Regional Programme Evaluation: Challenges and Perspectives

Guido de Blasio (Bank of Italy)
Alexander C. Lembcke (OECD)
Carlo Menon (OECD)

Rigorous Impact Evaluations in Europe, Torino May 2018

Outline

- The rationale and the empirics of place-based policy
- Challenges in evaluating place-based programmes (spatial shuffling)
- How to make research and policy more interlocked: ML targeting

Space and policies

- Should polices explicitly target particular places (World Bank, 2008; OECD, 2011)? Place-based policies versus people-centered policies, i.e., policies that do not take differences across cities or regions into account ("space blind")
- An abundant literature documents agglomeration economies and human capital externalities. These externalities suggest market failure, but they do not imply any particular spatial policy. Empiric analysis is needed to check whether intervention is welfare/equity enhancing
- Austin et al (2018) suggest that the rationale might be found in the heterogeneous response of employment rates to policy interventions. If Bartik shocks affect lagging regions more strongly, then in principle there is a case for regionally targeted support. Empirical analysis is required to check the plausibility of this rationale on a case-by-case basis and (ex post) whether interventions delivers on these premises

Financing of place-based policies

- Regional policy will remain one of the largest spending areas in the EU. For 2021-2027 EUR 273 billion are allocated to ERDF and CF (only CAP has a larger budget)
- Cities in the United States are offering incentive packages in the USD billions to attract Amazon's second headquarter
- In the US state and local governments spend USD 30-40 billion per year and the federal government another USD 8-12 billion on policies aimed at creating local jobs (Moretti, 2011)

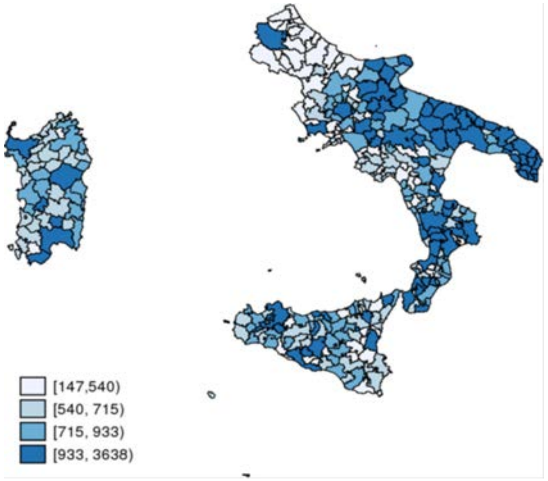
A (possibly) useful taxonomy

- Wanted effects (+ investment, - unemployment) & unwanted effects (+ corruption, +rent seeking)
- Even if wanted effects are positive & unwanted are zero the case for redistribution is not yet established.
- ➢ Efficiency requires that the benefits (for the targeted places) are higher than the costs (for the untargeted places, which pay the bill)
- At this stage, equity might come in and inform desirability (1 percent increase in unemployment in a high-unemployment area could be desirable wrt the same increase in a low-unemployment area)
- Below, few examples from the case of Italy, where wanted are zero, unwanted are likely to be positive, and thus equity considerations are trivial

Structural funds, 2007-14

- Based on Ciani, de Blasio (IZA JLP, 2015)
- Southern LLMs
- DID with continuous treatment
- LLM fixed effects, plus controls to deal with time-varying omitted at the LLM-level, also ML selected
- Outcomes: Employment, Population and House Prices

Quantiles of cumulative per capita payments (LLM map)



Effectiveness of Structural Funds

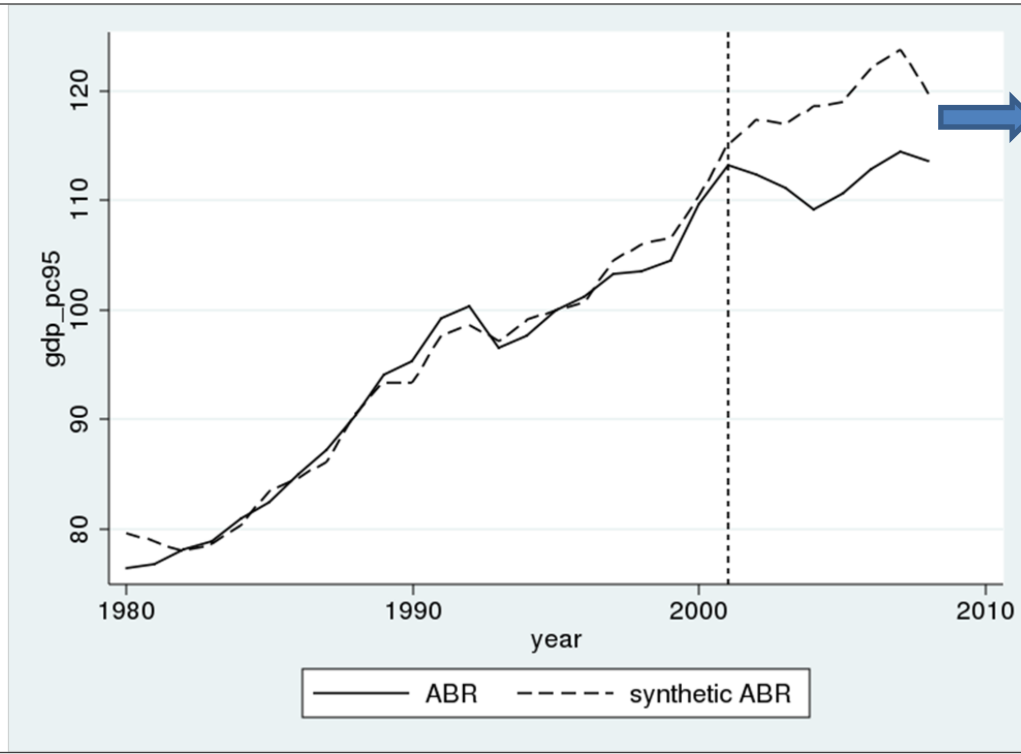
	Employment			Population			House Prices		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Year-to-year specifications									
EU Payments	.0028	0.0010	0.0014	-.0004	-.0000	-.0003	.0063	.0012	.0029
	(.0017)	(.0017)	(.0014)	(.0002)**	(.0002)	(.0002)	(.0025)**	(.0020)	(.0021)
Methods:	FE	fi'; fi'*t; fi'*t2	Double selection	FE	fi'; fi'*t; fi'*t2	Double selection	FE	fi'; fi'*t; fi'*t2	Double selection
Overtime average specifications									
EU Payments	-	-.0001	-.0035	-	.0002	-.0002	-	-.0118	-.0093
	-	(.0026)	(.0026)	-	(.0007)	(.0007)	-	(.0049)**	(.0044)**
Methods:		fi'	Double selection		fi'	Double selection		fi'	Double selection

Self-sustaining growth?

- Based on Barone, David, de Blasio (RSUE, 2016)
- Treatment and cure: “A treatment is an instance of treating someone, say, medically. A cure ends a problem. Sometimes, the treatment is a cure. Other times, it just keeps the problem under control without curing it: *if you remove the treatment, the problem comes back*” (Ozler, 2014)
- Study the unique case of the Italian region Abruzzi: entered the Objective 1 program in 1989 and exited in 1996 (financially in 2000) without any transitional support
- (Synthetic Control) results suggest that the gains under the program were completely eroded afterwards

Post-exit growth performance

Baseline result: GDP per capita 1980-2008
(index 1995=100)



In 2008: -5.5%

The graph reports the GDP per capita in real terms (1995=100) of the treated region (Abruzzi) and of the synthetic control. The weights used to build the synthetic controls are 0.641 (Molise), 0.200 (Campania) and 0.159 (Calabria).

Unwanted effects of Structural funds

- Based on De Angelis, de Blasio, Rizzica (2017, BIWP)
- EU safeguards to prevent fraud (co-financing by local authorities to improve accountability, centralized auditing mechanism, possibility of funds withdrawal if no adequate reporting and certification of use of funds).
- ➢ However, in Italy more than 90 percent of co-financing comes from national resources, although most projects are managed at the local level
- SDI archive, records of all the crimes committed in Italy taken from the IT system used by the police for investigations. Provides an instantaneous picture of the criminal activity in each municipality. Less subject to underreporting (which is pervasive in the case of corruption crimes)
- White collar crimes are defined as all crimes against the public administration and against public faith, which include corruption, bribery, embezzlement, abuse of authority and fraud
- # of crimes, no info on # people or amounts of money involved (both politicians and local bureaucrats)

White collars crimes and EU funds

	(1)	(2)	(3)
Log EU	0.084***	0.073***	0.042***
	(0.020)	(0.018)	(0.017)
FEt	YES	YES	YES
Time varying ctrls	NO	YES	YES
FEm	NO	NO	YES
Obs	6009	6009	6009

Notes: Poisson regressions, dependent variable is the number of white collar crimes in municipality i in year t. Controls are (log) population, labor market participation rate, unemployment rate, number of new college graduates per 1,000 inhabitants, yearly regional GDP growth, years from elections, second mandate major. Bootstrapped standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 .

Spatial shuffling

- The main threat to identification for evaluating place-based policies
- An area is targeted but the effects of the policy will depend on the endogenous relocation of L and K in the aftermath of the intervention (Duranton and Venables, 2018)
- Shuffling might be part of the policy objectives (see the examples below)

Example of spatial shuffling #1

- Based on Andini, de Blasio (JoEG, 2016)
- Major Italian place-based policy (Contratti di Programma), by means of which the state approves and finances investment projects proposed by private firms
- Using the areas to be exposed to the same policy at a later date as counterfactuals, it is estimated a limited impact on plant and employment growth rates, which is confined to a small area (a single municipality) and likely crowds out the economic growth of the surrounding areas
- Mechanisms for crowding-out: wages and rents
- It is hard to say that spatial shuffling in this case can be accounted as a welcome effect of the policy, since it is not a relocation from rich to poor areas (it is a within-poor areas spatial shuffling)

Example of spatial shuffling #2

- Based on de Blasio, Poy (JRS, 2017)
- The role of wage zones—compulsory wage differentials at the province level—on Italy's local labor markets during the 1950s
- Using spatial regression techniques, it is estimated that for the industrial sectors covered under wage zones there was an increase in employment when one crossed the border from a high-wage province into a low-wage one; the effect diminished, however, as the distance from the boundary increased
- The new job opportunities in low-wage provinces came at the expense of having reduced job opportunities in the bordering high-wage provinces. Moving costs were relevant and differences in remunerations did not compensate for that: between-province commuters (rather than migrants) seem to have taken the bulk of the adjustments (auxiliary results suggest that the residing population did not move)
- The policy ended in a 50km shuffling of K and L

Ex post evaluations and policy

- Ex-post evaluations deliver rigorous assessments of whether a programme has worked or not
- In some instances, a policy maker might want to know (ex-ante) how to design the program to maximize its effectiveness (targeting)
- Have a look at the results from similar interventions implemented in comparable context!
- However, it is very difficult to extrapolate from them: no guarantee that the same program produce the same effect in two different contexts, such as two different areas or even the same area in a dissimilar conjuncture
- Current identification strategies allow researchers to identify causal effects only for subpopulations, which might be quite different from the population of policy interest (LATE)
- ML tools helps to target the intervention towards the carries of effectiveness

ML prediction tools (in a nutshell)

- ML tools are useful for out-of-sample prediction exercises. They randomly split the dataset in a training sample and a testing sample
- The training sample is used to estimate a prediction model of Y on a set of Xs. The testing sample is used to evaluate the model performance, using realizations of the Xs not involved in the estimation. Once the model with best performance has been chosen, they allow to out-of-sample predict Y on new Xs
- ML tools rely on highly flexible functional forms, however: + flexibility in the model (+ complexity) → improve in-sample fit (-bias) vs reduce out-of-sample fit (+ variance). To address this trade-off, each ML algorithm comes with a regularization parameter, chosen through cross validation

ML and targeting

1) ML predict the carriers of effectiveness, by using data on those who most likely will ensure the achievement of the policy objective; use ex-post machinery to validate the prediction

This is the route followed by Andini et al (2017, 2018): ML prediction of consumption-constrained HHs; credit-constrained firms

2) Run a randomized experiment (selection on observables might do) and identify heterogeneous effects; use ML machinery to predict the observables mostly related to highest effectiveness (predict those units)

This is the route followed by Ascarza (2018) in the context of churning

In 1) need to have data on the carries of effectiveness. In 2) need to have a program implemented beforehand

Route 1): Andini et al (2017)

- The idea is to predict who is more likely to comply with a given policy, thus reducing the misallocation of benefits
- Andini et al, 2017 apply ML tools to the 80 euro bonus policy
- Identify consumption constrained households using the SHIW and show that targeting them would have been better

The program “80 euro”

- In Spring 2014 the Italian Government introduced a large scale tax credit
- The tax rebate aimed at providing a short-run boost to consumption (also reshaping labor taxation)
- It was channeled to employees with annual income between 8,145 and 26,000 euro (640 euro/year if income< 24,000 and declined to zero until the 26,000 threshold)
- The total transfer was almost 7 billion euro in 2014 (0:4% of Italian GDP)

The target group

- The impact of a fiscal policy on consumption clearly depends on the heterogeneity of marginal propensity to consume
- Neri et al (2018) find that the tax rebate had an impact on consumption, but it was stronger for consumption constrained households
- In the Bank of Italy's Survey on Household Income and Wealth (SHIW) we have a (reasonably good) proxy of consumption constraints
- We look at household reporting to make ends meet with at least some difficulty

Targeting with ML

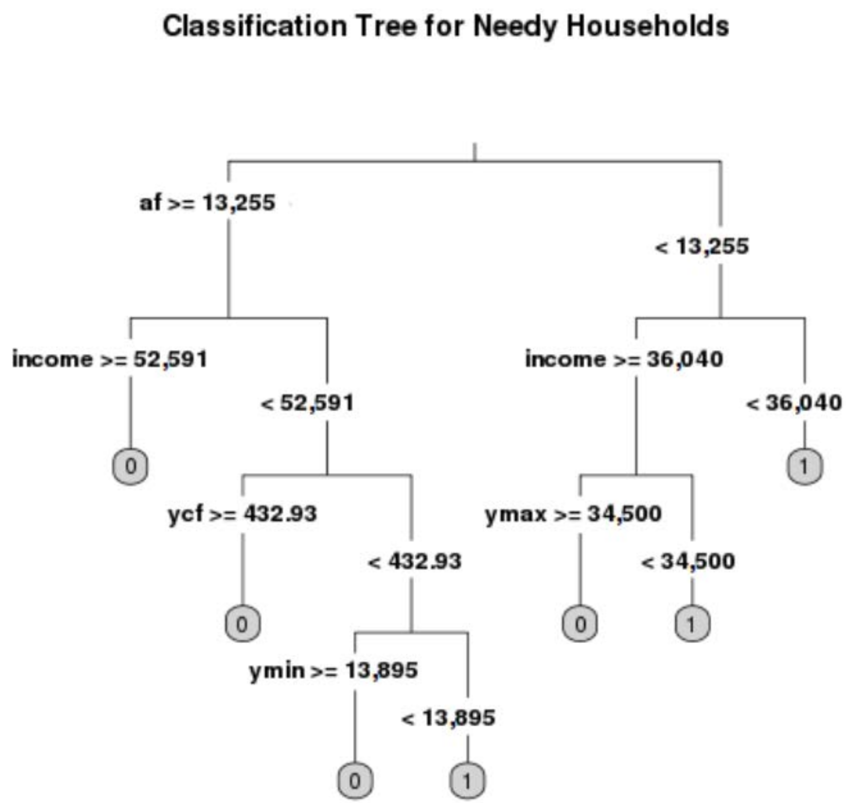
- The policy-maker cannot obviously know who are the households with difficulties in making ends meet
- Our targeting exercise aims at identifying them, through the following steps:

    1) We use the 2010 and 2012 SHIW to train and test a model for the constrained status on the basis of observed covariates

    2) We predict this status on the 2014 wave and we assess whether the impact of the tax rebate was larger among predicted constrained households

    3) We show how much the predicted needy households overlap with actual recipients of the rebate

80 euro, ML rule



80 euros, prediction

		Predicted Status			
		Not Needy	Needy	Total	Overlapping (%)
Real Status	Not Recipient	715	1446	2161	33.0%
	Recipient	431	1054	1485	70.9%
	Total	1146	2500	3646	
	Overlapping (%)	62.4%	42.1%		48.5%

The actual assignment rule incorrectly targeted about half of the households. 29% of actual expenditure has been allocated to non-constrained recipients

Actual consumption, out of the bonus

	Total consumption		Food consumption	
	(1)	(2)	(3)	(4)
Predicted status:	*Not consumption constrained*	*Consumption constrained*	*Not consumption constrained*	*Consumption constrained*
Bonus (amount)	-0.527	0.710**	0.009	0.369***
	(0.563)	(0.315)	(0.184)	(0.111)
N	1146	2500	1146	2500
R2	0.459	0.415	0.356	0.442

Notes. See Andini et al (2017) for details.

Possibilities for ML in the field of location-based policy

- Calculate local measures for consumption-constrained HHs or credit-constraints firms, allocate transfers accordingly (Route 1)
- Predict observables related to higher effectiveness, exploiting the ex-post evaluations already done (Route 2)
- Improve geo-targeting by using ML prediction of the evolution of economic activity at detailed levels of geo-stratification (Glaeser et al, 2017)
- Inform decision on where to locate infrastructures (Chin et al, 2017)

Open issues with ML

- Interpretability/transparency
    - Formal vs substantive transparency
- Omitted payoffs
- Manipulation
- Prediction stability