Finanziato dall'Unione europea
NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA

Consiglio Nazionale delle Ricerche

UK Stata Conference
London
3-4 September 2026

Optimal policy learning under budget and coverage constraints: A Stata implementation

Giovanni Cerulli
Research Director
IRCrES–CNR
Research Institute on Sustainable Economic Growth
National Research Council of Italy

DEFINITION OF OPL

- **What is policy learning?**

  Process of improving program **impact/welfare** achievements by re-iterating policies over time

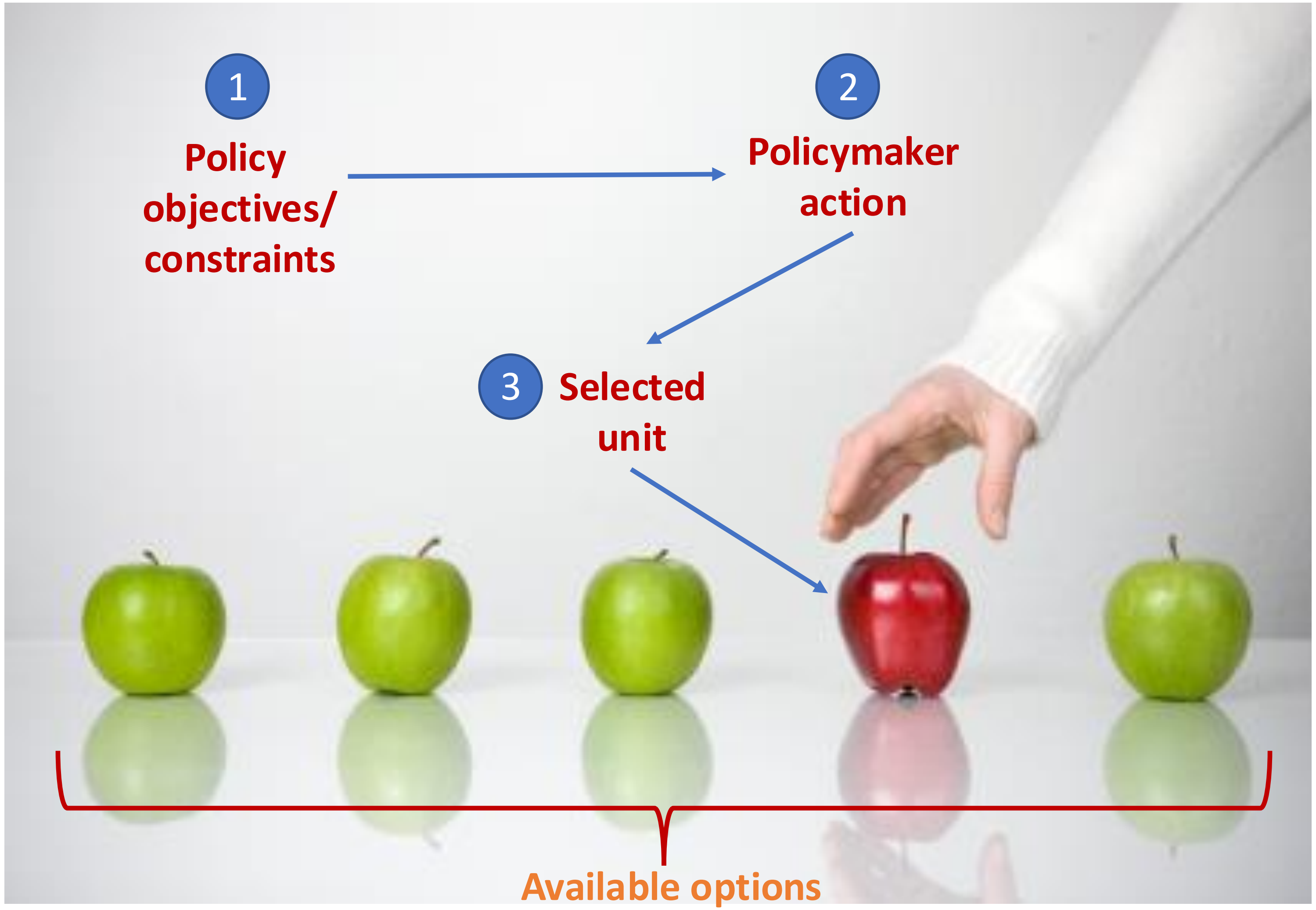
- **Optimal treatment assignment**

  Policymakers can **optimally fine-tune the treatment assignment** of a prospective policy using the results from an RCT or observational study. Assignment rules depends on the **class of policies** considered (here we focus on **threshold-based,** **linear-combination** and **decision-tree** policies)

- **Maximizing constrained impact/welfare**

  The policymaker hardly manage to reach the best solution (**unconstrained maximum impact/welfare**) because of institutional/economic contains of various sort

Policy as a *selection problem*



Example: Active Labor Policy

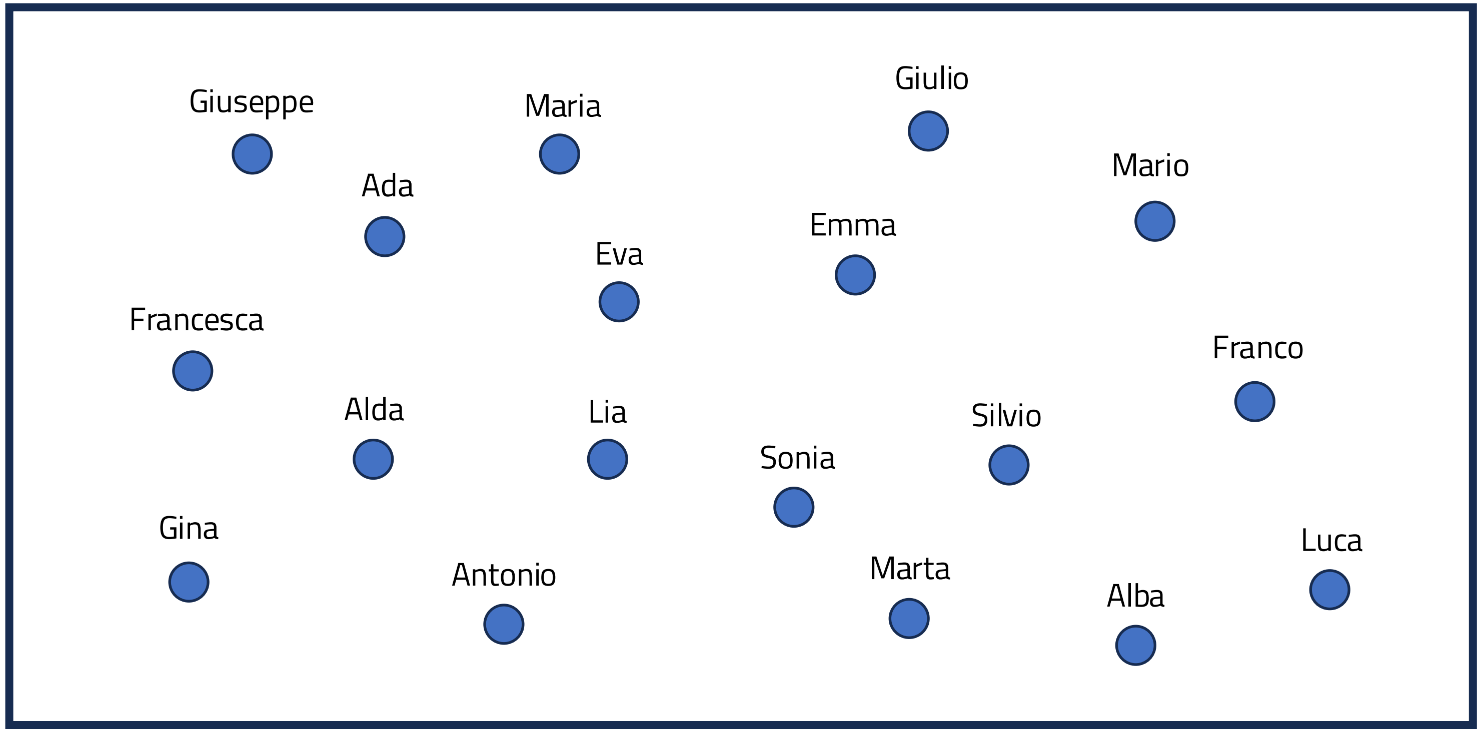
A 10k euro voucher to unemployed people for maximizing the probability to become employed in two years

Population = 18 unemployed people
Policy = [Voucher for training; Nothing]
Objective = Probability to become employed
Budget = 100K
Cost = 10K
Coverage = 100K/10K = 10 People to support

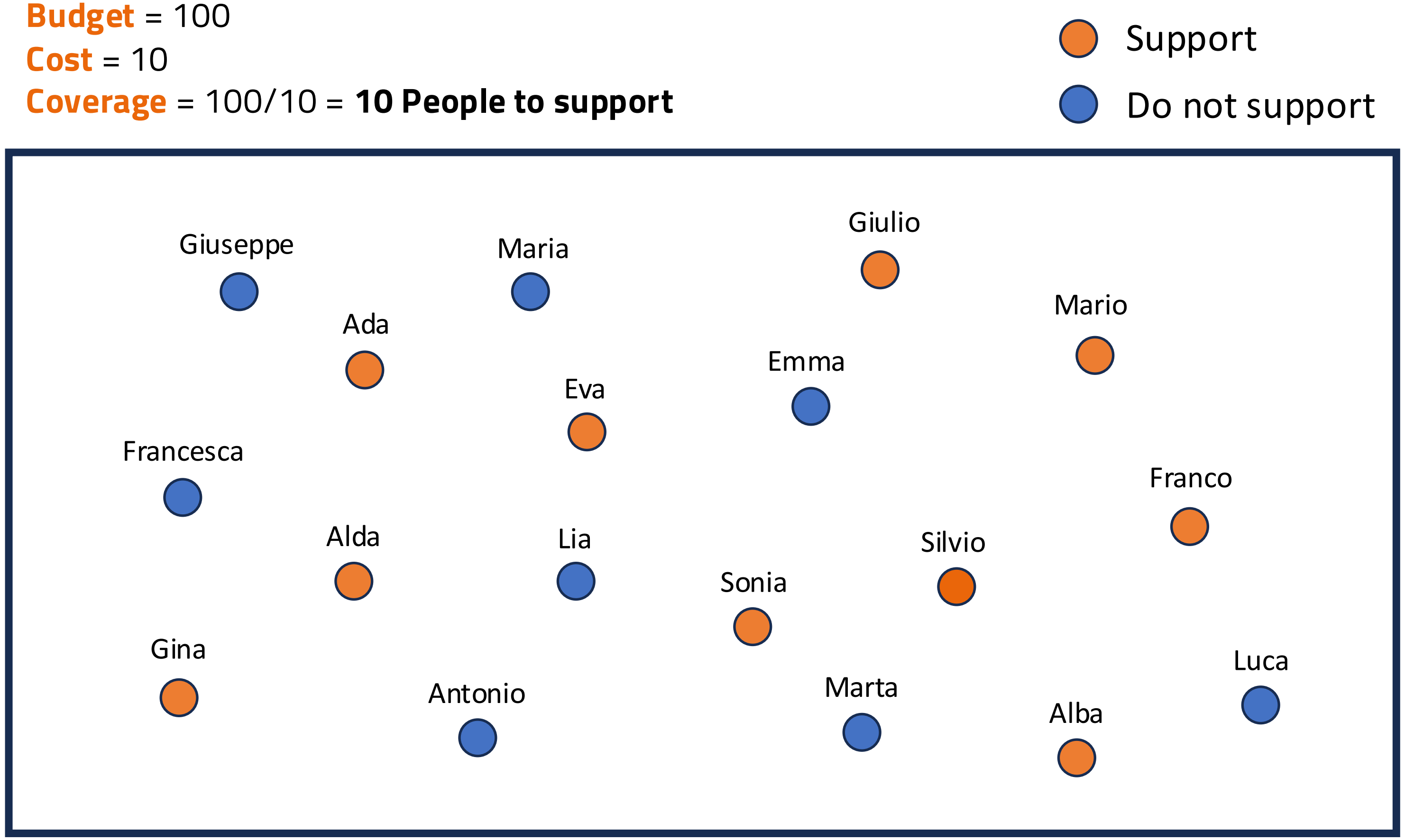
Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
Consiglio Nazionale delle Ricerche

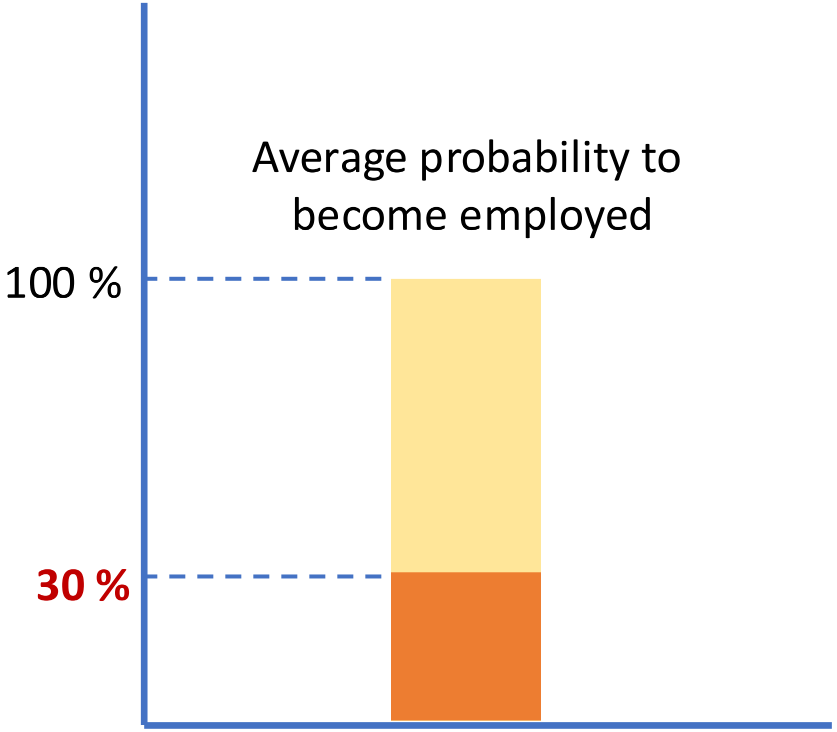
Population = 18 unemployed people
Policy = [Voucher for training; Nothing]
Objective = Probability to become employed in two years
Budget = 100
Cost = 10
Coverage = 100/10 = **10 People to support**

Population to select for support



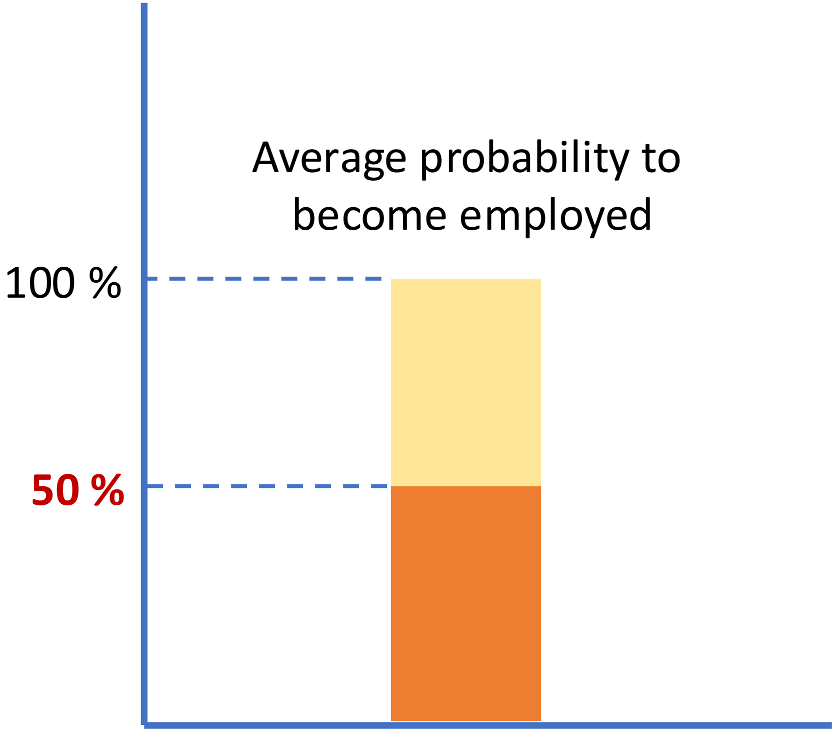
Population = 18 unemployed people
Policy = [Voucher for training; Nothing]
Objective = Probability to become employed in two years
Budget = 100
Cost = 10
Coverage = 100/10 = **10 People to support**





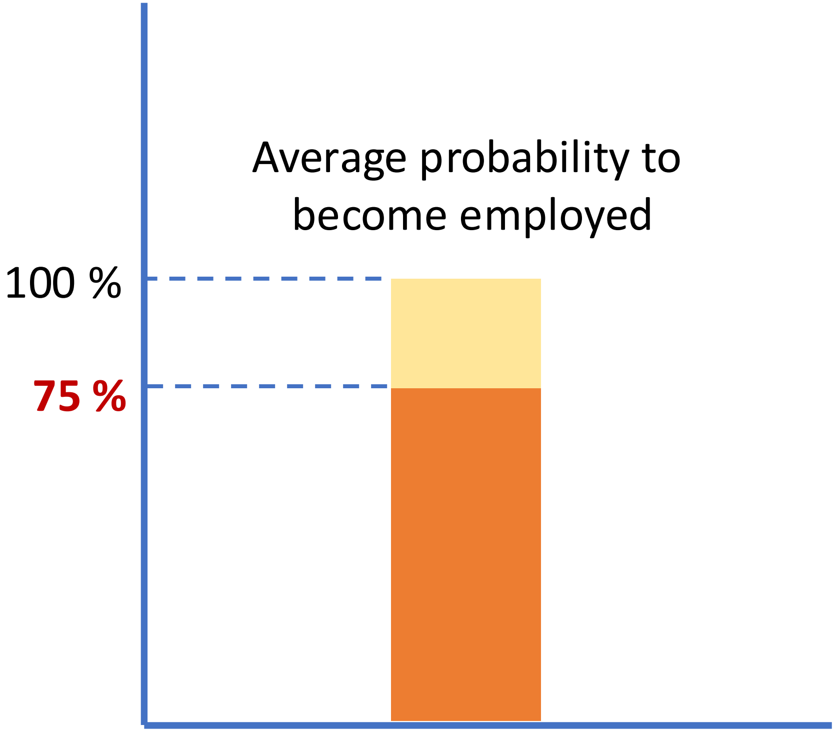
Population = 18 unemployed people
Policy = [Voucher for training; Nothing]
Objective = Probability to become employed in two years
Budget = 100
Cost = 10
Coverage = 100/10 = **10 People to support**

Support
Do not support



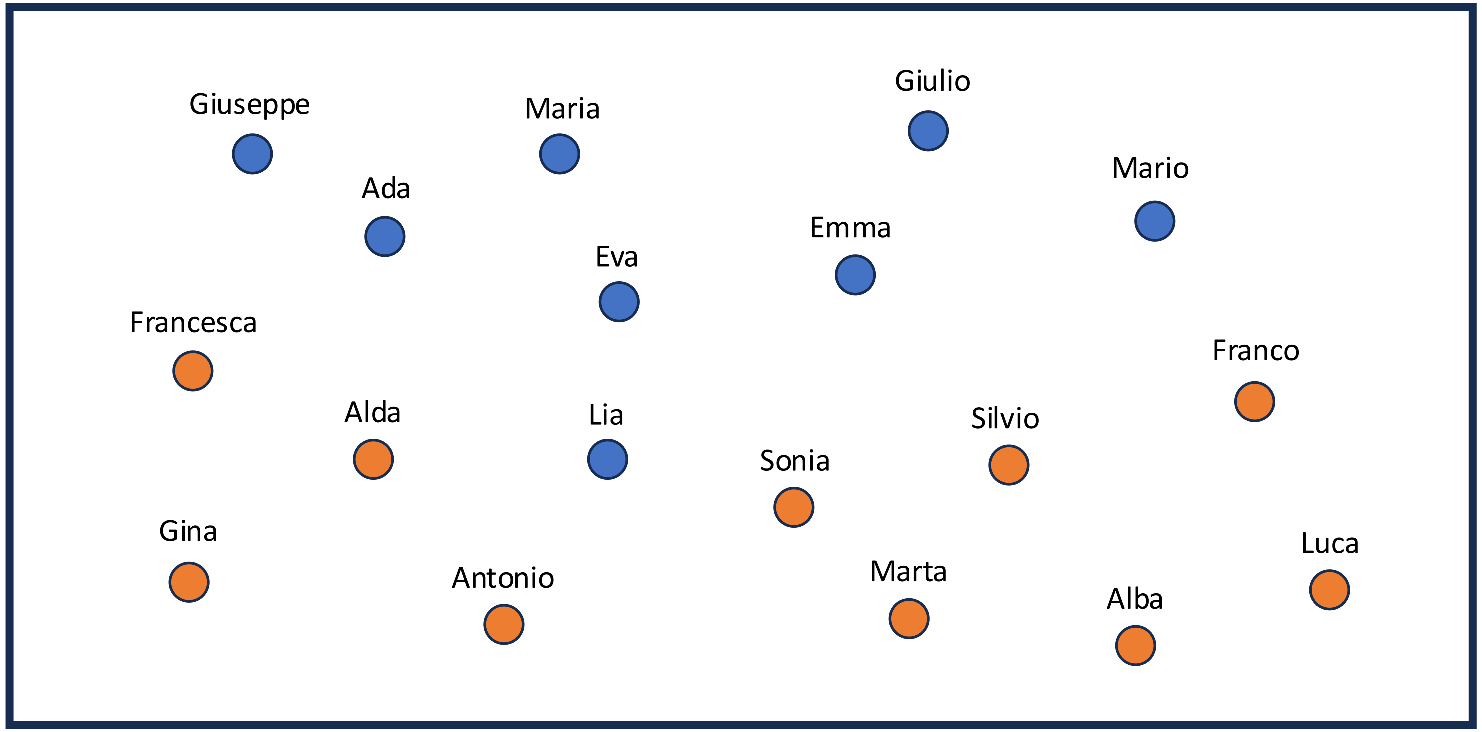
Population = 18 unemployed people
Policy = [Voucher for training; Nothing]
Objective = Probability to become employed in two years
Budget = 100
Cost = 10
Coverage = 100/10 = **10 People to support**

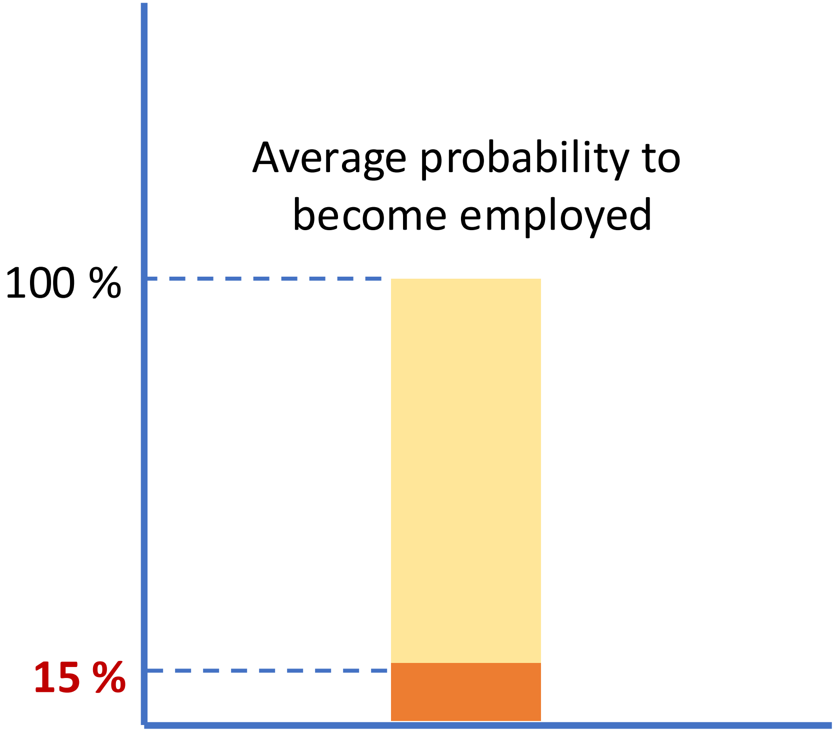
Support
Do not support



Population = 18 unemployed people
Policy = [Voucher for training; Nothing]
Objective = Probability to become employed in two years
Budget = 100
Cost = 10
Coverage = 100/10 = **10 People to support**

Support
Do not support

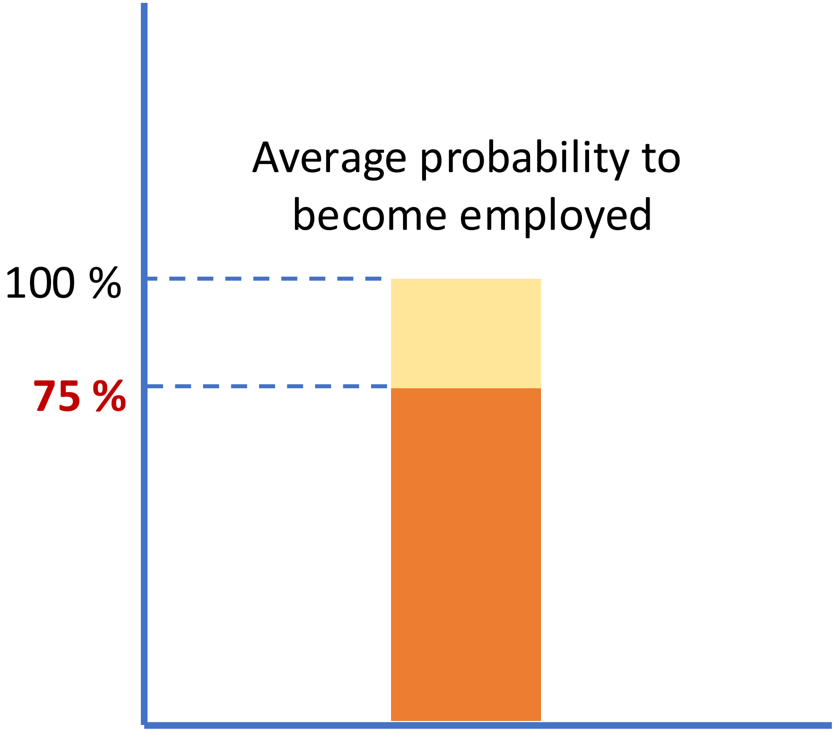




Population = 18 unemployed people
Policy = [Voucher for training; Nothing]
Objective = Probability to become employed in two years
Budget = 100
Cost = 10
Coverage = 100/10 = **10 People to support**

If this is the selection producing the largest employability, this will be the **optimal policy** to implement (**best selection of beneficiaries**)

Support
Do not support



What does OPL maximize?

Value-function
(Welfare)

or

Total Impact
(TTET)

Average Impact
(ATET)

A policy which maximizes *Welfare* maximizes also the *Total Impact*

Policy definition

Consider a decision maker who would like to allocate a set of M different treatments over a given population based on individual characteristics, with the aim of maximizing a given measure of welfare (income, employability, recovery from a disease, etc.).

For this decision maker, a policy $\pi(X)$ is defined as a decision rule, mapping individual characteristics onto one of M alternative treatments (or actions), that is:

$$\pi : X \longrightarrow \pi(X) = A \in \mathcal{A}$$

Binary Policy

A (binary) *policy* $\pi(X)$ is a mapping assigning a treatment to each individual based on her covariates:

$$\pi(X) : \mathcal{X} \longrightarrow \mathcal{A} = \{0; 1\},$$

- $\mathcal{X}$: space of individual characteristics (covariates).
- $\mathcal{A}$: set of two treatments, either 0 or 1.
- $\pi(X_i)$: rule assigning treatment to unit i under policy π.

Value-function

(Welfare)

The value-function (or *welfare*), $V(\pi)$, measures the effectiveness of a policy π in determining the expected outcome (Y):

$$V(\pi) = \mathbb{E}\left[Y(\pi(X))\right]$$

The *optimal policy* π^* is the one that maximizes this function:

$$\pi^* = \arg\max_{\pi} \mathbb{E}\left[Y(\pi(X))\right]$$

Individual Treatment Effect (ITE)

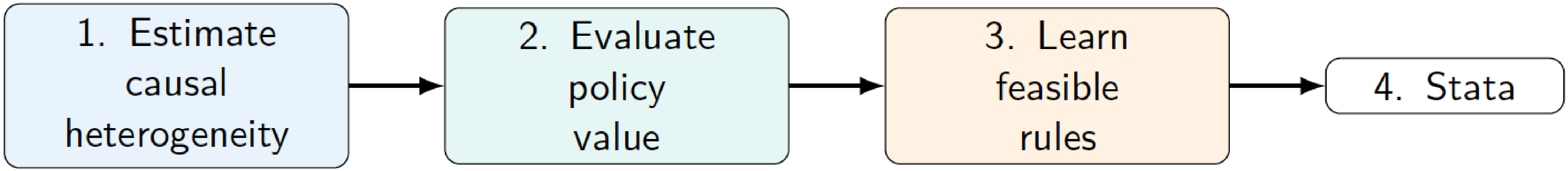
Let $Y_i(1)$ and $Y_i(0)$ denote the two potential outcomes when unit i is either treated, or untreated. We define, the **Conditional Average Treatment Effect** as:

$$\tau(X_i) = \mathbb{E}[Y_i(1) - Y_i(0) \mid X_i]$$

This is a measure of an **individual treatment effect** (ITE).

By the LIE, we have that $E_{X_i}[\tau(X_i)] = \text{ATE}$, the Average Treatment Effect.

Estimation pipeline



- Move from ATE to CATE, GATE and IATE.
- Understand first-best and constrained policies.
- Compare thresholds, scores, trees, budgets and multi-action rules.
- Translate methods into replicable Stata workflows.

Welfare maximization with Budget and Coverage Constrains

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
Consiglio Nazionale delle Ricerche

Motivation

Optimal Policy Learning (OPL) often requires selecting individuals to treat under realistic policy constraints.
Two key constraints frequently arise:

- A total budget constraint.
- A minimum coverage requirement (e.g. political or fairness constraints).

In addition, treatment costs may vary across individuals.
Goal: characterize the optimal treatment assignment rule when both constraints are present.

Finanziato dall'Unione europea
NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA
Consiglio Nazionale delle Ricerche

Setting

Consider N individuals indexed by $i = 1, \ldots, N$.
For each individual:

- $\hat{\tau}_i$ = estimated treatment effect
- $c_i > 0$ = individual treatment cost

Binary treatment decision:

$$d_i \in \{0, 1\}$$

where $d_i = 1$ indicates treatment.

Maximizing Total Welfare, under **budget constraint** and **minimum** number of treated units

Primal empirical welfare maximization:

TTET
(Total Treatment Effect on Treated)

$$\max_{d \in \{0,1\}^N} \sum_{i=1}^{N} \hat{\tau}_i d_i \quad \text{s.t.}$$

- budget constraint with budget $C > 0$:

$$\sum_{i=1}^{N} c_i d_i \leq C;$$

- minimum coverage (cardinality) constraint with $N^* \in \{1, \ldots, N\}$:

$$\sum_{i=1}^{N} d_i \geq N^*.$$

Joint Lagrangian

Rewrite the constraints in the standard form $g(d) \leq 0$:

$$g_1(d) = \sum_{i=1}^{N} c_i d_i - C \leq 0, \qquad g_2(d) = N^* - \sum_{i=1}^{N} d_i \leq 0.$$

Introduce Lagrange multipliers $\lambda \geq 0$ for $g_1(d)$ and $\nu \geq 0$ for $g_2(d)$. The joint Lagrangian is

$$\begin{aligned}\mathcal{L}(d, \lambda, \nu) &= \sum_{i=1}^{N} \hat{\tau}_i d_i - \lambda\Big(\sum_{i=1}^{N} c_i d_i - C\Big) - \nu\Big(N^* - \sum_{i=1}^{N} d_i\Big) \\ &= \sum_{i=1}^{N} (\hat{\tau}_i - \lambda c_i + \nu) d_i + \lambda C - \nu N^*.\end{aligned}$$

For fixed (λ, ν), maximizing (2) over $d \in \{0,1\}^N$ is separable across individuals.

Proposition

For fixed (λ, ν) *with* $\lambda \geq 0$ *and* $\nu \geq 0$, *the optimal relaxed decision satisfies*

$$d_i^*(\lambda, \nu) = 1\{\hat{\tau}_i - \lambda c_i + \nu \geq 0\}.$$

Interpretation:

$$\underbrace{\hat{\tau}_i + \nu}_{\text{Benefit}} \geq \underbrace{\lambda c_i}_{\text{Cost}}$$

The coverage constraint effectively increases the incentive to treat individuals.

The decision rule can be written as

$$\frac{\hat{\tau}_i}{c_i} \geq \lambda - \frac{\nu}{c_i}.$$

Cost-effectiveness

Key implication:

- The cutoff depends on c_i.
- A single cost-effectiveness threshold generally does not exist.

Solution of the optimal problem

The resulting allocation problem becomes a combinatorial optimization problem.

In particular, it corresponds to a:

0–1 knapsack problem with a cardinality constraint.

The empirical welfare maximization problem is therefore

$$\max_{d\in\{0,1\}^N} \sum_{i=1}^{N} \hat{\tau}_i d_i \quad \text{s.t.} \quad \sum_{i=1}^{N} c_i d_i \le C_N, \qquad \sum_{i=1}^{N} d_i \ge N_N^*.$$

This problem corresponds to a 0–1 knapsack with a minimum cardinality constraint. Let

$$B_N^*$$

denote the optimal value.

Greedy approach

↓

Rank&Cut Algorithm

A simple and natural allocation rule is:

$$\text{rank individuals by } \frac{\hat{\tau}_i}{c_i}$$

and select the largest feasible prefix.

However:

- the exact optimum generally does not correspond to a single ranking rule
- the problem is combinatorial

Question:

How good is the greedy ratio rule?

Near-optimality of two greedy algorithms

Problem

The optimal optimal allocation has a **knapsack-type structure**

Computational insight

The Linear Programming relaxation exhibits an $O(1)$ **integrality gap**
⇒ Asymptotically equivalent to the optimal discrete allocation

Algorithms

1) *Greedy–Lagrangian (GLC)*
Near-optimal in finite samples

2) *Rank-&-Cut (RC)*
Approximately optimal when coverage is non-binding or treatment costs are homogeneous. Performance deteriorates only when **cost heterogeneity** and **binding coverage** coexist

Rank-&-Cut algorithm

- Pick first the objects that give the most value for each unit of weight.
- Keep adding them until the bag is full.

Inputs

Estimated effects and costs $\{(\hat{\tau}_i, c_i)\}_{i=1}^N$, budget C, minimum coverage N^*.

Steps

1. Compute score $s_i = \hat{\tau}_i / c_i$.
2. Rank individuals by decreasing s_i.
3. For each k, compute
$$B_k = \sum_{j=1}^{k} \hat{\tau}_{(j)}, \qquad C_k = \sum_{j=1}^{k} c_{(j)}.$$
4. Keep k such that $k \geq N^*$ and $C_k \leq C$.
5. Select
$$k^* = \arg\max B_k \quad \text{(smallest if multiple).}$$
6. Treat top k^* ranked individuals.

Outcome

Return: $= B_{k^*}$, Total cost: $= C_{k^*}$.

Example

Rank-&-Cut
greedy algorithm

$$s_i = \frac{\hat{\tau}_i}{c_i}$$

i	$\hat{\tau}_i$	c_i	s_i
1	6	3	2.00
2	4	2	2.00
3	5	5	1.00
4	3	1	3.00
5	2	4	0.50

→ Rank of units

$[\,4,\ 1,\ 2,\ 3,\ 5\,]$ →

k	Treated units	B_k	C_k
1	4	3	1
2	4,1	9	4
3	4,1,2	13	6
4	4,1,2,3	18	11
5	4,1,2,3,5	20	15

- Budget: $C = 10$
- Minimum coverage: $N^* = 2$

Feasible choices

$k = 2, 3.$

Best choice

$k^* = 3$

Selected units

$\{4, 1, 2\}$

Stata package OPL

MODULE 1: BINARY TREATMENT

command	Description
make_cate	Estimation of the conditional average treatment effect (CATE)
opl_overlap	Assessing overlap between train and new data
opl_tb	Threshold-based optimal policy learning
opl_tb_c	Threshold-based policy learning at specific threshold values
opl_lc	Linear-combination optimal policy learning
opl_lc_c	Linear-combination policy learning at specific parameters' values
opl_dt	Decision-tree optimal policy learning
opl_dt_c	Decision-tree policy learning at specific splitting variables and threshold values
opl_budget	Optimal policy learning under budget constraint and minimum number of treated units

MODULE 2: BINARY COST-BENEFIT ANALYSIS

command	Description
opl_fb_cba	First-best cost-benefit analysis and policy-score frontier
opl_tb_cba	Welfare-maximizing cost-benefit analysis with threshold policies
opl_frontier_tb	Postestimation frontiers for threshold-based optimal policy learning
opl_policy_eval	Evaluation of a user-supplied binary treatment policy after **opl_tb_cba**

MODULE 3: MULTIVALUED TREATMENT

command	Description
opl_ma_fb	Optimal policy learning for multivalued treatment using first-best optimal policy and risk preferences
opl_ma_vf	Value-function estimation for multivalues optimal policy learning using RA, IPW, and DR estimators
opl_best_treat	Computing maximal reward and best treatment after **opl_ma_fb**
opl_plot_best	Plotting observed vs. maximal expected reward, and observed vs. optimal treatment

Stata **opl_budget** command

```
opl_budget tauvar costvar , budget(#) nmin(#) em0(#)
```

```
- TABLE OF RESULTS -
OPL WITH BUDGET CONSTRAIN
AND MINIMUM NUMBER OF TREATED
--------------------------------------------------
- General Info -
--------------------------------------------------
Number of obs                              267
Number of used obs                         267
Budget                                   80000
Minimum N. of requested treated             20
--------------------------------------------------
- Optimal Policy -
--------------------------------------------------
Parameter                                Value
--------------------------------------------------
Optimal number of treated                   33
Percentage of treated                       12
Total cost (TC)                     78950.0000
Average cost per treated (AC)        2392.4242
Total welfare (TW)                1239194.0710
Average welfare (AW)                 4641.1763
Average impact (ATET)                5126.4690
Total impact (TTET)                169173.4785
Policy net return (PNR)             90223.4785
PNR per treated (PNRT)               2734.0448
Policy variable created                pol_opt
--------------------------------------------------
Note: PNR = TTET-TC; PNRT = PNR/N1
--------------------------------------------------
```

Conclusion

Although the **exact optimal policy** under heterogeneous treatment costs and simultaneous budget and minimum coverage constraints is inherently **combinatorial**, a simple ***Rank-&-Cut*** **greedy allocation rule** based on **cost–effectiveness ratios** achieves good performance.

This result provides a theoretical justification for the use of **ranking-based allocation** procedures in large-scale policy learning applications, where computational simplicity and scalability are essential.

CONCLUSIONS

- Policy Learning: frontier of econometrics of prog eval
- Theory-driven and data-driven approaches can complement
- Machine Learning algorithms for estimating $\tau(X)$
- Budget and coverage constrains are relevant
- Generalization to various policy classes

[Submitted on 12 May 2026]

Optimal Policy Learning under Budget and Coverage Constraints

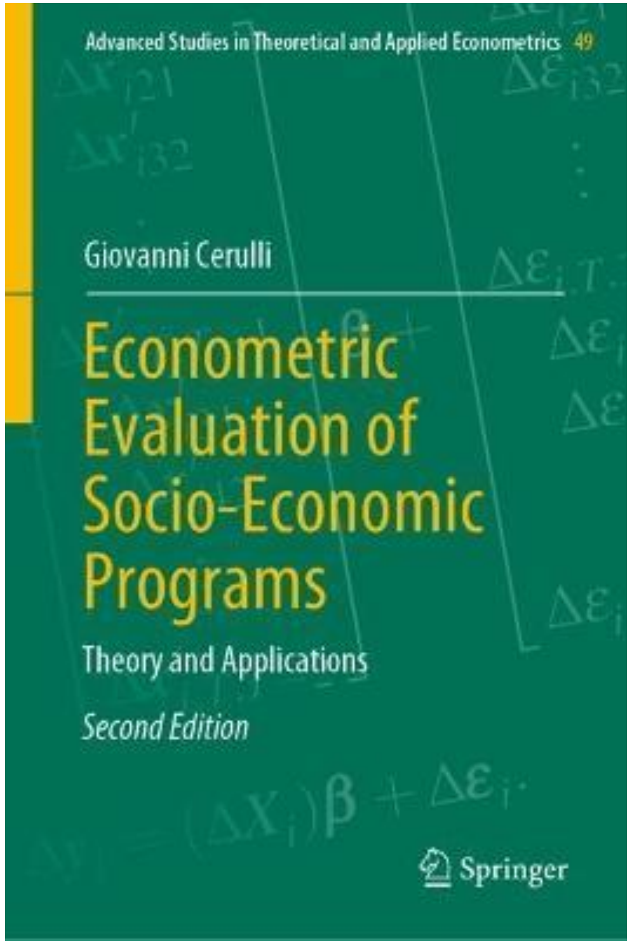
Giovanni Cerulli

We study optimal policy learning under combined budget and minimum coverage constraints. We show that the problem admits a knapsack-type structure and that the optimal policy can be characterized by an affine threshold rule involving both budget and coverage shadow prices. We establish that the linear programming relaxation of the combinatorial solution has an O(1) integrality gap, implying asymptotic equivalence with the optimal discrete allocation. Building on this result, we analyze two implementable approaches: a Greedy-Lagrangian (GLC) and a rank-and-cut (RC) algorithm. We show that the GLC closely approximates the optimal solution and achieves near-optimal performance in finite samples. By contrast, RC is approximately optimal whenever the coverage constraint is slack or costs are homogeneous, while misallocation arises only when cost heterogeneity interacts with a binding coverage constraint. Monte Carlo evidence supports these findings.

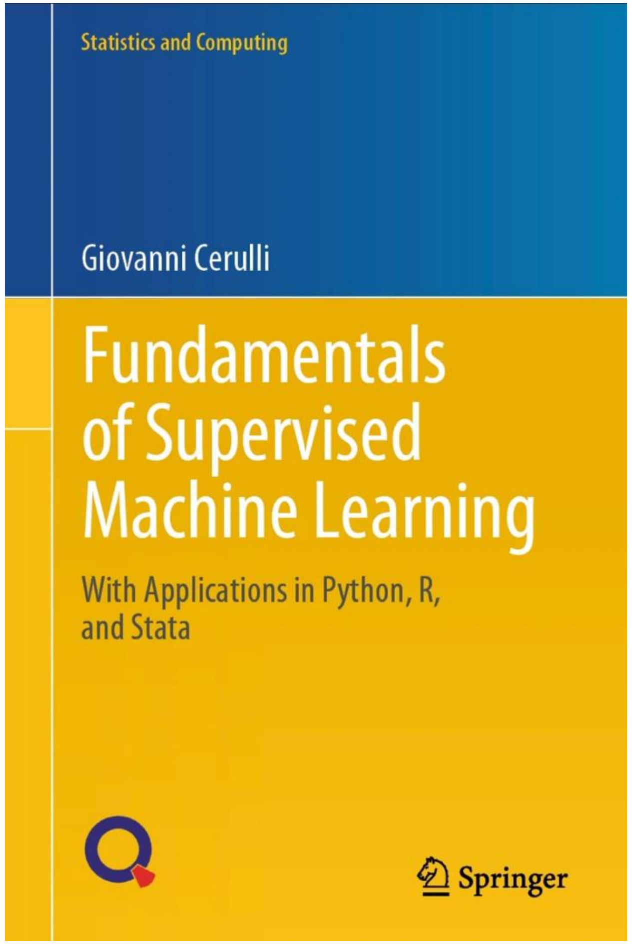
Subjects: **Machine Learning (stat.ML)**; Machine Learning (cs.LG)
Cite as: arXiv:2605.12235 **[stat.ML]**
(or arXiv:2605.12235v1 **[stat.ML]** for this version)
https://doi.org/10.48550/arXiv.2605.12235

Books for learning about Causal Inference and Machine Learning

https://link.springer.com/book/10.1007/978-3-662-65945-8



https://link.springer.com/book/10.1007/978-3-031-41337-7



Finanziato dall'Unione europea
NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA
Consiglio Nazionale delle Ricerche

THANK YOU!

fossr.dissemination@ircres.cnr.it

@fossrproject

fossr.eu

fossr-eu

@fossr

zenodo.org/communities/fossr