IRPET Istituto Regionale Programmazione Economica della Toscana

Evaluating public supports to the investment activities of business firms: A meta-regression analysis of Italian studies

Marco Mariani

marco.mariani@irpet.it

Getting into the debate

- A recent (recurrent)debate about the effectiveness of industrial policies:
  - Is public intervention useful or is it a complete waste of taxpayers' money?
  - What types of policy instruments work better?

- Evaluation can provide some of the needed answers. However, in order to be reliable, evaluation must be performed with the appropriate statistical tools (so-called *econometrics of program evaluation*)

Outline

- What is quantitative programme evaluation?
  - ✓ “treatments”, outcomes, quantities of interest
  - ✓ Experiments vs. real-world observational settings
  - ✓ conditions under which the causal effect of a treatment can be identified and estimated in observational, real-world, settings

- A meta-analysis of Italian evaluation studies:
  - ✓ what is a meta-analysis?
  - ✓ what is the probability of success of Italian enterprise and innovation policies?
  - ✓ what type of policy works better?

Quantitative program evaluation

After the programme has been implemented…

did the program change the participants' "behaviour" in the desired way?

Three examples:

	Firm R&D subsidy (S)	Training (T) for the unemployed	University grants (G) for less rich students
Treatment	1 if subsidised, 0 if not	1 if trained, 0 if not	1 if delivered, 0 if not
Outcome Y	R&D investment one year later	Occupational status one year later	Dropout status at the end of 1° year
Expected result	Higher average investment if S=1	Higher probability of being employed if T=1	Lower probability of dropout if G=1

Quantities of interest

T : treatment status = 1 if treated; =0 if untreated

Y: outcome of interest

Treatment effect for unit *i*: $Y_i(T=1) - Y_i(T=0)$

Problem:

if *i* receives treatment we only know $Y_i(1)$ but not $Y_i(0)$

if *i* does not receive treatment we only know $Y_i(0)$ but not $Y_i(1)$

Therefore, attention shifts on estimable average quantities, such as

Average treatment effect: $E[Y(1)-Y(0)]$

Average treatment effect on the treated: $E[Y(1)-Y(0) \mid T=1]$

Simple comparisons

Let us focus, for example, on the Average treatment effect: E[Y(1)-Y(0)]
It is a very simple comparison between two averages

Firm id	Subsidy	R&D investment (Th. Euros) one year later	$E[Y_{+1}(.)]$
1	Yes	130	
2	Yes	100	
3	Yes	90	104
4	Yes	120	
5	Yes	80	
6	No	100	
7	No	80	
8	No	70	78
9	No	75	
10	No	65	

$E[Y_{+1}(1)-Y_{+1}(0)] =$
$= E[Y_{+1}(1)]-E[Y_{+1}(0)]=$
$= 104-78 = 26$

Is 26 the correct average treatment effect?

Under which conditions it is so?

The experimental ideal

In an experiment, firms are assigned to the subsidy at random

This means that subsidy assignment is by construction independent of firm's *i* potential outcomes $Y_i(1)$ and $Y_i(0)$

Random subsidy assignment guarantees that
E[Y(1)], the average outcome in the treatment group, and
E[Y(0)] , the average outcome in the control group
are not systematically different if the subsidy does not exist

Under this independence (exogeneity) condition
E[Y(1)-Y(0)] = 26
represents a correct estimate of the average subsidy causal effect on R&D investment

Although policy experiments are sometimes implemented on relatively small numbers in order to understand if a particular programme deserves mainstreaming …

Complications in observational settings

… real-world policies are more often evaluated after their implementation, to see if they worked

In this situation we can only observe ex-post which firms received the subsidy, as a result of
i) their choice of applying for the programme
ii) the agency decision to subsidise some of the R&D project that firms have submitted

Since it is based on firm and agency choices, subsidy assignment cannot be assumed as independent of firm's *i* potential outcomes $Y_i(1)$ and $Y_i(0)$. For example:

- if already innovative firms apply for and receive the subsidy and non-innovative ones do not, then E[Y(1)] is probably higher than E[Y(0)] also if the subsidy does not exist

This "selection" problem makes the two groups incomparable and therefore E[Y(1)-Y(0)] = 26 is no longer a correct average treatment effect, in that it is biased by differences in the level of Y that would have been there anyway!

Revisiting the previous example

Firm id	Subs.	Investment one year earlier	$E[Y_{-1}(.)]$	Investment one year later	$E[Y_{+1}(.)]$
1	Yes	100		130	
2	Yes	95		100	
3	Yes	90	92	90	104
4	Yes	105		120	
5	Yes	70		80	
6	No	95		100	
7	No	70		80	
8	No	60	68	70	78
9	No	55		75	
10	No	60		65	

Figures suggest that selection bias is at least 92 - 68 = 24

If we assume that this difference is sufficient to account for selection bias, then the average treatment effect is much smaller now: 26 – 24 = 2

Possible solutions in observational settings

If treatment assignment is not exogenous by construction, treatment effects can be identified and estimated at the price of some assumptions:

- ❑ parallelism: if differences prior to treatment are assumed to be constant over time, as in the previous example, then 26 – 24 = 2 → Difference-in-difference approach (**DID**)
- ❑ selection on observables: we might also assume that, for firms having the same observable characteristics, treatment assignment is as good as random, and compare the E[Y(1)] to an E[Y(0) | twins] → **matching**, also in combination with DID
- ❑ suppose the subsidy is assigned if the merit score compiled by the agency after assessing firm proposals is at least equal to a certain threshold value that remains out of the control of firms. Projects just above this threshold can be compared to projects just below the threshold that show very similar merit but are unsubsidised, under the very mild assumption that Y would have been a smooth function of merit, in the small region around the threshold, if the subsidy was not there → Regression Discontinuty Design (**RDD**)
- ❑ other approaches rely on specific assumptions on treatment assignment (e.g. instrumental variables, selection models, etc.)

A review of ‘counterfactual’ studies

The approaches recalled so far are intended to reconstruct ‘counterfactual’ situations in an explicit and convincing way

Once we select for consideration these studies only, we must decide how to pool their results and carry on a review. Two approaches are possible:

Narrative	Systematic
Depend on authors’ inclination (bias?)	Scientific approach to a review article
Author gets to pick any criteria	Criteria determined at outset
	Comprehensive search for relevant articles
Methods not usually specified	Explicit methods of appraisal and synthesis
Vote count or narrative summary	Meta-analysis may be used to combine data

What is a systematic review?

A comprehensive review of the evidence on a clearly formulated question that uses systematic and explicit methods to identify, select and critically appraise relevant primary research, and to extract and analyse data from the studies that are included in the review

Advantages over narrative reviews:

- Reduces subjective bias
- Replicability
- Helps resolve controversy between conflicting studies
- Identifies gaps in current research
- Provides reliable **basis** for decision making

Once one has collected all studies and categorised the information they contain in an appropriate way, the choice is how to produce a meaningful summary of this information

Systematic reviews & Meta-analysis

- Systematic review is the entire process of collecting, reviewing and presenting all available evidence

- Meta-analysis is the statistical technique involved in extracting and combining data to produce a summary result of the systematic review

Meta-analysis

- Allows the reviewer to quantitatively combine and analyse the results from multiple studies
- Thus a MA becomes a research study on research studies, hence the term "meta".
- MA collects the empirical results from multiple studies and draw conclusions about the “overall” effect across studies no matter what the original study conclusions were.
- In other words, MA tries to detect the “truth” lying behind a number of studies focusing on the same phenomenon, all of which can be affected by some bias/error
- Meta-regression is a statistical tool to perform a meta-analysis. It enables us …
  - ✓ to assess the influence of some programme or study characteristics on the size / probability of particular results, e.g. size of treatment effects or probability of positive treatment effects
  - ✓ to test whether the influence found in the sample of studies under scrutiny is statistically significant, i.e. to establish if it caused by something other than mere random chance

Data

Creation of the database with our systematic review:

- ✓ literature search , investigation of the reference list of the retrieved studies, questions to colleagues
- ✓ 43 published and unpublished articles written at any time, adopting the tools of the conterfactual approach (Imbens and Wooldridge, 2009) * 478 estimates

Categorisation into variables of the information contained in the studies:

- ✓ Outcome variable: treatment effect $y_i = \begin{cases} 1 & \text{if the estimate is significantly positive} \\ 0 & \text{otherwise} \end{cases}$

- ✓ Predictors: type of policy, policy level at which the intervention is implemented, target of the interventions, type of incentives, year in which the programme is implemented, type of outcome on which treatment effects are estimated, timing of estimated impact, number of firms involved in the estimation, and basic methodology used for estimation, publication status of article, …

Note that the outcome variable and some predictors are measured at the level of estimates, while other predictors are defined /constant at the study level!
Note also that each study usually contains a number of estimates (11 on average)

	At the level of estimates	At the level of studies
	Mean	Group mean
Response variable: treatment effect is significantly positive	0.337	
At least one treatment effect is significantly positive		0.907
Variables that are constant within studies		
Study was published in a journal	0.536	0.651
Study uses administrative rather than survey data	0.900	0.837
Programme type		
R&D	0.559	0.512
investments	0.343	0.372
bank loans	0.098	0.116
Variables that are not always constant within studies		
Outcome directly affected by the programme	0.297	0.356
Non simultaneous treatment effect	0.609	0.442
N. of firms involved in estimation	4158.255	5085.834
Target firms		
Target all firms	0.776	0.605
Target SMEs only	0.140	0.244
unspecified	0.084	0.151

	At the level of estimates	At the level of studies
	Mean	Group mean
Government level delivering the programme		
national	0.362	0.430
regional	0.554	0.419
unspecified or mixed	0.084	0.151
Incentive type		
unspecified or mixed	0.109	0.197
loan	0.289	0.201
grant	0.554	0.528
tax credit	0.048	0.074
Basic methodology used for estimation		
DID	0.201	0.205
RDD	0.098	0.128
matched DID	0.425	0.209
matching	0.218	0.322
other	0.059	0.136
Year of the programme		
late 2000s	0.149	0.209
earlier	0.851	0.791
Number of observations	478	43

Vote counts & other descriptive statistics

Type of programme	Significantly positive	Insignificant	Significantly negative	Total
promotes R&D	76 (28.5%)	183 (68.5%)	8 (3.0%)	267 (100%)
Investment	59 (36.0%)	87 (53.0%)	18 (11.0%)	164 (100%)
Bank loans	26 (55.3%)	16 (34.0%)	5 (10.6%)	47 (100%)
Total	161 (33.7%)	286 (59.8%)	31 (6.5%)	478 (100%)

Type of incentives	R&D	Investment subsidies	Bank loans	Total
Unspecified or mixed	51 (19.1%)	1 (0.6%)		52 (10.9%)
Loans	88 (33.0%)	3 (1.8%)	47 (100.0%)	138 (28.9%)
Grants	124 (46.4%)	141 (86.0%)		265 (55.4%)
Tax credits	4 (1.5%)	19 (11.6%)		23 (4.8%)
Total	267 (100.0%)	164 (100.0%)	47 (100.0%)	478 (100.0%)

Goals and form of the meta-regression

- We are interested in the probability that the response is 1 as a function of:

  i) the predictors $\boldsymbol{x}_i$ and ii) a term of unobserved heterogeneity at the study level u_s

  $E(y_i|\boldsymbol{x}_i, u_s)=Pr(y_i= 1 |\boldsymbol{x}_i, u_s)$

- since the response variable is binary, we must use an appropriate regression model, such as a logit

- We estimate the following random-intercept multilevel model

$$logit\{\Pr(y_{is} = 1|\boldsymbol{x}_{is}, u_s)\} = \beta_0^C + \boldsymbol{\beta}^C \boldsymbol{x}_{is} + u_s$$

where coeffcients β^C represents the change in the log odds ratio of having a significantly positive treatment effect estimate for a one unit increase in the predictor, conditional on u_S. The latter refers to the random error component for the deviation of the intercept of a group from the overall intercept.

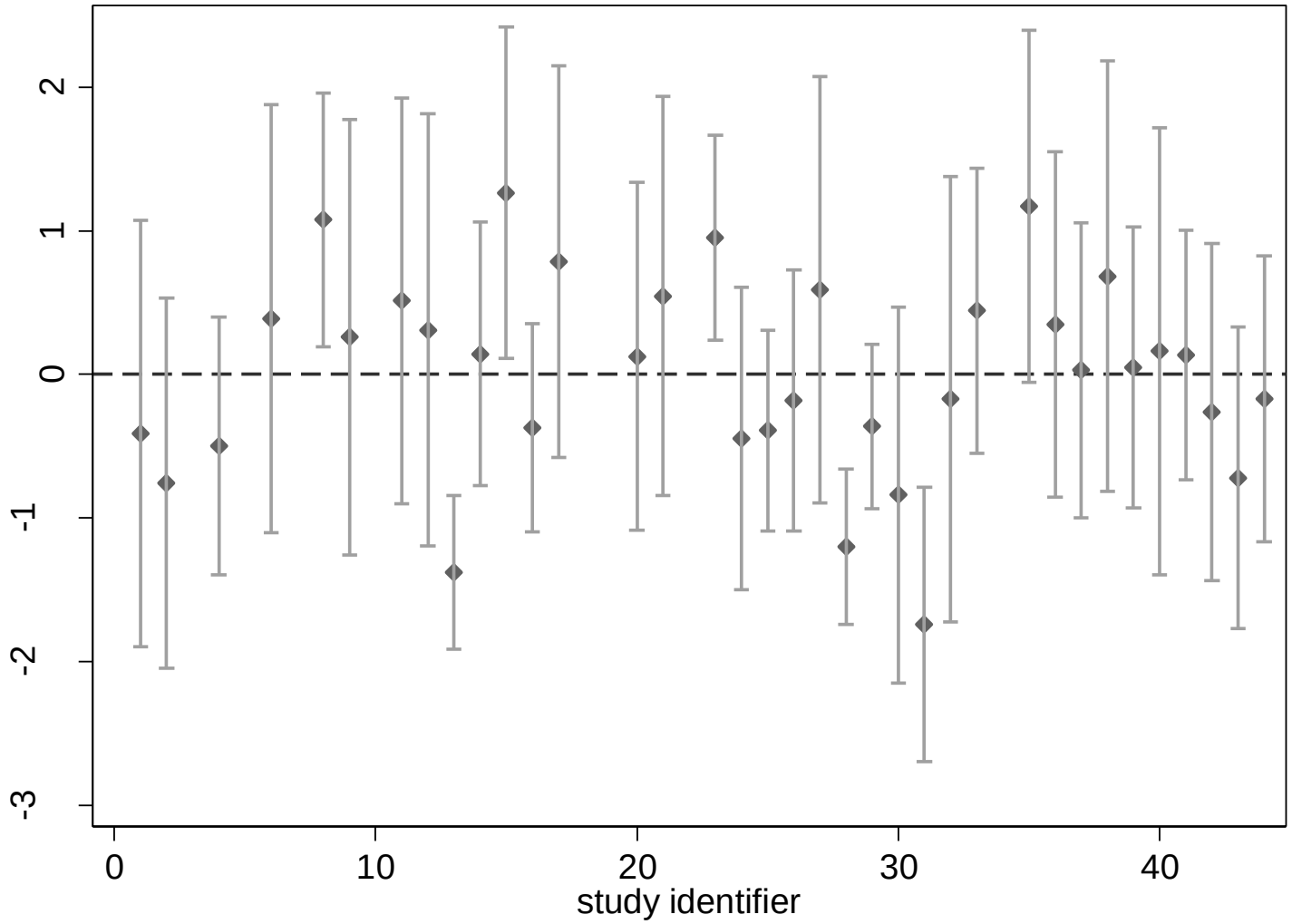
By means of the following nonlinear transformation we can use coefficients to compute probabilities

$$\Pr(y_{is} = 1|\boldsymbol{x}_{is}, u_s) = \frac{\exp(\beta_0^C + \boldsymbol{\beta}^C \boldsymbol{x}_{is} + u_s)}{1 + \exp(\beta_0^C + \boldsymbol{\beta}^C \boldsymbol{x}_{is} + u_s)}$$

More on unobserved study heterogeneity

- could be due, for example, to the unobserved ability of the authors in framing the study or obtaining credible estimates, or also it might depend on their determination to search for particular results
- explanations of u_s can be only hypothetical, since it captures the "joint average" influence on *Y* exerted by all aspects that are not represented by observable predictors
- in order to estimate the study-specific deviation from the overall intercept, we must hypothesise that it follows some particular distribution. The usual prior is $u_s \sim N(0, \sigma_u^2)$
- once having estimated variance σ_u^2 we test whether it is significantly different from zero. Intuitively, the idea is that the greater this variance, the less negligible unobserved study heterogeneity is
- if one is interested in probability computations that are net of the term of unobserved study heterogeneity, these can be obtained by fixing all u_s at their mean value of zero

Estimated study heterogeneity



$\sigma_u^2 = 0.957$

This variance is statistically different from zero

Results for some common policy schemes

- it may be interesting use the coefficients estimated net of the terms of unobserved study heterogeneity to compute the predicted probability of success associated with particular combinations of predictors, corresponding to some of the most common schemes in the area of industrial and innovation policies
- in order to obtain these probabilities we fix some predictors at particular values representing policy schemes, we also fix all u_s at their mean value of zero

THREE POLICY SCHEMES

A. R&D grant, targeting both small and larger firms
B. Guaranteed loan for SMEs only
C. Investment grant, targeting both small and larger firms

We predict probabilities of success depending on the fact that:

- the outcome variable to which the treatment effect refers is a variable that the programme in question is intended to modify in a direc way
- the government level delivering the programme is national or regional

R&D grant for all firms

Average adjusted probability predictions; random effects fixed at zero

	(A) whatever level	*(B) national level*	*(C) regional level*	*(C - B) difference*
DIRECTLY AFFECTED OUTCOME	0.732***	0.596**	0.813***	0.217
	(0.070)	(0.232)	(0.083)	(0.145)
OTHER OUTCOME	0.188***	0.100*	0.245***	0.145*
	(0.061)	(0.056)	(0.083)	(0.080)

Notes. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Guaranteed loan for SMEs

Average adjusted probability predictions; random effects fixed at zero

	(A) whatever level	*(B) national level*	*(C) regional level*	*(C - B) difference*
DIRECTLY AFFECTED OUTCOME	0.715***	0.575***	0.799***	0.224
	(0.161)	(0.215)	(0.145)	(0.139)
OTHER OUTCOME	0.461**	0.309	0.557**	0.248*
	(0.214)	(0.203)	(0.233)	(0.137)

Notes. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Investment grant for all firms

Average adjusted probability predictions; random effects fixed at zero

	(A) whatever level	*(B) national level*	*(C) regional level*	*(C - B) difference*
DIRECTLY AFFECTED OUTCOME	0.675***	0.527***	0.764***	0.238*
	(0.112)	(0.146)	(0.116)	(0.131)
OTHER OUTCOME	0.501***	0.346***	0.599***	0.253*
	(0.105)	(0.115)	(0.126)	(0.137)

Notes. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$