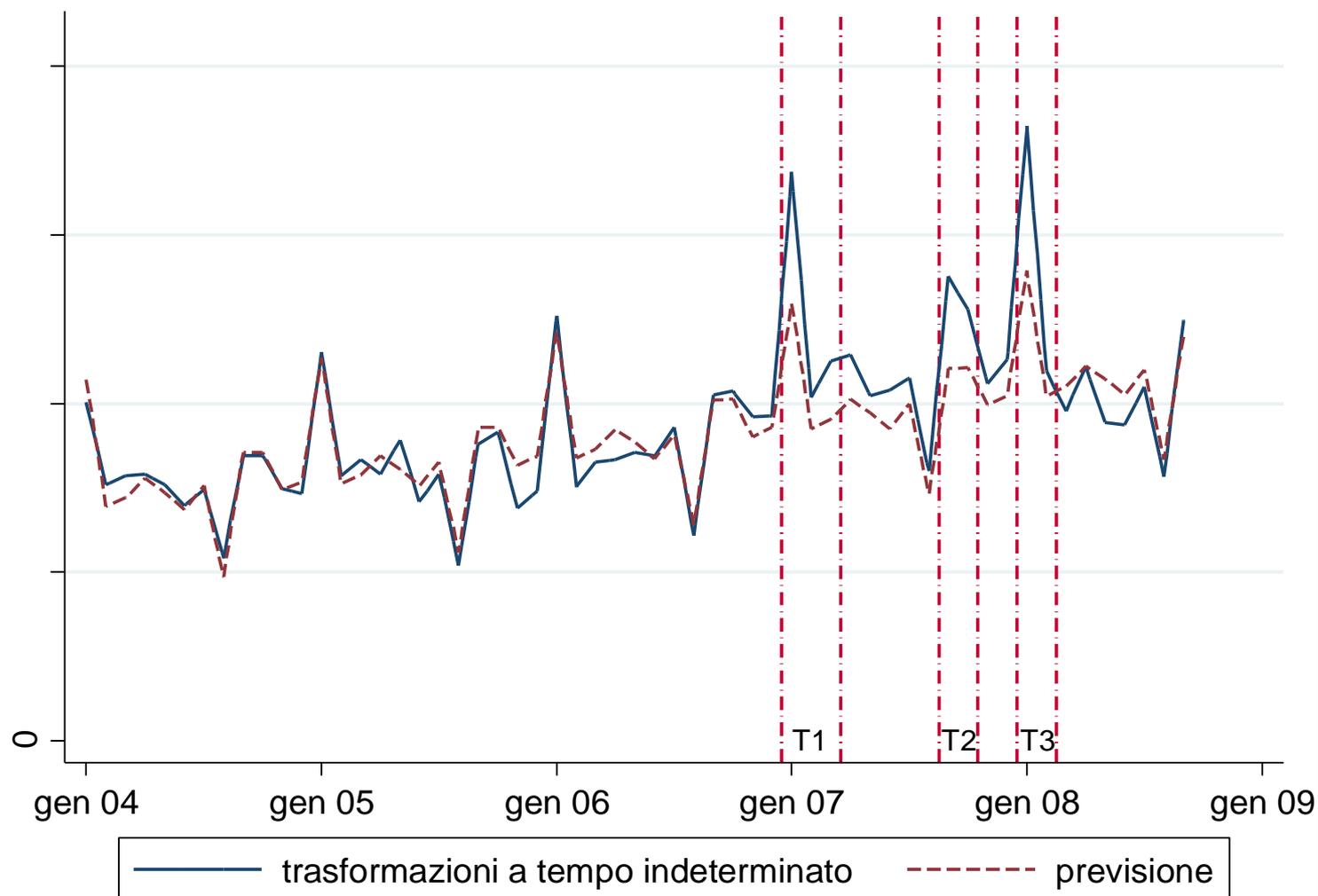
Le altre 7 province piemontesi (senza incentivi)

77



Le altre 7 province piemontesi (senza incentivi)

78



Nei periodi in cui a Torino si incentivano le stabilizzazioni, queste aumentano anche dove non si incentiva nulla

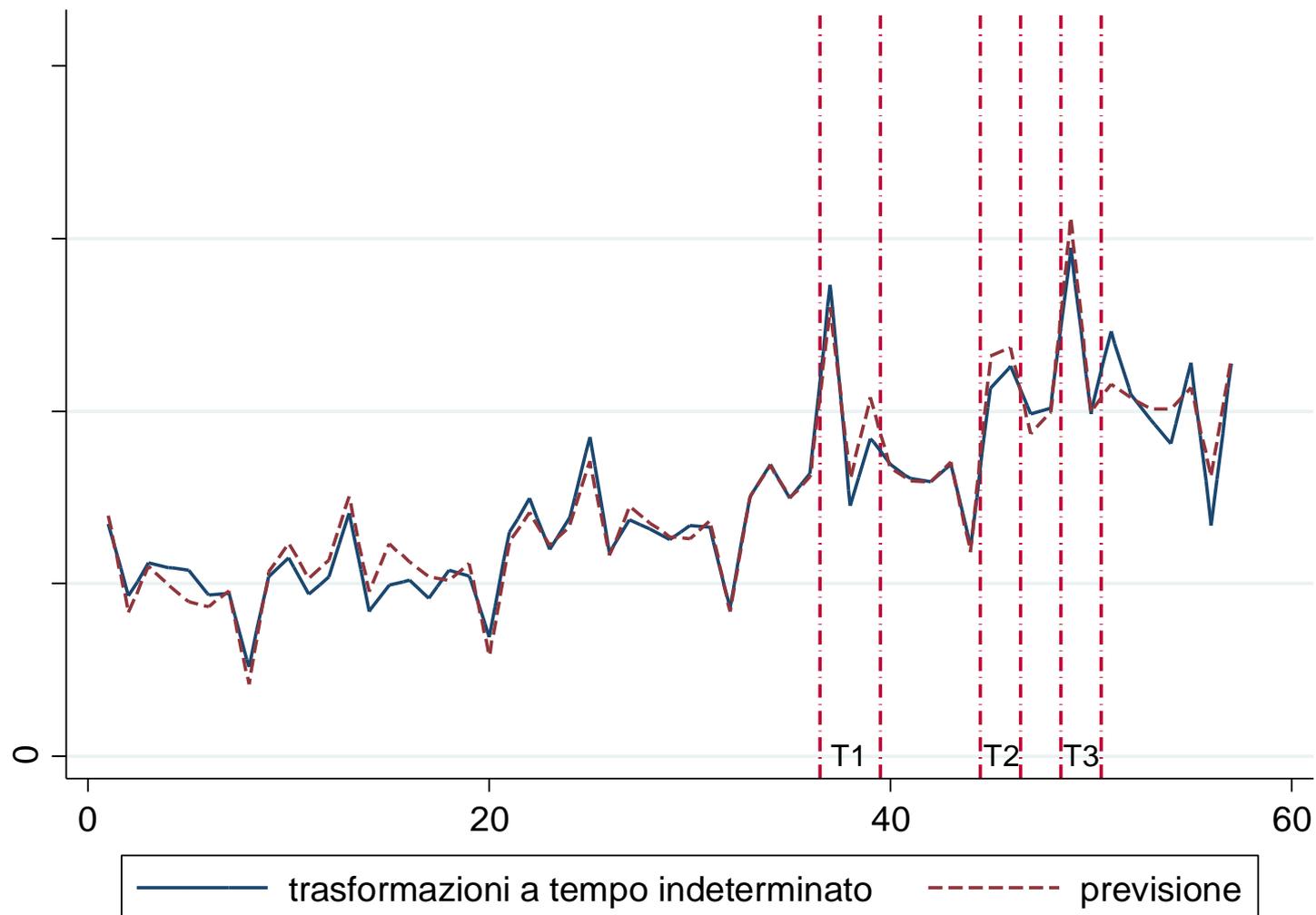
Quindi c'è l'effetto congiunto di “qualcos'altro”

Che cerchiamo di controllare con una strategia
Difference in Differences

La variazione incontrollata nelle 7 province di controllo è applicata a Torino. Le differenze residue saranno effetto degli incentivi

La stima *DinD* per la provincia di Torino

80



GRAZIE

LMOCOSTABELLA@ASVAPP.ORG