Making Subsidies Work: Rules vs. Discretion
Supplementary Materials

Federico Cingano [*] Filippo Palomba [†] Paolo Pinotti [‡] Enrico Rettore [§]

November 22, 2023

S1 Construction of sub-rankings of L488/92 applications

As explained in Section 2, the final ranking of L488/92 applicants mainly depends on three criteria in the first two calls for projects (*skin in the game*, *job creation*, *no waste*), plus two additional criteria in subsequent calls (*political discretion* and *environmental responsibility*). In addition, separate rankings were formed by (i) firm size, (ii) activity in the service sector, (iii) eligibility to receive EU funds, and (iv) EU objective area in which a firm operates. These four additional criteria entered the formation of the final ranking by either reserving part of the total budget for specific categories of firms (i-ii) or by making additional EU funds available for specific types of projects (iii-iv).

Firm size. Each region had to commit 50% of its L488/92 budget to small and medium enterprises (i.e., fewer than 250 employees, turnover under €50 million, or balance sheets below €43 million). Figure S1 provides one example from the second call, as published in the Official Journal. The projects are sorted in decreasing order according to the final score (in column L). Looking at funds allocation (column T) reveals that the projects ranked 90th and 92nd (ID 75995 and 7939) were declared eligible, while those ranked 88th (ex-aequo, ID 90634 and 38259) were not, despite their higher score. This is because the first two were submitted by a medium and a small firm, while the other two were submitted by large firms (see column N: "G" stands for large, "M" for

[*] Bank of Italy, Economic Research Department, e-mail: federico.cingano@bancaditalia.it.
[†] Princeton University, e-mail: fpalomba@princeton.edu.
[‡] Bocconi University, Social and Political Science Department and BAFFI-CAREFIN Center, CEPR, CESifo, e-mail: paolo.pinotti@unibocconi.it.
[§] University of Padua, Department of Economics and Management and FBK-IRVAPP, IZA, e-mail: enrico.rettore@unipd.it.

Fig. S1: Extract of the ranking published in the Official Journal.

A Posiz. in grad.	B Numero di prog.	C RAGIONE SOCIALE	D I1 Capitale proprio	E I2 Occupazione attività	F I3 Agevolazione richiesta	G I1N Capitale proprio	H I2N Occupazione attività	I I3N Agevolazione richiesta	L Somma indicatori normalizz.	M Sett. serv.	N Dim	O Ob	Q Cof	Q Esito finale	R Cod esc.	S Risor	T Agevolaz concessa L mil.
80	75299		07300000	00103263	11111111	07150427	04611723	-02334272	094278780		M	1	S	A		N	994 62
81	90303		07300000	00101891	11111111	0,7150427	04451254	-02334272	0,92674090		M	1	S	A		N	326,55
82	15165		07356807	00068273	11764706	0,7428728	00519287	01297118	0,92451330		M	1	S	A		N	2.545,50
83	8219		06347110	0,0103466	1,1904762	02482163	04635466	0,2075272	091929010		P	1	S	A		N	200,79
84	38337		07450000	0,0062719	1,1764706	07685286	-00130310	0,1297118	0,90520940		P	1	S	A		N	670,38
85	45619		07480000	0,0053331	1,1904762	08032257	-01228332	0,2075272	0,88791970		P	1	S	A		N	1 358,73
86	64729		07037122	0,0041324	1,2500000	05862572	-02632673	0,5382429	0,86123280		P	1	S	A		N	8 191,05
87	75998		0,5503532	0,0105457	1,2500000	0,1650575	0,4868334	0,5382429	0,86001880		M	1	S	A		N	754,89
88	90634		0,8000000	0,0000000	1,2500000	1,0579768	-0,7465935	0,5382429	0,84962620		G	1	S		4		0,00
88	38259		0,8000000	00000000	1,2500000	1,0579768	-0,7465935	0,5382429	0,84962620		G	1	S		4		0,00
90	75995		0,5698718	00213675	1,0000000	-0,0694347	1,7525530	-0,8507631	0,83235520		M	1	S	A		N	833,16
91	50826		0,5441235	00063659	1,3333333	-0,1955771	-0,0020367	1,0012447	0,80363090		G	1	S		4		0,00
92	7939		0,7540000	00029062	1,2195122	0,8326201	-0,4066839	0,3688519	0,79478810		P	1	S	P	1	N	2.390,31
93	1707		0,5388774	00095278	1,2658228	-0,2212781	0,3677796	0,6261547	0,77265620		M	1	S		1		0,00
94	1681		0,5152374	0,0140829	1,1904762	-0,3370918	0,9005449	0,2075272	0,77098030		P	1	S		1		0,00

Notes: This is a snapshot from a ranking of the second call published in the Official Journal. The first column (A) shows the position in the ranking, the second (B) the ID of the project, and the third (C) the company name, which we omit. Then there are 7 columns (D-L) that contain data on the raw sub-indexes, normalized sub-indexes and aggregated index presented in Section 2. The last columns indicate: whether the firm is active in the service sector (M), the size of the firm (N), the EU Objective area where the firm operates (O), the firm's eligibility to receive EU funding (P), the outcome of the application (Q), the reason for non-selection (R), the source of funding received (S), the amount of funding (T). Source: *Gazzetta Ufficiale, SG 174 of 28.07.1997, SO 151, p.68*.

Medium and "P" for small).

Had these projects been selected for funding, the 50% quota reserved for small and medium-sized firms would have been violated.

Activity in the service sector. Firms operating in the service sector could receive at most 5% of the regional budget. Therefore, a project could be selected to receive funds even if it had a lower score than another project submitted by a company operating in the services sector. This case is illustrated in Figure S2.

As before, projects are sorted by the score received (column L). However, the project in 7th place with ID 67085-11 was funded even though it had a lower score than the project in 6th place with ID 20788-11. This is because the latter was submitted by a service provider and the 5% upper bound had been reached (see column M, where "S" stands for service provider).

Eligibility for EU funds. Projects meeting certain criteria – in terms of location and type of activities, duration of investment, and the amount of eligible expenses – were eligible for co-funding from the European Regional Development Funds (ERDF). These projects might be selected over higher-ranked projects that were eligible for national funds only.

This case is portrayed in Figure S3. The projects ranked 171st and 172nd (IDs 40416 and 12997)

Fig. S2: Extract of the ranking published in the Official Journal.

LEGGE 488/92 - BANDO DEL 2000 (8°) DEL SETTORE INDUSTRIA - GRADUATORIA ORDINARIA DELLA REGIONE LIGURIA **Allegato 2/10**

NUMERO INIZIATIVE IN GRADUATORIA 113

	Indicatore 1	Indicatore 2	Indicatore 3	Indicatore 4	Indicatore 5
MEDIE	0,5582615451	0,0035209584	1,1729751788	19,8230088496	6,8716814159
DEVIAZIONI STANDARD	0,2193934539	0,0055085503	0,2925841006	5,3838881471	3,6245558601

A	B	C	D	E	F	G	H	I	L	M	N	O	P	Q	R	S	T
				1	2	3	4	5									
Posiz. in grad.	Numero di progetto	Ragione Sociale	Prov.	Capitale proprio	Occupazione attivata	Agevolazione richiesta	Indicatore Regionale	Indicatore Ambientale	Somma indicatori normalizzati	Sett. Serv.	Dimen-sione	Ob.	Cofin	Esito con-clusivo	Cod. escl.	Agevolaz. Concedibile (LM)	Agevolaz. Concedibile (Euro)
1	52111 - 11		GE	36,4458900	0,045	1,0526 %	20	10,00000	07,0763729	S	P	2	SI	A		85,52	44.167
2	66443 - 11		GE	51,6017900	0,001	2,9412 %	20	10,00000	06,3481123		M	2	SI	A		1.299,06	670.908
3	68960 - 11		GE	40,6230200	0,001	2,8571 %	20	7,0000000	04,6858531	S	G	2		P	2	638,88	329.954
4	67097 - 11		SP	89,6628000	0,024	1,5385 %	20	0,0000000	04,5825445		G	2		A		99,74	51.511
5	40226 - 11		GE	70,4517800	0,010	1,1111 %	30	10,00000	04,4082812	S	P	2	SI	N	2	-	-
6	20788 - 11		GE	30,7919600	0,002	2,0000 %	30	10,00000	04,1099450	S	P	2	SI	N	2	-	-
7	67085 - 11		GE	85,7800000	0,002	1,2658 %	30	10,00000	04,1029045		P	2	SI	A		871,86	450.278
8	20903 - 11		GE	84,7000000	0,002	1,1765 %	30	10,00000	03,8635997	S	P	2	SI	N	2	-	-
9	20709 - 11		GE	83,5347900	0,003	1,1364 %	30	10,00000	03,8450711		P	2	SI	A		254,94	131.665
10	20649 - 11		SV	67,7305500	0,007	1,1111 %	30	10,00000	03,6940373		P	2	SI	A		477,15	246.427

Notes: This is a snapshot from a ranking of the eighth call published in the Official Journal. The first column (A) shows the position in the ranking, the second (B) the ID of the project, the third (C) the company name, which we omit, and the fourth (D) the province where the company was located. Then there are 6 columns (E-L) that contain data on the five normalized sub-indexes presented in Section 2, as well as the overall index. The last columns indicate whether the firm is active in the service sector (M), the size of the firm (N), the EU Objective area where the firm operates (O), the firm's eligibility to receive EU funding (P), the outcome of the application (Q), the reason for non-selection (R), the amount of funding received in millions Italian Lire (S), the same amount in euros (T). Source: *Gazzetta Ufficiale, SG 186 of 11.08.2001, SO 208, p.29*.

Fig. S3: Extract of the ranking published in the Official Journal.

A	B	C	D	E	F	G	H	I	L	M	N	O	P	Q	R	S	T
			1	2	3	4	5										
Posiz. in grad.	Numero di prog.	Ragione Sociale	Capitale proprio	Occupazione attivata	Agevolazione richiesta	Ind. reg.	Ind. amb.	Somma indicatori normalizz.	Sett. serv.	Dim.	Ob.	Cof.	Esito finale	Cod. Escl.	Risorse	Agevolaz. concedibile L. mil.	Agevolaz. concedibile Euro
163	12380		0,6586178	0,0000000	2,0000000	0	4	0,52555960		G	2		N	1		0,00	0
164	15042		0,5190234	0,0026975	1,1111111	1	6	0,51754010		P	2	S	A		C	283,53	146.734
165	3955		0,4065680	0,0020112	1,2500000	1	6	0,48736070		P	5		N	1		0,00	0
166	5814		0,3706947	0,0061005	1,2500000	1	5	0,48524370		P	2	S	A		C	146,60	75.869
167	15338		0,9300000	0,0000000	1,4285714	0	6	0,44496170		G	2		N	1		0,00	0
168	40967		0,1657733	0,0033984	1,1764706	1	8	0,44400580		M	2	S	A		C	835,28	432.279
169	16944		0,3165459	0,0022000	1,0526316	1	8	0,39952000		P	5	S	A		C	325,35	168.377
170	40418		0,2326934	0,0058173	1,2500000	1	6	0,38718090		P	2	S	A		C	94,71	49.015
171	40416		0,1642957	0,0005917	1,6666667	1	5	0,38703570		P	2		N	1		0,00	0
172	12997		0,2303593	0,0012955	1,1764706	1	8	0,38296540		P	2	S	A		C	233,42	120.801

Notes: This is a snapshot taken from one ranking of the eight calls published in the Official Journal. The first column (A) shows the position in the ranking, the second one (B) the ID of the project, and the third one (C) the company name, which we omit. Then, there are 6 columns (D-I) containing data on the five normalized sub-indexes presented in Section 2, and the aggregate index. The last columns report: whether the firm operates in the services sector (L), the dimension of the firm (M), the EU Objective area the firm operates in (N), the firm's eligibility for EU funding (O), the outcome of the application (P), the reason for not being selected (Q), the source of funds received (R), the amount of funds received expressed in millions of Italian Lire (S), the same amount expressed in Euro (T). Source: *Gazzetta Ufficiale, SG 54 of 06.03.1999 54, SO 47, p.28*.

were both presented by small firms. However, only the second, lower scoring project received funding. This is because it had access to EU funds while the first one did not, and the national funds were already exhausted (eligible projects are marked with an "S" in column O; the "C" in column R indicates that the funds received were co-financed, whilst "N" denotes national funding).

EU Objective Area. Even projects eligible for EU funding could be subject to constraints on the type of ERDF program. In particular, firms in Northern and Central regions could tap either Objective 2 funds (if located in areas in industrial decline) or Objective 5b funds (if in disadvantaged rural areas), and the budget available for either source of funds would typically be different. Figure S4 shows an example in which all projects submitted by firms operating in an Objective 5b area were not selected due to exhaustion of the corresponding funds, while all Objective 2 projects were selected, even if such projects received a lower score.

Fig. S4: Extract of the ranking published in the Official Journal.

			INDICATORI NON NORMALIZZATI				INDICATORI NORMALIZZATI										
A Posiz in grad	B Numero di prog	C RAGIONE SOCIALE	D I1 Capitale proprio	E I2 Occupazione attività	F I3 Agevolazione richiesta	G I1N Capitale proprio	H I2N Occupazione attività	I I3N Agevolazione richiesta	L Somma indicatori normalizz	M Sett serv	N Dim	O Ob	Q Cof	Q Esito finale	R Cod esc	S Risor	T Agevolaz concessa L. mil
129	25360		0,8505263	0,0000000	1,0000000	0,7008209	-0,4351545	-0,7296810	-0,46401460		P	5B	S		1		0,00
130	65825		0,7917975	0,0010292	1,0526316	0,5186830	-0,3709085	-0,6155161	-0,46774160		M	2	S	A		C	363,81
131	42605		0,7877161	0,0029590	1,0000000	0,5060252	-0,2504442	-0,7296810	-0,47410000		P	2	S	A		C	691,05
132	22406		0,8000000	0,0022222	1,0000000	0,5441217	-0,2964376	-0,7296810	-0,48199690		P	2	S	A		C	295,77
133	34725		0,3582400	0,0000000	1,6666667	-0,8259257	-0,4351545	0,7164069	-0,54467330		P	2	S	A		C	123,24
134	35710		0,3737235	0,0031519	1,5384615	-0,7779062	-0,2384027	0,4383129	-0,57799600		M	2	S	A		C	1270,80
135	8717		0,1890212	0,0155360	1,4285714	-1,3507306	0,5346528	0,1999468	-0,61613100		P	2	S	A		C	88,83
136	8686		0,8000000	0,0000000	1,0000000	0,5441217	-0,4351545	-0,7296810	-0,62071380		P	2	S	A		C	381,00
137	45047		0,7065861	0,0000000	1,1235955	0,2544136	-0,4351545	-0,4615861	-0,64232700		M	2	S	A		C	167,85
138	4708		0,5756229	0,0043018	1,1764706	-0,1517477	-0,1666222	-0,3468930	-0,66526290		M	2	S	A		C	584,43
139	31930		0,4007181	0,0036424	1,4285714	-0,6941867	-0,2077841	0,1999468	-0,70202400		P	5B	S		1		0,00
139	31931		0,4007181	0,0036424	1,4285714	-0,6941867	-0,2077841	0,1999468	-0,70202400		P	5B	S		1		0,00
141	41867		0,5150862	0,0121228	1,0000000	-0,3394925	0,3215899	-0,7296810	-0,74758360		P	5B	S		1		0,00
141	41870		0,5150862	0,0121228	1,0000000	-0,3394925	0,3215899	-0,7296810	-0,74758360		P	5B	S		1		0,00

Notes: This is a snapshot from a ranking of the first call published in the Official Journal. The first column (A) shows the position in the ranking, the second (B) the ID of the project, and the third (C) the company name, which we omit. Then there are 7 columns (D-L) that contain data on the raw sub-indexes, normalized sub-indexes and aggregated index presented in Section 2. The last columns indicate: whether the firm is active in the service sector (M), the size of the firm (N), the EU Objective area where the firm operates (O), the firm's eligibility to receive EU funding (P), the outcome of the application (Q), the reason for non-selection (R), the source of funding received (S), the amount of funding (T). Source: *Gazzetta Ufficiale, SG 288 of 09.12.1996, SO 215, p.34*.

Cell construction. A ranking is defined by six elements:

(1) *call* – in our final sample, we consider the following calls: $1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20, 31, 32, 33$

(2) *region* – Italy has 20 regions

(3) *firm size* – we create two different rankings along this dimension, one for small-medium enterprises and one for large firms

(4) *service sector* – there is one ranking for service providers and another one for firms that are not active in this sector

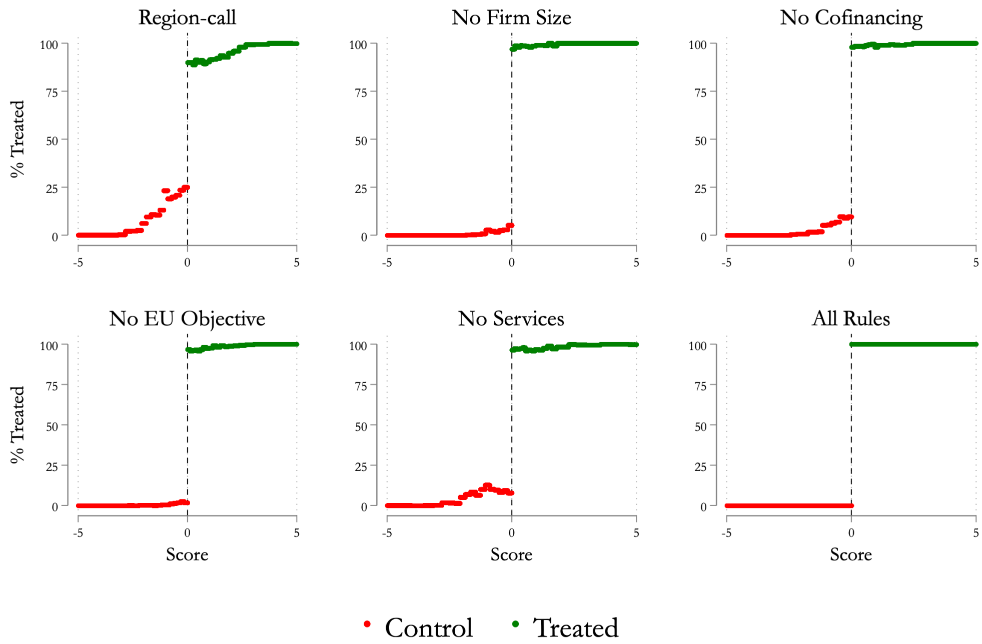
(5) *eligibility for EU funding* – there is one ranking for eligible firms and another for those not eligible

(6) *EU Objective* – there are four ranking types: one for Objective 1, one for Objective 2, one for Objective 5b, and one for the areas that are not part of the program and are considered "Out of Objective"

We define a *cell* as the interaction of elements (1) to (6). For example, a cell in our specification could be: projects submitted during the 2nd call in the Tuscany region by small and medium-sized enterprises not active in the service sector, eligible for EU funds, and operating in an Objective 2 area.

Considering only elements (1) and (2), as in previous evaluations of L488/92, introduces significant measurement error in treatment assignment near the cutoff (top-left panel in Figure S5). When we consider the additional rules that determine assignment to treatment, we retrieve a sharp discontinuity at the pooled cutoff (lower right panel in Figure S5). The other panels in Figure S5 show that each and any of the four dimensions described above (in addition to call and region) is necessary to recover the sharp discontinuity in treatment assignment.

Fig. S5: Measurement error in treatment assignment due to errors in the construction of rankings



S2 Total and direct effects when applicants can re-apply

The outcome of applications submitted in year t may affect the probability of re-applying for funds – and, therefore, obtaining the subsidy – in later years, say at $t + \Delta$. In this case, the dynamic treatment effects on outcomes from $t + \Delta$ onwards would reflect both the direct effect of the subsidy obtained at time t, and the indirect effect through a different probability of obtaining subsidies in subsequent years. The sign of the indirect effect is a priori unclear. On the one hand, firms obtaining funds in year t may not have additional (promising) projects to submit in year $t + \Delta$, or they may be constrained in the amount of own resources that could be invested. In this case our estimates provide a lower bound for the direct effect of obtaining the subsidy at time t. On the other hand, obtaining funds in year t may improve the chances of succeeding in year $t + \Delta$, due for example to increased availability of resources or reputation effects, in which case we would be over-estimating the direct effects of the subsidy.

In practice, we sign the (indirect) effect of obtaining a subsidy on the probability of obtaining additional funds in the following years using our baseline RDD specification (3). Figure S6 shows that applicants scoring just above the cutoff in year t have a 23 percentage point lower probability of re-applying for funds in year $t + 1$, and a 16 percentage point lower probability of actually obtaining such funds. These differences decrease markedly in year $t + 2$ to eventually disappear from $t + 3$ onward. Therefore, the estimated coefficients in Table 3 and Figure 5 under-estimate the direct, dynamic treatment effects of the subsidy.

This is not an issue for the internal validity of our estimates, as receiving less subsidies between t and $t + \Delta$ is itself a causal effect of the subsidy received at time t. In terms of external validity, however, we may want to distinguish between direct and indirect effects, as the latter would not apply in the context of one-off interventions. We thus extend the estimating equation (3) to allow for dependence of firm outcomes on subsidies received in *all* previous calls. We illustrate our procedure with reference to a two-period case. Let the model for the call in period $t = 1$ be the standard one:

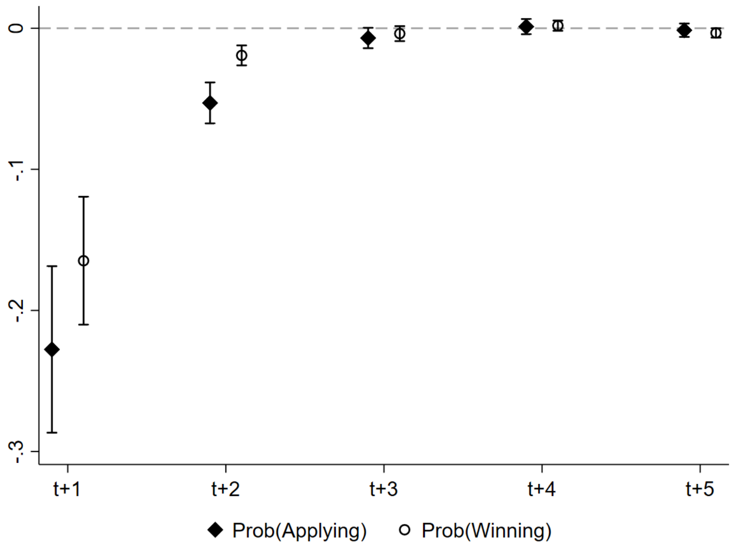
$$Y_1 = \tau_1 D_1 + \gamma_1 S_1 + \delta_1 D_1 \cdot S_1 + \varepsilon_1 \tag{1}$$

where all variables are defined as in equation (3), and the sub-index "1" denotes the period.[1] With repeated interventions, the causal effect of the subsidy received in period $t = 1$ on the outcome in period $t = 2$ would read as

$$Y_2 = \tau_2 D_2 + \gamma_2 S_1 + \delta_2 D_1 \cdot S_1 + \tilde{\tau}_2 D_1 + \varepsilon_2,$$

[1]We consider the case of a linear regression in S to simplify notation (i.e., $k = 1$ in equation 3), but it is immediate to allow for higher-order polynomials in S.

Fig. S6: Direct and indirect effects for re-applicants



Notes: The graph shows the estimated effect of obtaining the L488/92 subsidy in year t on the probability of re-applying for the same subsidy (black markers) and obtaining it (grey marker) in subsequent years, as estimated from the RD regression 3. 95% confidence intervals are also shown in the graph.

where we explicitly take into account that in period 2 some units among those applying for the subsidy in $t = 1$ might apply to the new call and possibly receive the subsidy in $t = 2$, which would have an effect on Y_2 as large as τ_2. If we knew τ_2, the following regression would be suitable to properly estimate $\tilde{\tau}_2$ (i.e., the causal effect of D_1 on the outcome in $t = 2$):

$$Y_2 - \tau_2 D_2 = \gamma_2 S_1 + \delta_2 D_1 \cdot S_1 + \tilde{\tau}_2 D_1 + \varepsilon_2. \tag{2}$$

An estimate of τ_2 could be recovered from a regression analogous to (1), run on firms participating in the call issued in period $t = 2$ but not in the previous call.

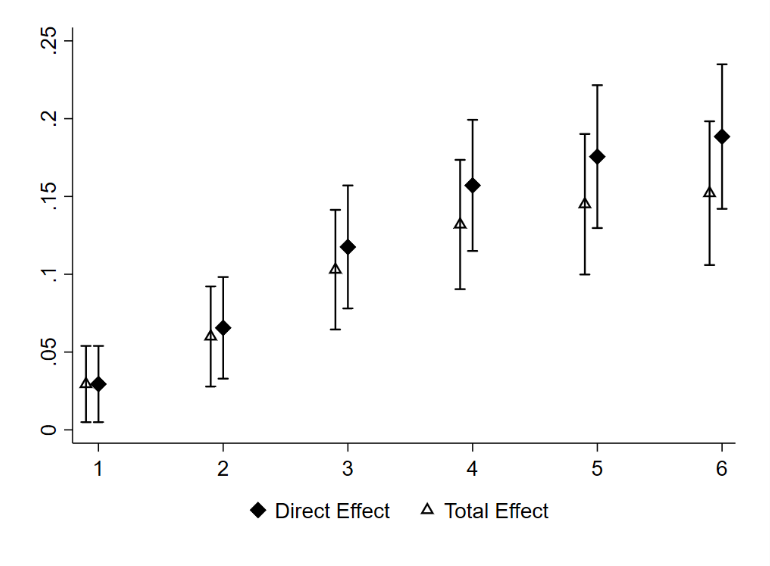
In practice, with calls issued across several subsequent years, we estimate (1) allowing for year-specific coefficients τ_1^t ($t = 1996, ..., 2006$) in a sample including only firms applying for the first time. Year-specific contemporaneous coefficients are then used to "net" outcomes of firms applying in two consecutive years: $\tilde{Y}_2 = Y_2 - \tau_1^t D_2$.[2] Finally, the one-year-ahead direct effect of the subsidy $\tilde{\tau}_2$ is obtained by RDD using $\tilde{Y}_2$ on the left-hand-side of equation (2). The procedure is then iterated to estimate the direct effects of the policy at further horizons.

Figure S7 compares the total effect of the subsidy received at time t on employment growth at different time horizons, as reported also in Table 3 and Figure 5, with the direct effect obtained by subtracting the effect of subsequent subsides, estimated following the procedure described

[2]For example, the outcomes of a firm applying for the first time in 2001 and then also in 2002 would be Y_{2001} and $\tilde{Y}_{2002} = Y_{2002} - \tau_1^{2002} D_2$

above. As expected, in light of the evidence in Figure S6, the direct effect is larger than the total effect, as the latter also includes the indirect, negative effect going through a lower probability of re-applying for subsidies after obtaining it. However, the difference between direct and total effects remains small.

Fig. S7: Total and direct effects for re-applicants



Notes: The graph compares the total effect of obtaining a subsidy, as estimated in Table 3 and Figure 5 (second graph), with the direct effect obtained by subtracting the contemporaneous effect of any subsidy obtained in subsequent calls, as detailed in equations (1) and (2).

S3 Data-driven selection of covariates

We implement a data-driven algorithm that searches for a vector of covariates satisfying the CIA condition in the spirit of Imbens and Rubin (2015). Formally, assume that we have a set of k covariates $\mathcal{C}$, which is the union of two disjoint sets:

- a set $\mathcal{C}_1 \subset \mathcal{C}$ made up of $k_1 < k$ variables which must be included in the CIA regressions (7), but are not sufficient to make the running variable ignorable. These variables may be justified by some economic theory and, in principle, it could be that $\mathcal{C}_1 = \varnothing$.
- a set $\mathcal{C}_2 \subseteq \mathcal{C}$ made up of $k_2 \leq k$ *candidate* variables which could be included in the CIA regressions (7) with the only purpose of making the running variable ignorable.

The algorithm searches for a set $\tilde{\mathcal{C}} \subseteq \mathcal{C}_2$ such that $\tilde{\mathcal{C}} \cup \mathcal{C}_1$ makes the running variable ignorable.

Algorithm

1. Run the following set of regressions for $j = 1, \ldots, k_2$,

$$Y = \sum_{\ell=1}^{p} \gamma_\ell^0 S^\ell + \mathbf{z}'\tau^0 + w_j\mu_j^0 + FE_c^0 + \nu^0, \quad \text{if} \quad -h \leq S < 0,$$
$$Y = \sum_{\ell=1}^{q} \gamma_\ell^1 S^\ell + \mathbf{z}'\tau^1 + w_j\mu_j^1 + FE_c^1 + \nu^1, \quad \text{if} \quad 0 \leq S \leq h, \tag{3}$$

   where $\mathbf{z}$ is the vector of k_1 covariates that are always included; w_j is the j-th candidate covariate; and the other terms are defined as in equations 3 and (7), but allowing for different parameters on the two sides of the cutoff.

2. For each regression run the F-test for the null hypothesis that the CIA holds (separately) on each side of the cutoff

$$H_0^{(L)} : \gamma_1^0 = \cdots = \gamma_p^0 = 0 \quad \text{and} \quad H_0^{(R)} : \gamma_1^1 = \cdots = \gamma_q^1 = 0.$$

   and store the F-tests $F^{j,L}$ and $F^{j,R}$.

3. Select the two variables associated with the smallest F-statistics in the two sets $\mathcal{F}^L = \{F^{1,L}, F^{2,L}, \ldots, F^{k_2,L}\}$ and $\mathcal{F}^R = \{F^{1,R}, F^{2,R}, \ldots, F^{k_2,R}\}$. Notice that nothing prevents the variable with the smallest F-statistic on the left of the cutoff to differ from one on the right of the cutoff.

4. Add these two variables to the regressions in (3) and repeat steps 1-3 for the other candidate covariates.

5. Repeat step 4 until one of the following stopping criteria is reached:

   - the null hypothesis that the running variable is not significantly different from 0 cannot be rejected at the α% level
   - all the covariates in $\tilde{\mathcal{C}}$ have been included in the (3)

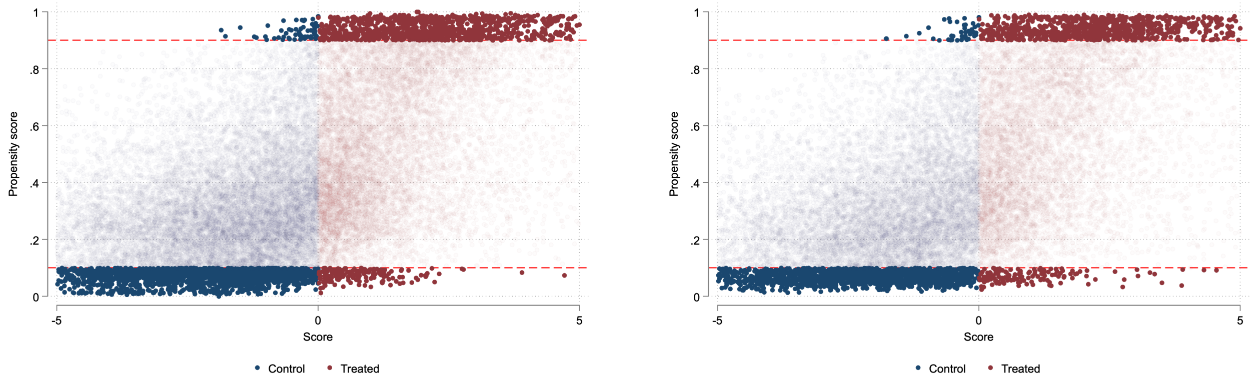
The basic idea behind the algorithm is to implement a *greedy approach*. An approach is greedy when it is myopic, in the sense that the best variable is selected at each particular step, rather than looking ahead and picking a variable that will lead to a larger reduction in the loss function in some future step. This is done to avoid testing all the possible combinations of the elements of $\mathcal{C}_2$.[3]

[3]This exercise would soon become intractable from a computational point of view as it involves estimating $\sum_{i=1}^{k_2}$

S4 Sensitivity to trimming the sample on the propensity score

The procedure by Angrist and Rokkanen (2015) for extrapolating treatment effects away from the RDD cutoff leverages common support in the propensity score between treated and untreated units, which we test in Figure 6 (right graph). Figure S8 below provides additional evidence of common support over the joint distribution of the running variable S and the (estimated) propensity score $\hat{e}(X)$, including for extreme values of the latter – below 0.1 and above 0.9. In any event, Figure S9 shows that results are unaffected when eliminating observations with propensity score outside $[0.1, 0.9]$.[4]

Fig. S8: $\hat{e}(X)$ as a function of the running variable S.



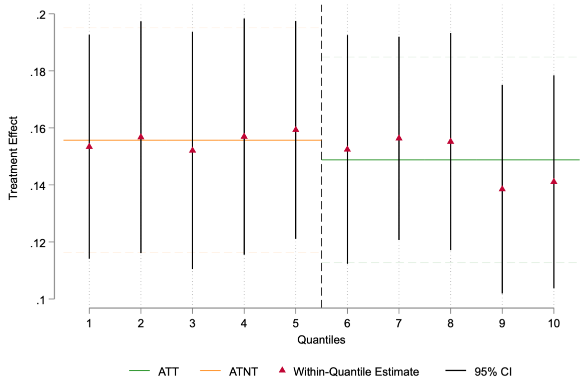
(i) using X for cumulated employment growth **(ii) using X for cumulated investment**

Since we are particularly interested in how treatment effects vary across projects selected on rules vs. discretion (Section 6.3), Figure S10 plots the fraction of units with an estimated propensity score above outside $[0.1, 0.9]$ across quintiles of SD and SR, by treatment arm. The distribution of such observations is quite sparse, and excluding them from the sample does not affect the results of our heterogeneity analysis, see Figure S11.

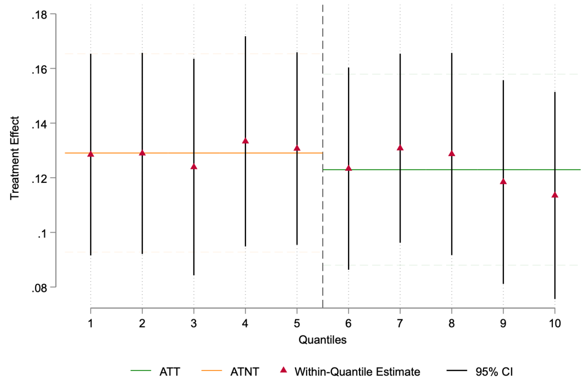
$\binom{k_2}{i}$ different regressions. To quantify this issue, with 10 covariates, the number of different combinations to be tested for is 1023. This case is still tractable. However, adding just 10 other covariates drives the number of combinations over 1 million.

[4]Crump et al. (2009) recommend discarding observations with propensity scores outside the range $[\alpha, 1-\alpha]$, where α is defined according to an optimal selection criterion, since such observations are often associated with unreliably large or small estimated treatment effects. In our case, the optimal threshold computed according to Theorem 1 in Crump et al. (2009) equals $\alpha^* = 0.10$. Indeed, Crump et al. (2009) show that, for a wide range of applications, the rule of thumb $\alpha = 0.1$ provides a good approximation of the optimal criterion.

Fig. S9: Treatment effects within quantiles of the running variable S.

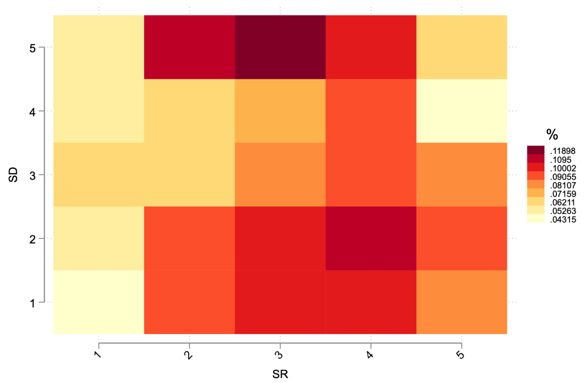


(i) baseline

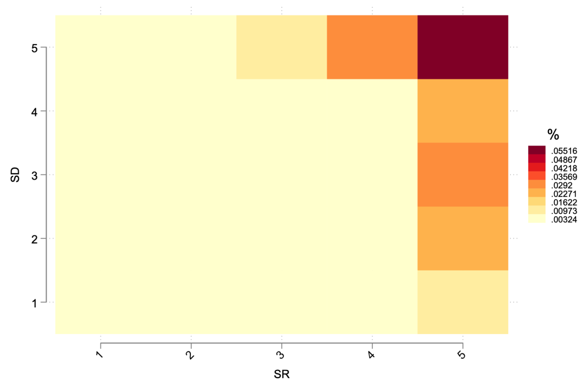


(ii) $\hat{e}(X_i) \in [0.1, 0.9]$

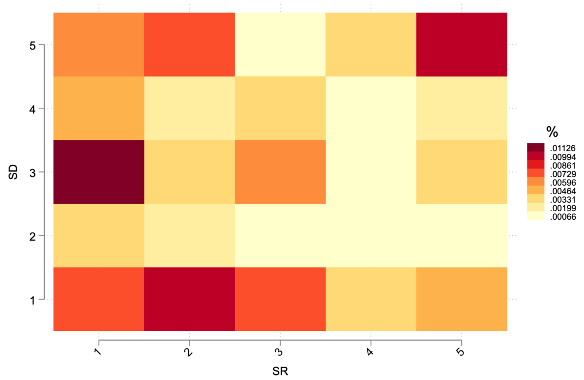
Fig. S10: $\hat{e}(X)$ as a function of the quantiles of SR and SD.



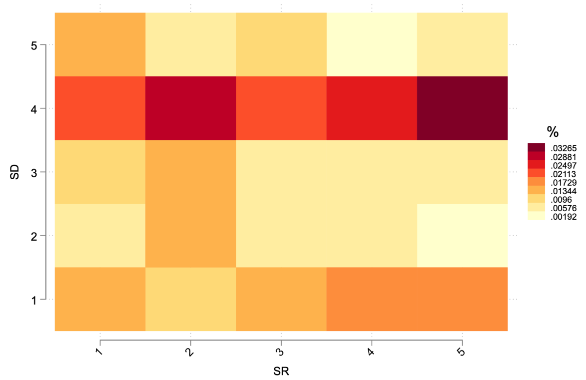
(i) fraction of control units with $\hat{e}(X_i) < \alpha$



(ii) fraction of treated units with $\hat{e}(X_i) < \alpha$

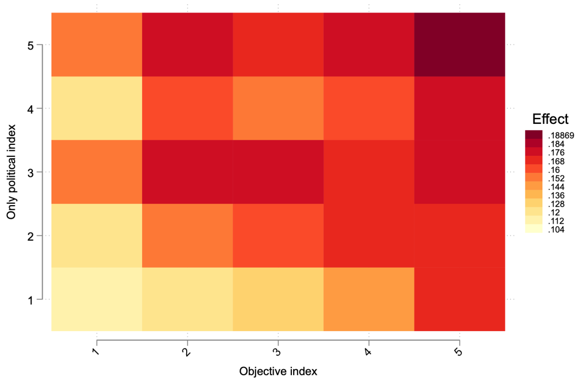


(iii) fraction of control units with $\hat{e}(X_i) > 1 - \alpha$

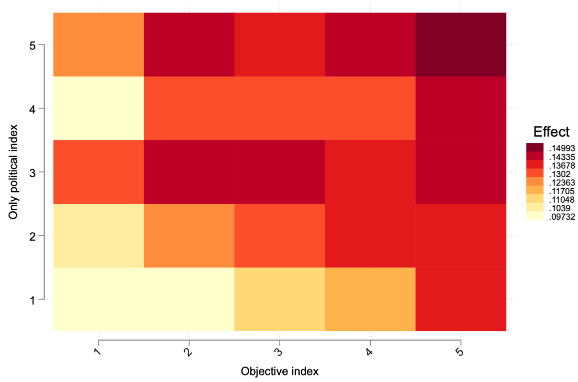


(iv) fraction of treated units with $\hat{e}(X_i) > 1 - \alpha$

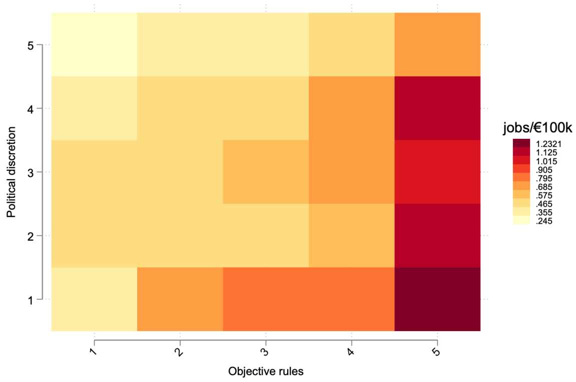
Fig. S11: Treatment effects and new jobs created per €100,000, rules v. discretion.



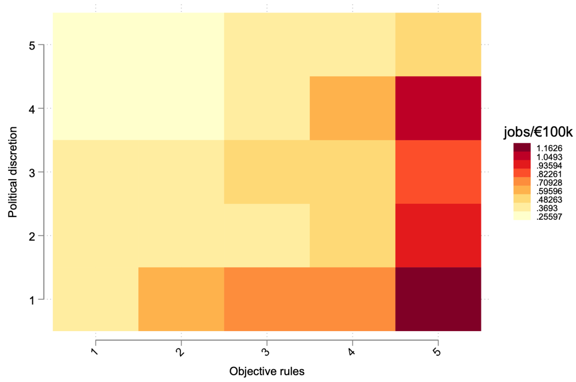
(i) baseline



(ii) $\hat{e}(X_i) \in [0.1, 0.9]$



(iii) baseline



(iv) $\hat{e}(X_i) \in [0.1, 0.9]$

S5 Additional Empirical Evidence on the Politicians' response function

To measure the – possibly non-linear – degree of dependence between SD and SR, we evaluate their longitudinal rank correlation ρ separately in each point of the support of Z – the triple municipality-industry-type of project over which SD is assigned. We then use such statistic to test the null hypothesis of independence between SR and SD. More precisely, for each value of Z, we consider the values of SD and SR over time and compute the Spearman's rank correlation of those values, that is

$$\rho(SD, SR) = \frac{\text{Cov}(R(SD), R(SR))}{\sigma_{R(SD)}\sigma_{R(SR)}},$$

where $R(\cdot)$ is a function assigning to each element in a vector its rank and σ_X denotes the standard deviation of X. The Spearman's correlation corresponds to the classic Pearson's correlation between the ranks of SD and SR. By conditioning on Z and exploiting only longitudinal variation, finding a non zero rank correlation would be consistent with the hypothesis that the regional authorities set the value of SD based on their expectations on SR.

Table S1 breaks down the total number of points in the support of Z we end up with – 3520 - by the number of time periods we observe each of them, e.g., 2455 points in the support of Z are observed in two different calls, 770 in three, and so on. Let $\widehat{F}(\rho|n), n = 2, \ldots, 7$, be the empirical distribution of the longitudinal rank correlation for those Z-types observed in n periods and let $w(n)$ be the empirical relative frequency (column (3) of Table S1). We recover the unconditional empirical distribution $\widehat{F}(\rho)$ as the weighted average of the conditional distributions, i.e.

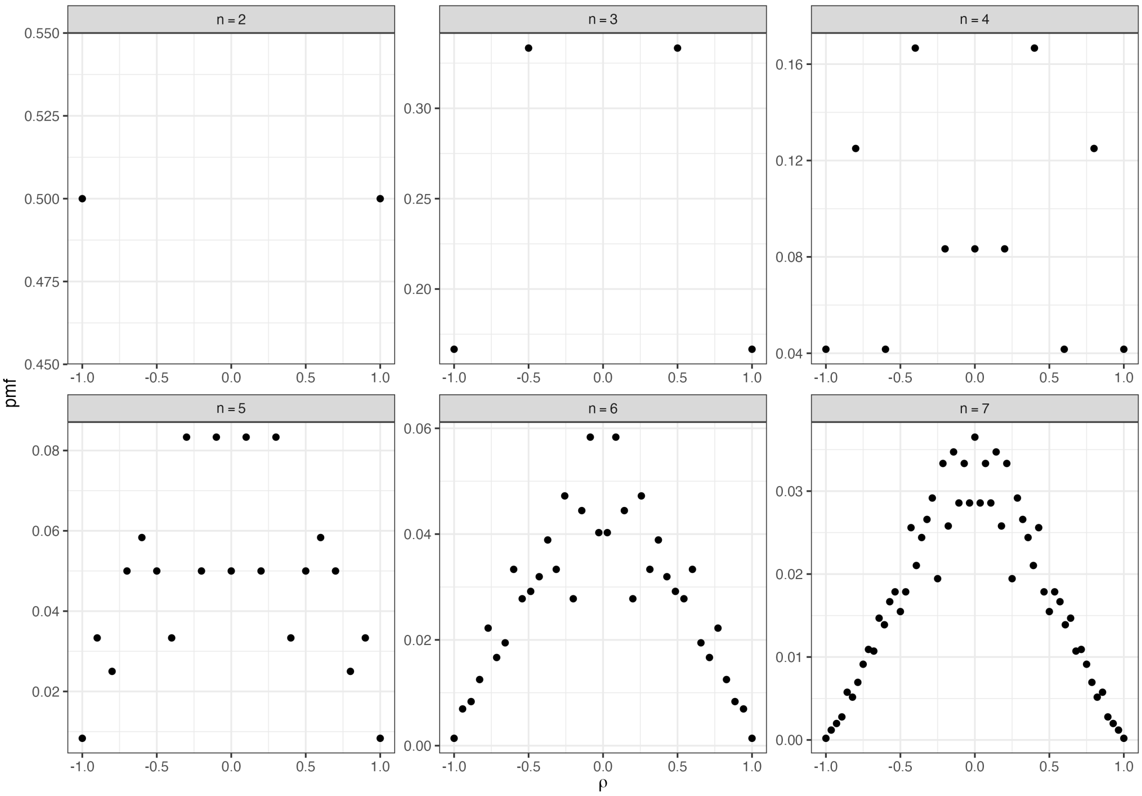
$$\widehat{F}(\rho) = \sum_{n=2}^{7} \widehat{F}(\rho|n)w(n).$$

To test whether $\widehat{F}(\rho)$ is consistent with the null hypothesis of zero rank correlation, we also derive the exact distribution $F(\rho|n)$ under the null hypothesis for each value of n. To do so, for a given n, we take $x, y \in \mathbb{N}_+^n$, compute $\rho(x, \pi(y))$ for all the possible permutations π, and count the number of times we observe a particular value for ρ. For example, we get that $\mathbb{P}[\rho = 1 \mid n = 2] = \mathbb{P}[\rho = -1 \mid n = 2] = 0.5$ and $\mathbb{P}[\rho = -1 \mid n = 3] = \mathbb{P}[\rho = 1 \mid n = 3] = 1/6, \mathbb{P}[\rho = -1/2 \mid n = 3] = \mathbb{P}[\rho = 1/2 \mid n = 3] = 1/3$. Figure S12 displays $F(\rho \mid n)$ for different values of n. Again, we get the overall distribution of ρ under the null hypothesis by taking the weighted average of the distributions conditional on n.

Table S1: Distribution of Z-types by number of periods of observation and politicians' expectations.

perfect foresight			adaptive expectations		
n	#	%	n	#	%
2	2455	69.74	2	886	68.90
3	770	21.79	3	297	23.09
4	236	6.68	4	94	7.31
5	56	1.62	5	7	0.54
6	4	0.11	6	2	0.16
7	2	0.06	7	0	0
Total	3520	100.00	Total	1286	100.00

Fig. S12: Exact distribution of ρ under the null hypothesis.



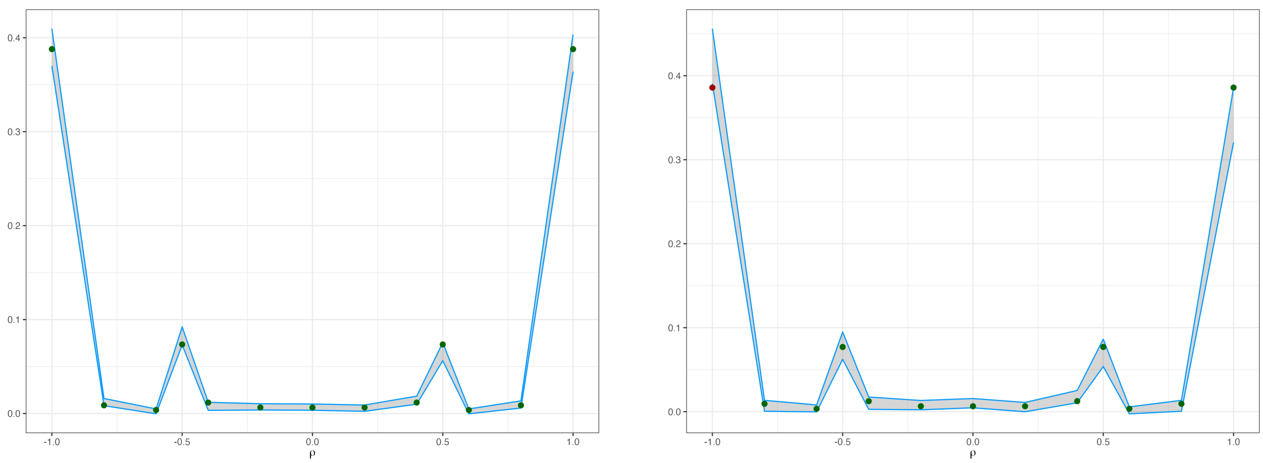
Our test is a simple comparison between the exact distribution of the Spearman's rank correlation under the null and the empirical distribution, $F(\rho)$ and $\widehat{F}(\rho)$, respectively. Intuitively, if the null hypothesis of no dependence between SD and SR is true, then the behavior of ρ in our sample, described by $\widehat{F}(\rho)$, should not be statistically different from the theoretical one, indicated by $F(\rho)$. Practically speaking, we check whether

$$F(r) \in \left[\widehat{F}(r) \pm 1.96 \cdot \widehat{\text{se}}(\widehat{F}(r))\right],$$

for some values of the support of ρ that we select as the ones for which at least two of the six conditional distributions $\{F(\rho \mid n), n = 2, \dots, 7\}$ have non-zero support.[5]

Figure S13 shows the results. The dots represent the exact distribution $F(\rho)$ under the null hypothesis, whilst the blue shaded area depicts the 99% confidence interval around the empirical distribution of ρ. Green dots indicate points of the support for which we fail to reject the null, whereas red dots highlight points for which we reject the null. The two large peaks at 1 and -1 are due to the Z-types for which we observe only two time periods - 70% of the total number of Z-types (see Table S1): in this case the possible values of ρ are only 1 and -1.

Fig. S13: Distribution of the rank correlation under the null hypothesis of no correlation $F(\rho)$ (black dots) and 99% confidence interval associated to the corresponding empirical distribution $\widehat{F}(\rho)$ (shaded blue area).



(i) perfect foresight **(ii) adaptive expectations**

The remarkable result is that there is no evidence of violation of the null hypothesis neither under assumption 1a nor under assumption 1b. We emphasize that those reported in Figure S13 are confidence intervals and not confidence bands. The latter would take into account the fact that we perform multiple hypothesis. However, failing to adequately control for multiple hypothesis testing leads to over-rejection of the null which, if anything, would play against us.

[5]These values are $-1.0, -0.8, -0.6, -0.5, -0.4, -0.2, 0.0, 0.2, 0.4, 0.5, 0.6, 0.8, 1.0$.

S6 Additional figures and tables

Table S2: List of calls in the L488/92 data

Call	Type	Ministerial Decree	Official Journal	Projects	€ 2010 bln
1°	Industry I	M.D. 20.11.1996	SG 288 of 09.12.1996, SO 215	7459	4.55
2°	Industry II	M.D. 30.06.1997	SG 174 of 28.07.1997, SO 151	5988	3.06
3°	Industry III	M.D. 14.08.1998	SG 207 of 05.09.1998, SO 149	12364	2.54
⋆	Correction	M.D. 11.09.1998	SG 219 of 19.09.1998, SO 161		
4°	Industry IV	M.D. 18.02.1999	SG 54 of 06.03.1999 54, SO 47	8766	2.46
5°	Special	M.D. 16.07.1999	SG 174 of 27.07.1999	528	-
6°	Tourism I	M.D. 07.12.1999	SG 297 of 20.12.1999, SO 223	2575	0.63
7°	Special	M.D. 29.10.1999	SG 276 of 24.11.1999	791	0.13
8°	Industry V	M.D. 09.04.2001	SG 121 of 26.05.2001, SO 129	8716	2.14
⋆	Correction	M.D. 10.07.2001	SG 186 of 11.08.2001, SO 208		
9°	Tourism II	M.D. 30.11.2001	SG 2 of 03.01.2002, SO 4	2290	0.40
10°	Trade I	M.D. 10.12.2001	SG 12 of 15.02.2002, SO 9	658	0.17
11°	Industry VI	M.D. 12.02.2002	SG 65 of 18.03.2002, SO 47	3870	1.44
12°	Tourism III	M.D. 12.07.2002	SG 185 of 08.08.2002, SO 165	1695	0.40
13°	Trade II	M.D. 10.07.2002	SG 186 of 09.08.2002, SO 167	485	0.15
14°	Industry VIII	M.D. 27.05.2003	SG 157 of 09.07.2003, SO 105	2936	1.00
15°	Tourism IV	M.D. 14.10.2003	SG 278 of 29.11.2003, SO 186	1127	0.32
16°	Trade III	M.D. 14.10.2003	SG 278 of 29.11.2003, SO 186	492	0.05
17°	Industry VIII	M.D. 15.11.2004	SG 281 of 30.11.2004, SO 172	5845	0.72
⋆	Correction	M.D. 14.01.2005	SG 43 of 22.02.2005, SO 23		
18°	Special	M.D. 07.07.2004	SG 170 of 22.07.2004	117	-
19°	Tourism V	M.D. 05.07.2005	SG 185 of 10.08.2005, SO 141	3097	0.27
20°	Trade V	M.D. 05.07.2005	SG 186 of 11.08.2005, SO 142	2103	0.05
22°	Special	M.D. 16.03.2005	SG 110 of 13.05.2005, SO 89	292	0.06
23°	Craftwork	M.D. 23.12.2004	SG 24 of 31.01.2005, SO 13	2036	-
27°	Special	M.D. 09.04.2004	SG 95 of 12.04.2004	12	0.04
28°	Tourism	M.D. 15.11.2005	SG 276 of 26.11.2005	15	0.04
29°	Industry-Tourism	M.D. 04.08.2006	SG 190 of 17.08.2006	15	0.01
31°	Industry	M.D. 30.12.2006	SG 35 of 12.02.2007, SO 34	1957	0.72
32°	Tourism	M.D. 30.12.2006	SG 42 of 20.02.2007, SO 44	685	0.41
33°	Trade	M.D. 30.12.2006	SG 42 of 20.02.2007, SO 45	332	0.08
34°	Craftwork	M.D. 30.12.2006	SG 37 of 14.02.2007, SO 37	549	-
35°	Special	M.D. 29.12.2006	SG 31 of 07.02.2007	19	0.02
Tot				77286	21.82

Notes: This is a list of the calls included in the L488/92 data supplied by the Ministry of Economic Development. The original data did not include 5 of the 35 calls $(21, 24, 25, 26, 30)$, while for 4 other calls we cannot retrieve the total amount of subsidy $(5, 18, 23, 34)$. The rows denoted with a ⋆ indicate corrections to the final official rankings published on the Official Journal. In our analysis we consider the rankings published in the corrections. The 5th, 7th, 18th, 22nd, and 35th calls do not fall within the usual characterization of L488/92, as they were issued to intervene quickly against natural disasters, or tackle particular issues. For example, call 5 targeted projects in the regions of Umbria and Marche hit by the September 1997 earthquake. Call 18 targeted environmentally sustainable projects. The 22nd call was restricted to firms in minor islands, whilst call 7 was limited to Veneto, Marche, Emilia-Romagna, Liguria, and Umbria. Finally, Call 35 was limited to a subset of firms in the province of Salerno.

References

Angrist, Joshua D and Miikka Rokkanen, "Wanna get away? Regression discontinuity estimation of exam school effects away from the cutoff," *Journal of the American Statistical Association*, 2015, *110* (512), 1331–1344.

Crump, Richard K, V Joseph Hotz, Guido W Imbens, and Oscar A Mitnik, "Dealing with limited overlap in estimation of average treatment effects," *Biometrika*, 2009, *96* (1), 187–199.

Imbens, Guido W and Donald B Rubin, *Causal inference in statistics, social, and biomedical sciences*, Cambridge University Press, 2015.