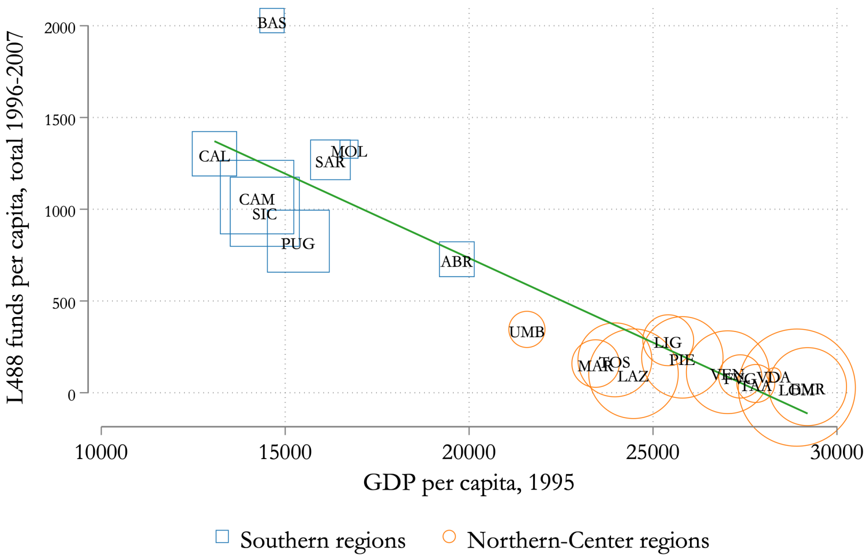
Making Subsidies Work: Rules vs. Discretion
Online Appendix

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

November 17, 2023

1 Additional figures and tables

Fig. A1: L488/92 funds and GDP per capita across regions



Notes: This figure plots the total amount of L488/92 per capita received over the period 1997-2007 (vertical axis) against the GDP per capita in 1995 (horizontal axis), across Italian regions. Both variables are expressed in euros at constant 2010 prices. The size of markers is proportional to region population.

*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. A2: Total L488/92 funds by area-year (left), and industry-year (right)

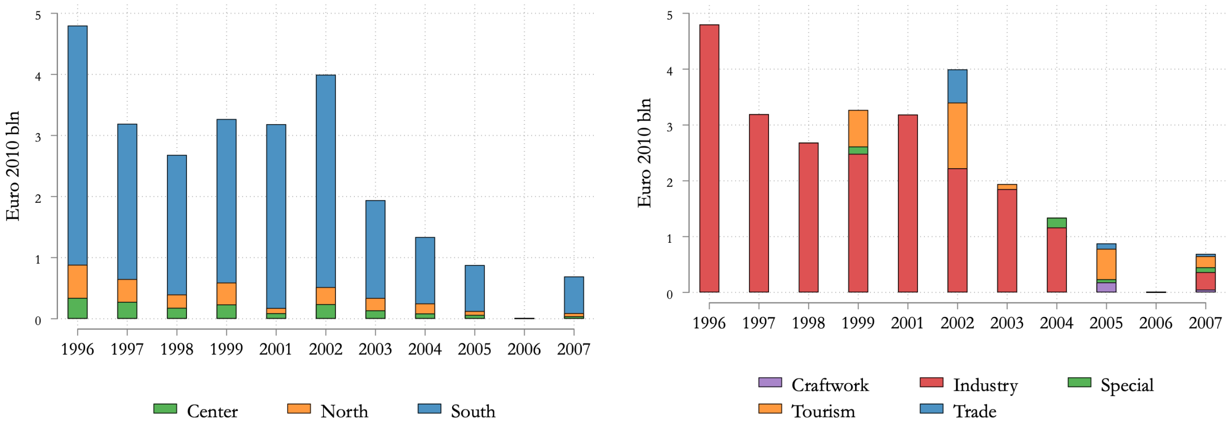


Fig. A3: Total L488/92 funds by region, source, and economic sector

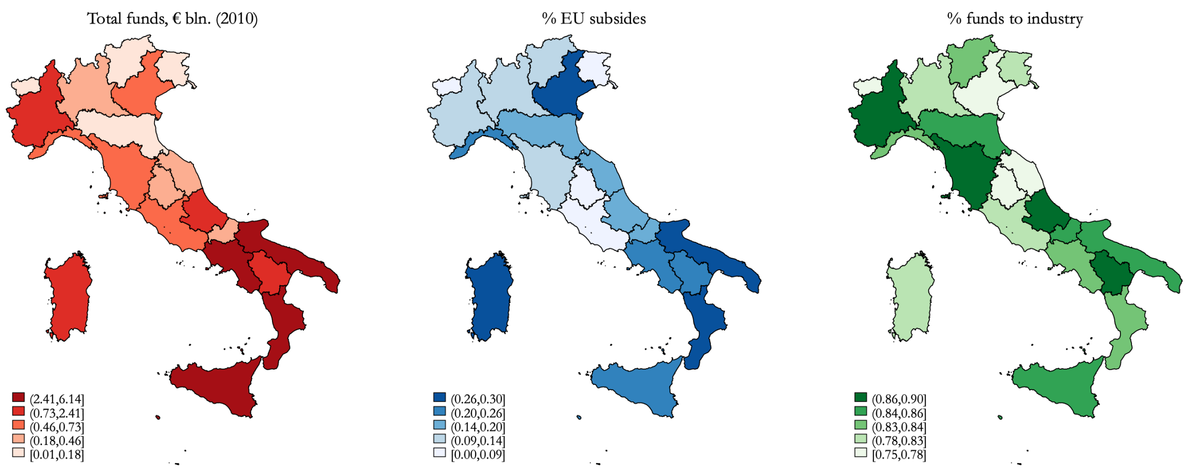
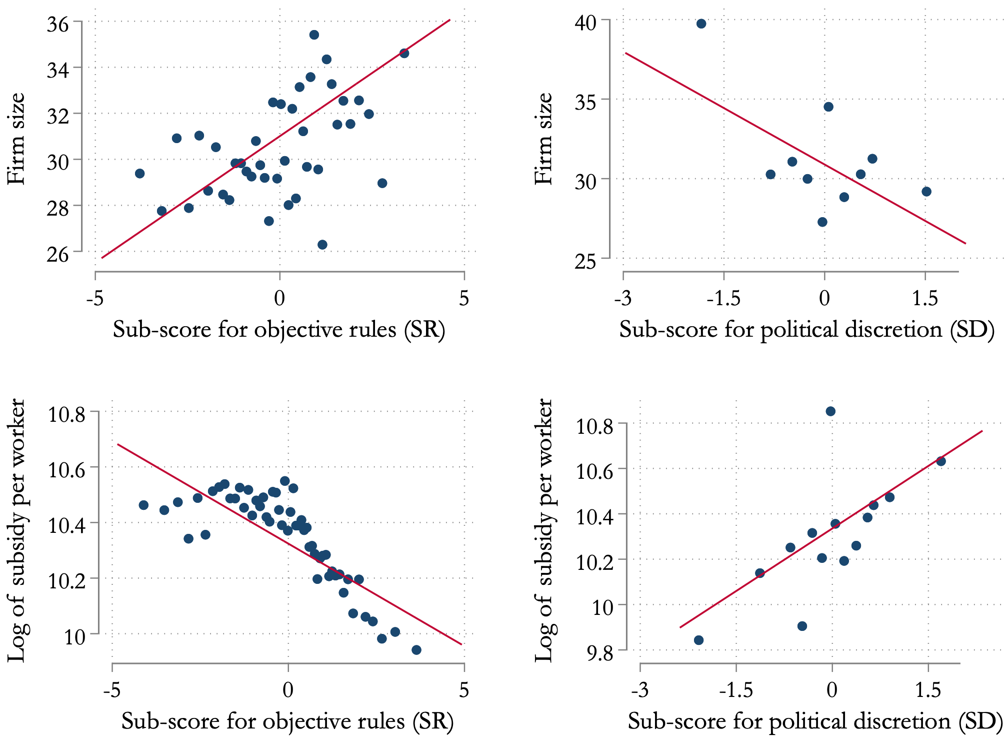
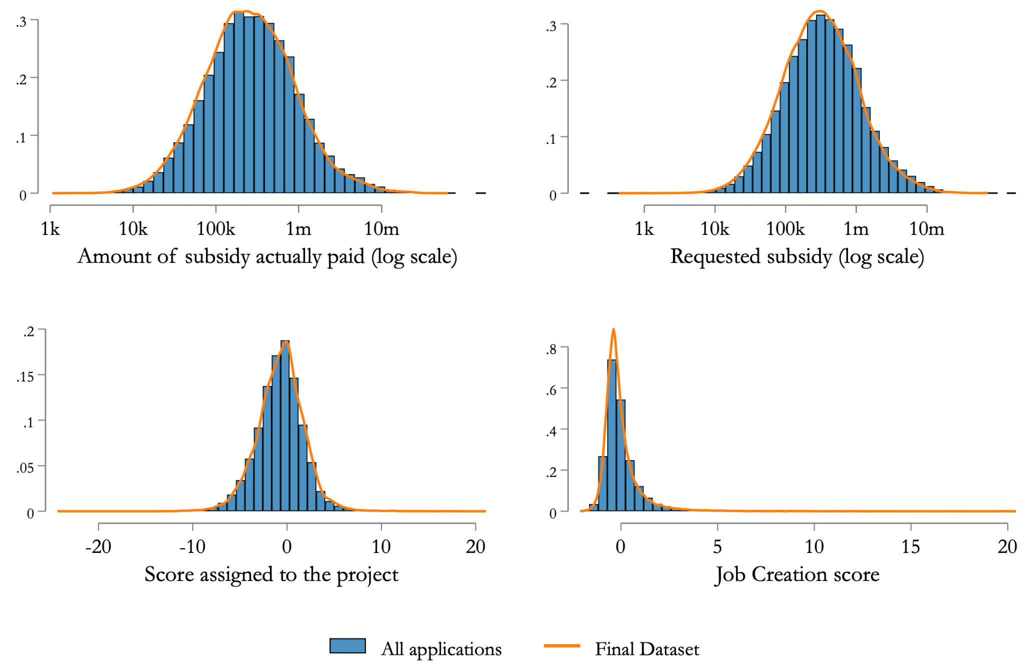


Fig. A4: Political discretion, objective rules, and applicant characteristics



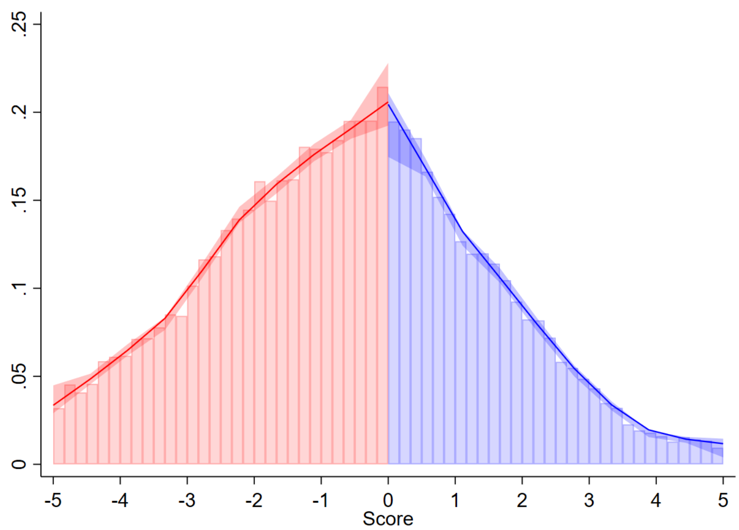
Notes: These graphs plot the sub-scores for political discretion and objective rules (on the horizontal axis), against the size of applicant firms and the amount of subsidies they applied for (on the vertical axis), controlling for cell fixed effects, across quantile-spaced bins. Covariate adjustment and the choice of the optimal number of bins are performed according to Cattaneo, Crump, Farrell and Feng (2022).

Fig. A5: Distribution of selected variables across all applications and within the sub-sample of matched applications



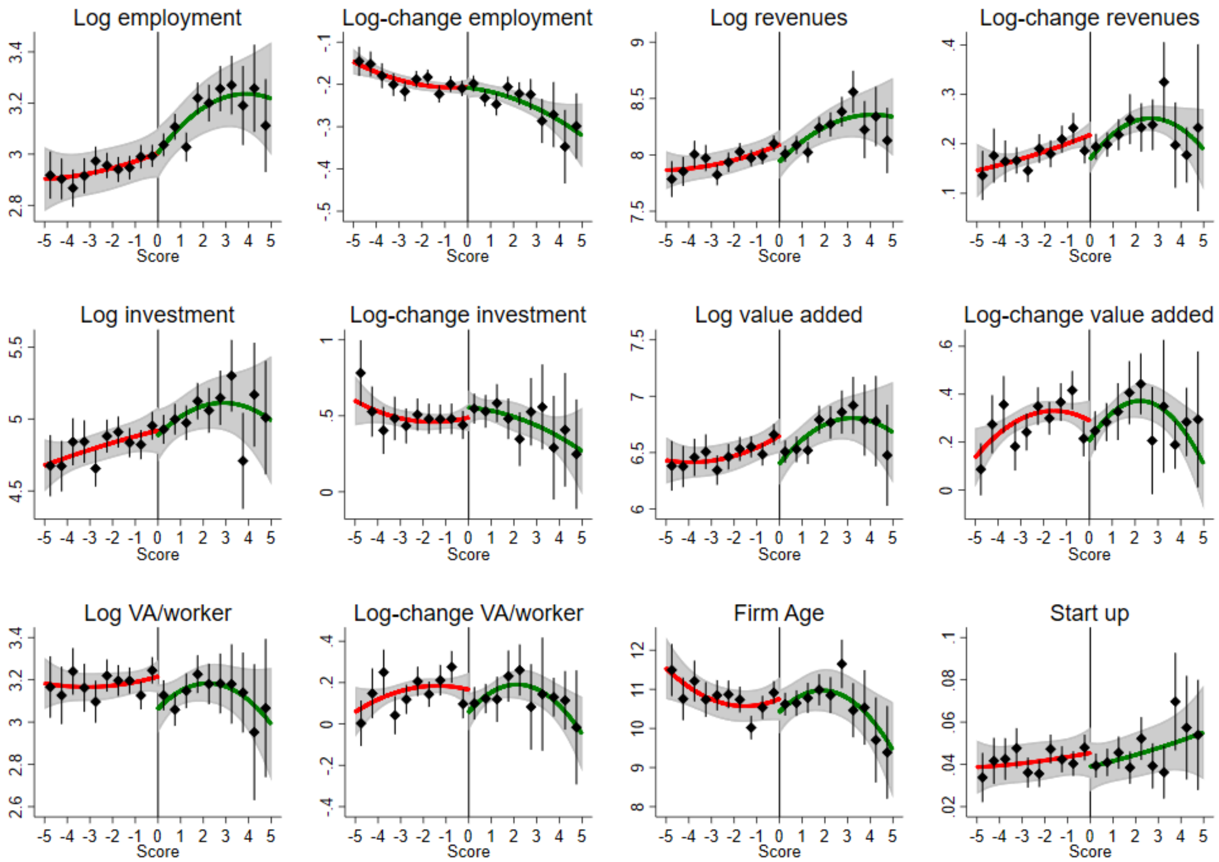
Notes: This figure shows the distribution of some variables across the entire sample of applicants and across the final sample of applicants for which we have complete information on employees and balance sheet data.

Fig. A6: Density of applicant scores



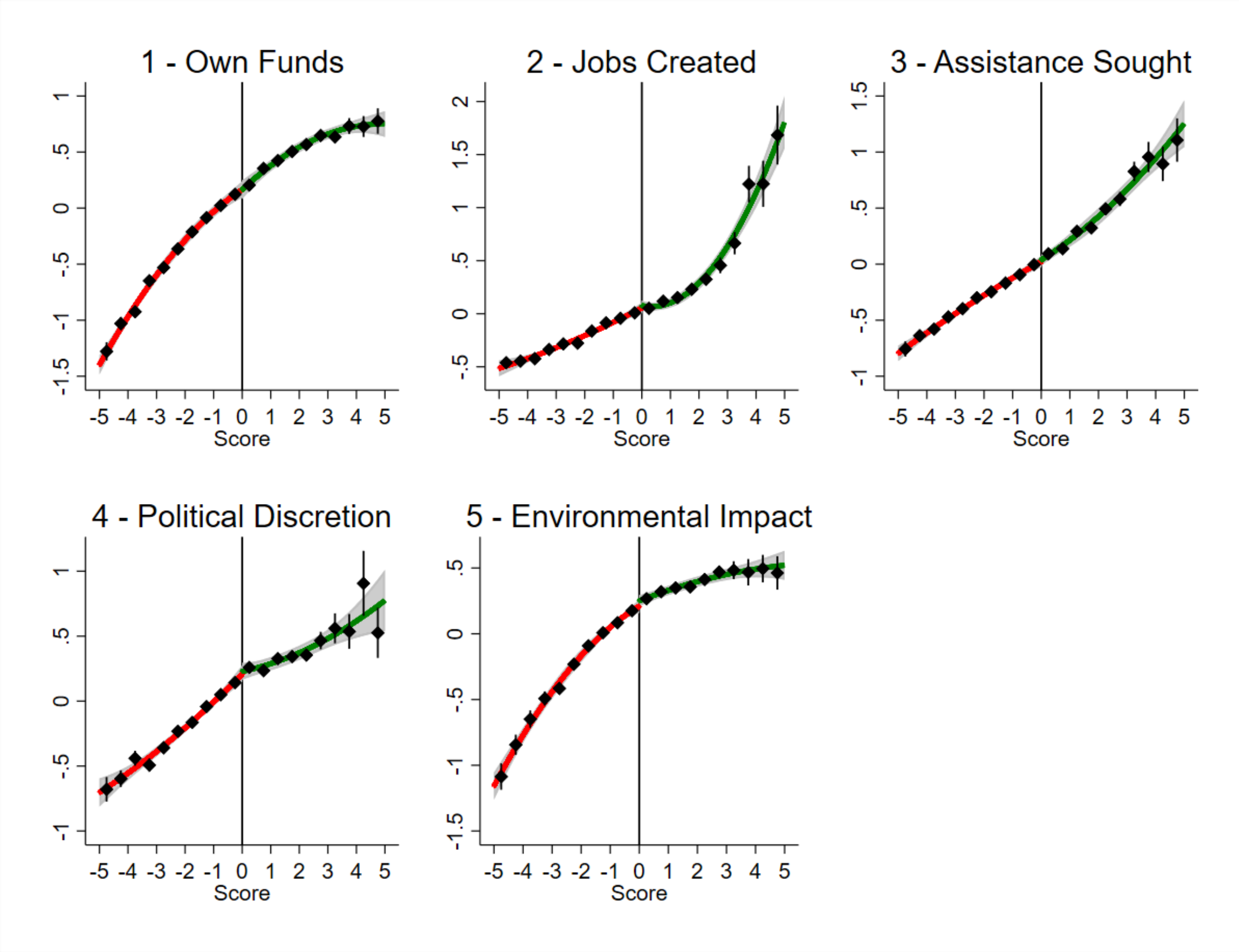
Notes: The histogram shows the distribution of applicant scores. Local polynomial density estimates (solid lines) and robust bias-corrected 95% confidence intervals (shaded areas), computed according to Cattaneo, Jansson and Ma (2020), are also reported in the figure.

Fig. A7: Balance of firm characteristics one year before the call



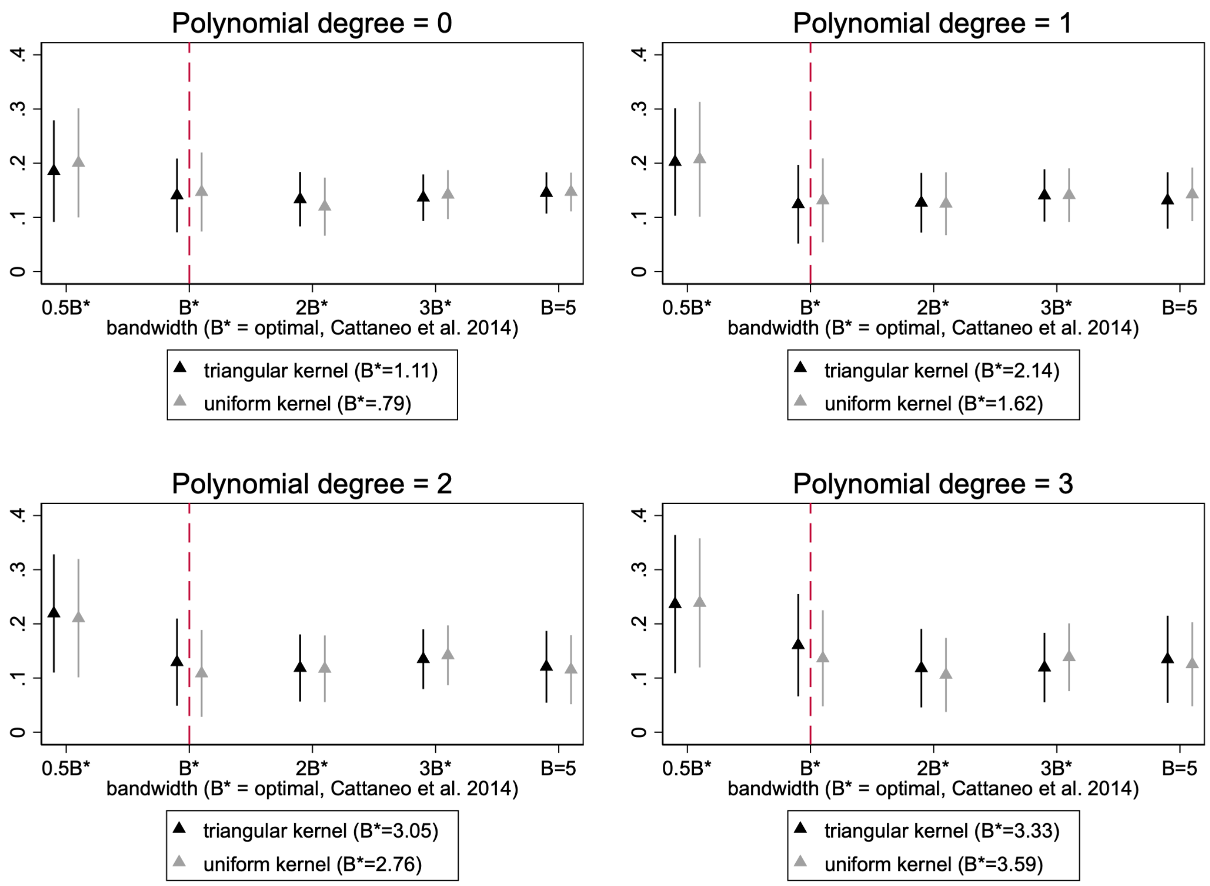
Notes: These graphs show the relationship between the standardized score obtained by firm applications for L488/92 funds, on the horizontal axis, and several firm characteristics measured one year before the call – log and yearly log-change in revenues, value added, value added per worker, investment, firm age and being a start-up. Bins represent averages over equally-spaced intervals, and confidence intervals (at the 90% significance level) are also shown by vertical lines. The predicted relationships between each variable and the score are estimated using a quadratic polynomial regression, controlling for cell-specific fixed effects. 90% confidence bands for the predicted relationship (in grey) are computed based on heteroskedasticity-robust standard errors clustered by cell.

Fig. A8: Balance of the score components



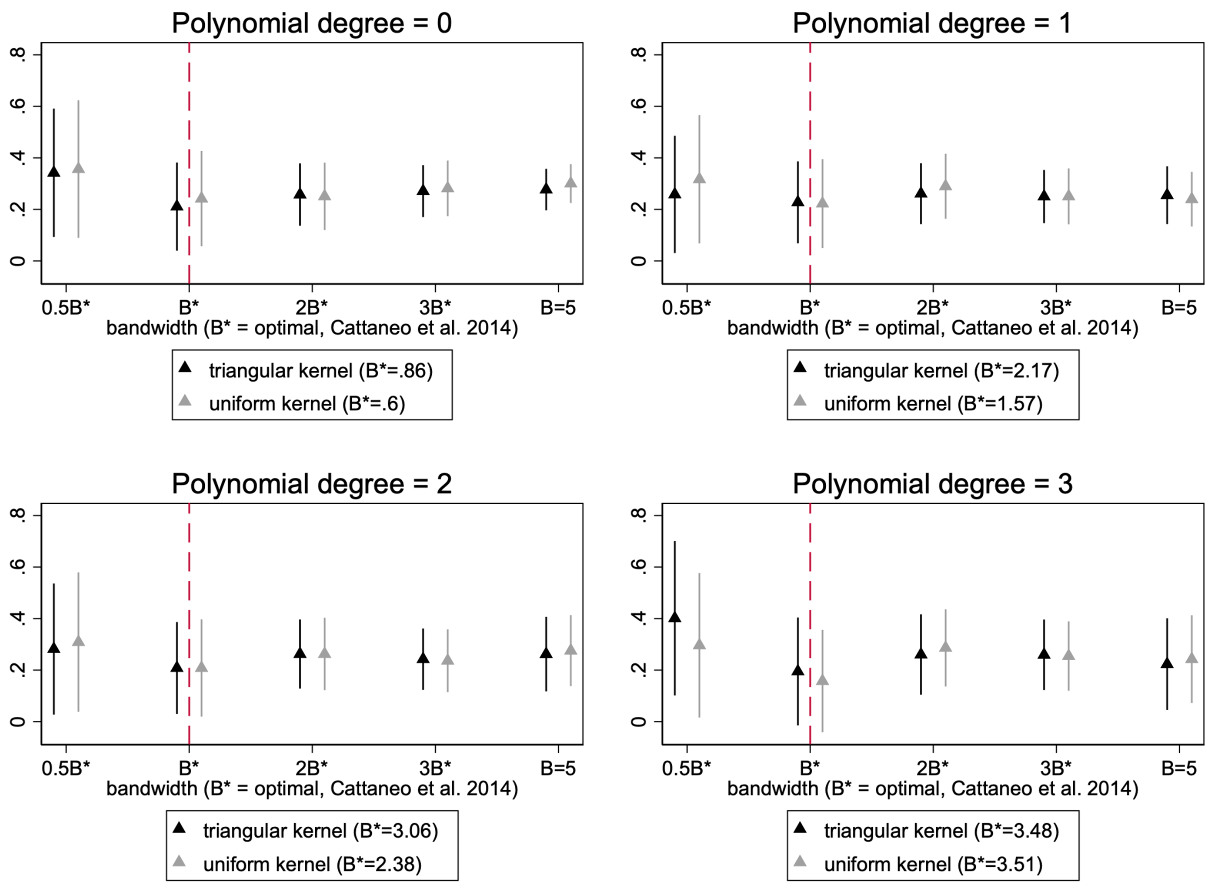
Notes: These graphs show the relationship between the standardized score obtained by firm applications for L488/92 funds, on the horizontal axis, and its five components (described in previous Section 2). Bins represent averages over equally-spaced intervals, and confidence intervals (at the 90% significance level) are also shown by vertical lines. The predicted relationships between each variable and the score are estimated using a quadratic polynomial regression, controlling for cell-specific fixed effects. 90% confidence bands for the predicted relationship (in grey) are computed based on heteroskedasticity-robust standard errors clustered by cell.

Fig. A9: The effect of obtaining the subsidy on firm employment, non-parametric RDD estimates



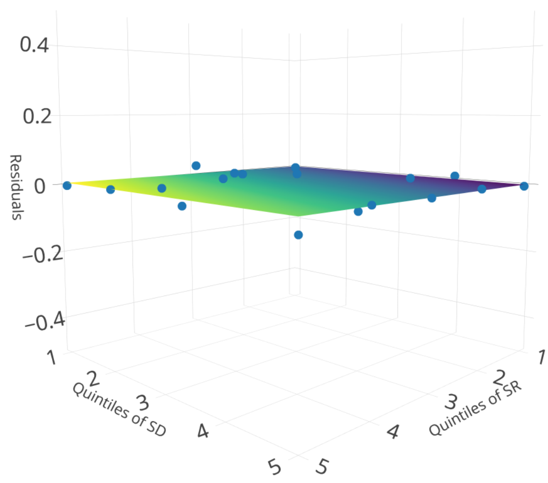
Notes: This figure plots the estimated effect of being eligible for the subsidy (i.e., scoring above the cutoff) on the log-change of firm employment 6 years after the award of subsidies, for different specifications of non-parametric RDD. In particular, each graph shows point estimates and confidence intervals when using triangular and uniform kernels, for different degrees of the polynomial in the running variable (reported on top of each graph) and different bandwidths around the cutoff (on the horizontal axis). The optimal bandwidth B^*, as well as point estimates and 90% confidence intervals are computed following the approach proposed in Calonico, Cattaneo and Titiunik (2014).

Fig. A10: The effect of obtaining the subsidy on firm investment, non-parametric RDD estimates

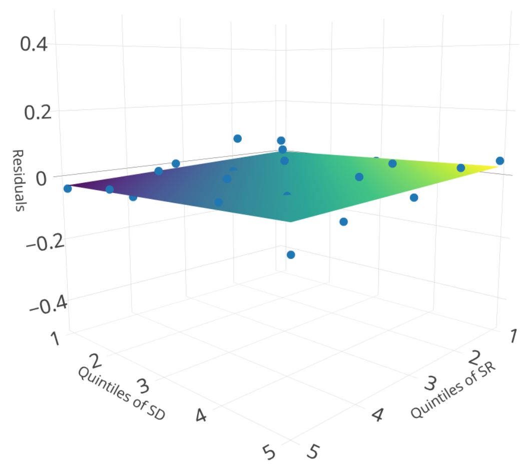


Notes: This figure plots the estimated effect of being eligible for the subsidy (i.e., scoring above the cutoff) on the log of cumulated investment over the 3 years after the award of subsidies, for different specifications of non-parametric RDD. In particular, each graph shows point estimates and confidence intervals when using triangular and uniform kernels, for different degrees of the polynomial in the running variable (reported on top of each graph) and different bandwidths around the cutoff (on the horizontal axis). The optimal bandwidth B^*, as well as point estimates and 90% confidence intervals are computed following the approach proposed in Calonico, Cattaneo and Titiunik (2014).

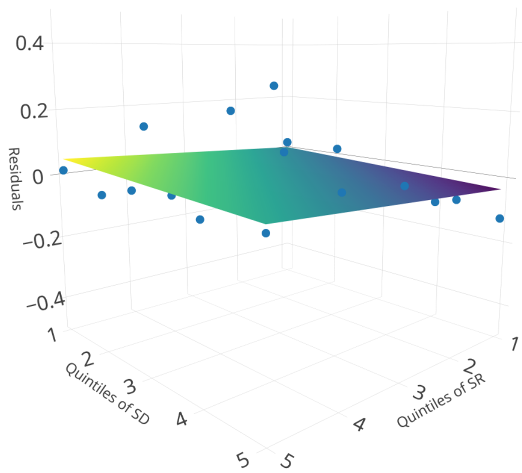
Fig. A11: Testing the CIA: plotting residuals on SR and SD.



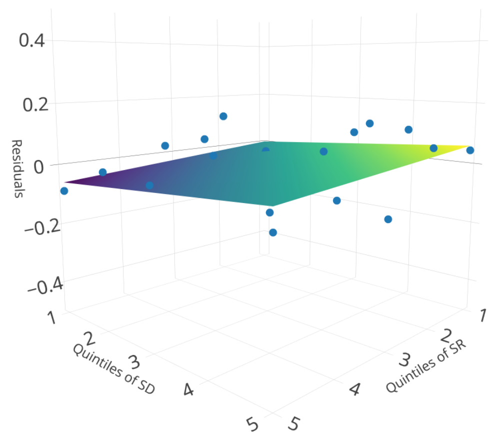
(i) Employment Growth - Left of cutoff



(ii) Employment Growth - Right



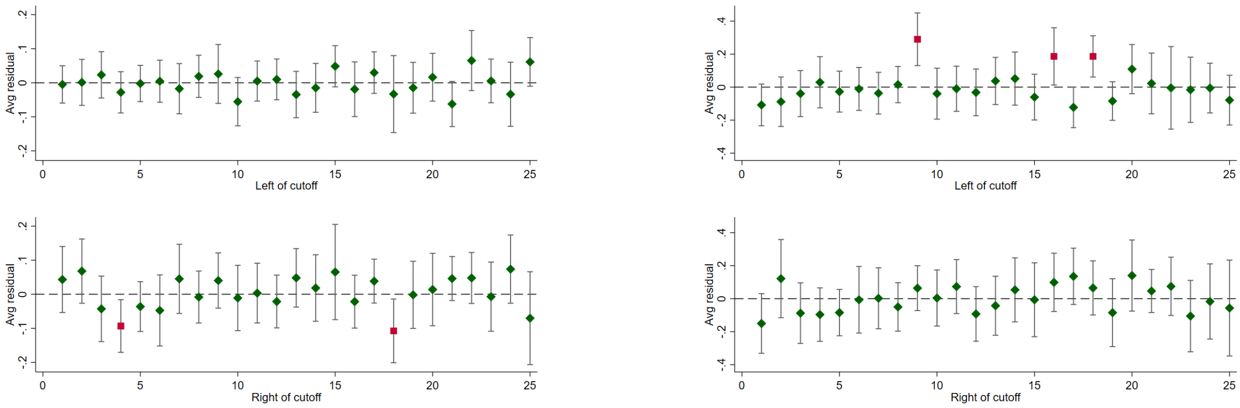
(iii) Investment - Left of cutoff



(iv) Investment - Right

Notes: This figure reports the best interpolating hyperplane of the residuals of the residuals obtained in a regression of the outcome (either employment growth or investment) on $X^\star$ on the quintiles of the sub-scores for objective rules (SR) and political discretion (SD).

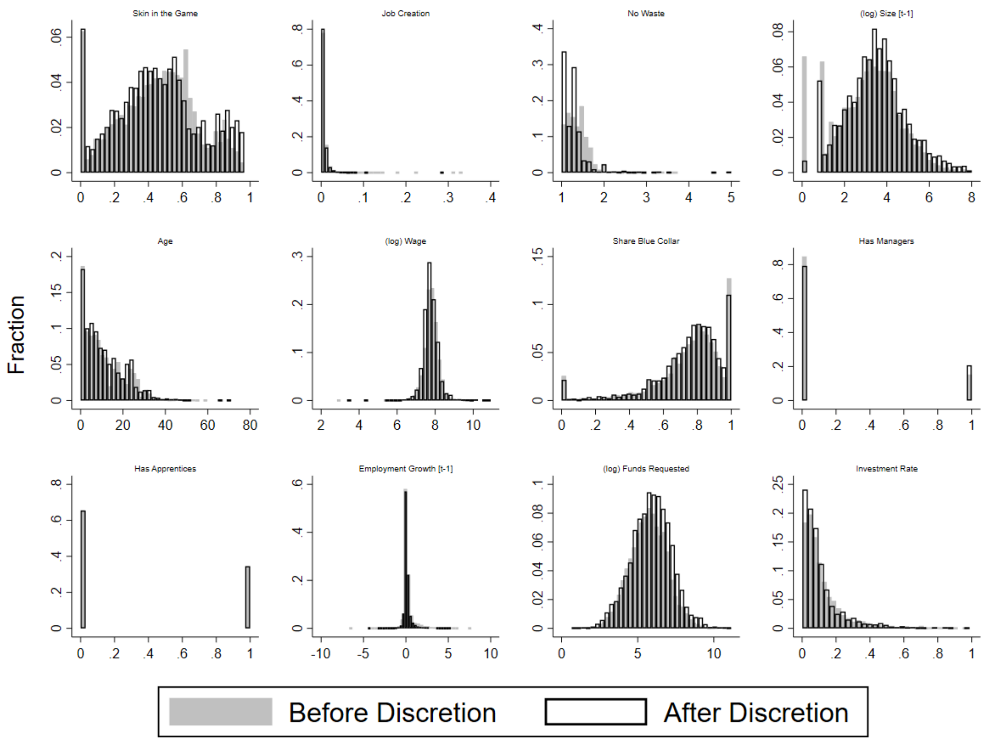
Fig. A12: Testing the CIA: zero-mean residuals test.



(i) Employment Growth **(ii) Investment**

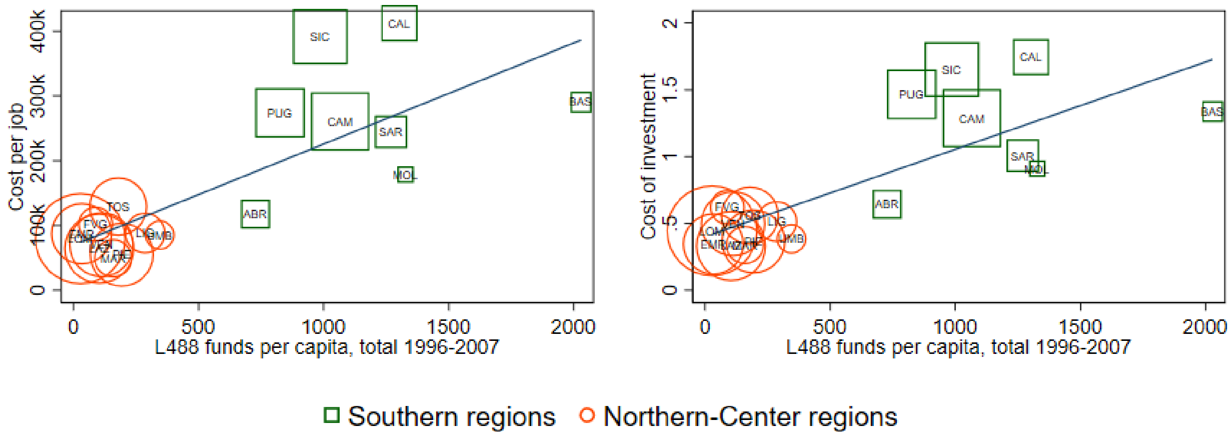
Notes: This figure reports the within-cell average of the regression of the outcome (either employment growth or investment) on $X^\star$ with 95% confidence intervals. The 25 cells are defined as the intersection of 5 quintiles of SR and 5 quintiles of SD.

Fig. A13: Balancing of applicants' characteristics before and after the change in the selection rule



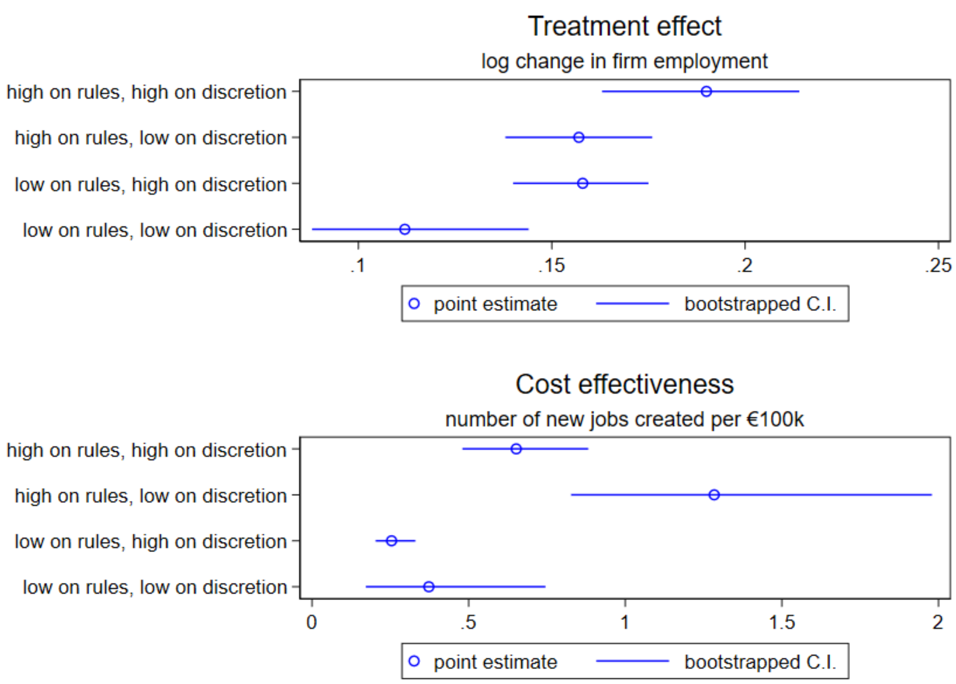
Notes: This figure reports the distribution of the applicants' characteristics in the auctions right before (gray bars) and after (transparent bars) the introduction of political discretion. Only auctions concerning industry are included.

Fig. A14: Cost per job and cost of investment across regions



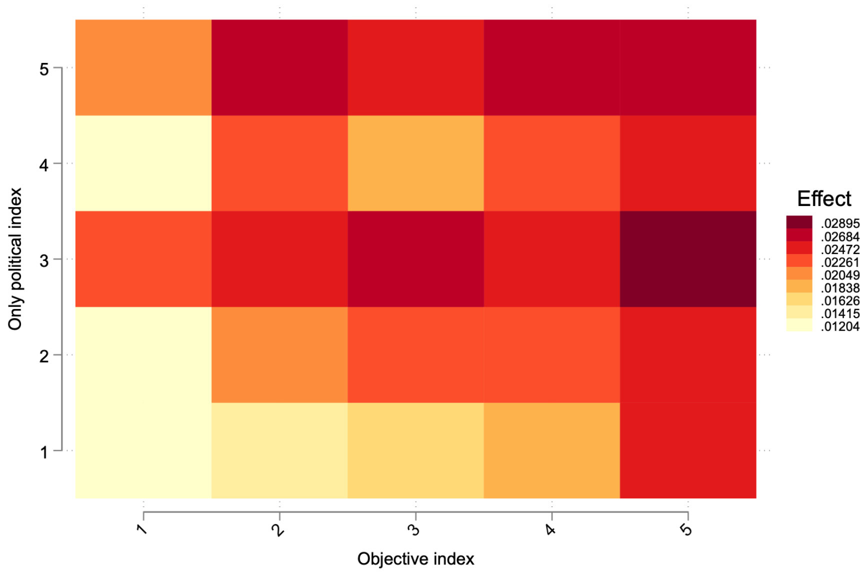
Notes: These are graphs of the estimated cost per job (left graph) and the cost of additional investment (right graph) against the total amount of L488/92 per capita across Italian regions. The size of markers is proportional to regional population.

Fig. A15: Treatment effect and new jobs created per €100,000, rules vs. discretion



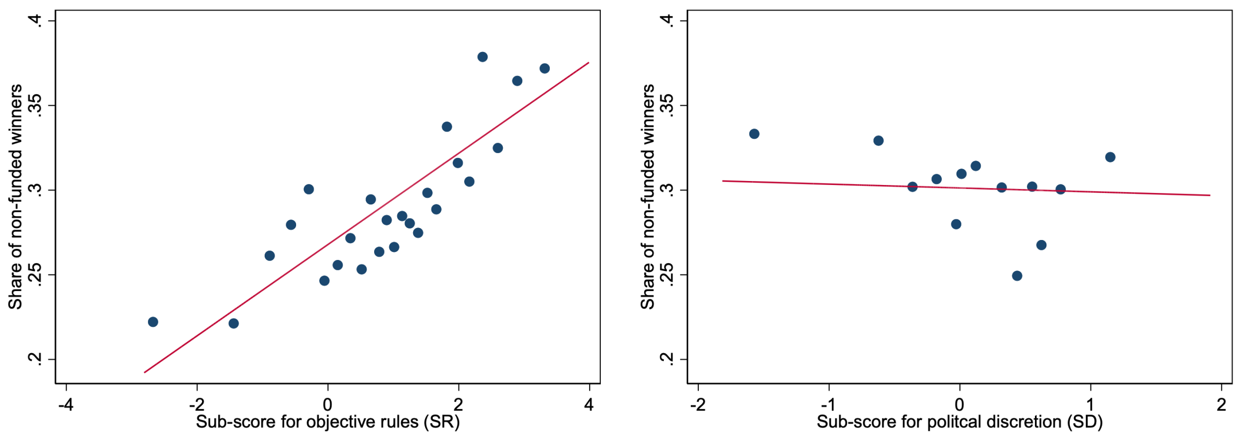
Notes: This figure plots point estimates and bootstrapped confidence intervals for treatment effects on firm employment growth (top graph) and the cost effectiveness of subsidies (bottom graph), for four groups of applicants. Applicants "high on rules" ("low on rules") are those in the top (bottom) quintile of the objective sub-score SR; similarly, applicants "high on discretion" ("low on discretion") are those in the top (bottom) quintile of the discretionary sub-score SD. In practice, the four estimates in each graph refer to the four "corners" of the heatmaps in Figure 8, and 90% confidence intervals are bootstrapped as in Online Appendix Table A8.

Fig. A16: Treatment effect on survival probability



Notes: This figure shows the heterogeneity in treatment effects on firms' survival probability by quintiles of the sub-scores for objective rules (SR) and political discretion (SD). The treatment effect for each bin ($SR = r, SD = d$) is estimated as $\mathbb{E}\left[Y(1) - Y(0) \mid SR = r, SD = d\right] = (\beta_1 - \beta_0) \cdot \mathbb{E}\left[X^\star \mid SR = r, SD = d\right]$. The covariates included in $X^\star$ are listed at the beginning of Section 6.

Fig. A17: Share of projects scoring above the cutoff that are not funded



Notes: This figure shows the relationship between the share of eligible projects scoring above the cutoff that are not funded eventually and, respectively, the sub-score for objective rules (left graph) and the sub-score for political discretion (right graph), controlling for cell fixed effects, across quantile-spaced bins. Covariate adjustment and the choice of the optimal number of bins are performed according to Cattaneo, Crump, Farrell and Feng (2022).

Table A1: Balance of firm characteristics one year before the call

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Specification:	linear				quadratic			
Kernel:	uniform		triangular		uniform		triangular	
Group fixed effects	no	yes	no	yes	no	yes	no	yes
Log employment	0.044	0.002	0.027	0.006	0.000	0.017	0.020	0.026
	(0.043)	(0.034)	(0.04)	(0.034)	(0.048)	(0.04)	(0.048)	(0.04)
	[0.721]	[0.995]	[0.92]	[0.989]	[0.999]	[0.959]	[0.912]	[0.871]
Log-change employment	0.016	0.013	0.009	0.011	-0.001	0.005	0.010	0.016
	(0.013)	(0.014)	(0.014)	(0.015)	(0.018)	(0.018)	(0.019)	(0.019)
	[0.721]	[0.887]	[0.92]	[0.906]	[0.998]	[0.959]	[0.912]	[0.871]
Log revenues	-0.071	-0.004	-0.102	-0.041	-0.151	-0.094	-0.120	-0.076
	(0.06)	(0.049)	(0.061)	(0.051)	(0.078)	(0.063)	(0.079)	(0.064)
	[0.721]	[0.995]	[0.486]	[0.906]	[0.342]	[0.642]	[0.63]	[0.806]
Log-change revenues	-0.021	-0.030	-0.032	-0.038	-0.048	-0.051	-0.036	-0.037
	(0.017)	(0.018)	(0.018)	(0.019)	(0.023)	(0.025)	(0.025)	(0.026)
	[0.721]	[0.557]	[0.457]	[0.33]	[0.282]	[0.276]	[0.642]	[0.687]
Log investment	0.022	0.049	0.001	0.022	-0.034	-0.009	-0.027	-0.009
	(0.079)	(0.071)	(0.083)	(0.077)	(0.107)	(0.098)	(0.108)	(0.098)
	[0.79]	[0.887]	[0.997]	[0.989]	[0.981]	[0.959]	[0.912]	[0.938]
Log-change investment	0.124	0.088	0.102	0.065	0.066	0.045	0.109	0.088
	(0.065)	(0.067)	(0.064)	(0.066)	(0.078)	(0.081)	(0.084)	(0.086)
	[0.409]	[0.733]	[0.486]	[0.906]	[0.849]	[0.959]	[0.712]	[0.871]
Log value added	-0.112	-0.088	-0.165	-0.133	-0.249	-0.208	-0.214	-0.188
	(0.079)	(0.07)	(0.08)	(0.073)	(0.103)	(0.093)	(0.103)	(0.094)
	[0.614]	[0.733]	[0.293]	[0.41]	[0.141]	[0.215]	[0.296]	[0.342]
Log-change value added	-0.065	-0.071	-0.073	-0.077	-0.084	-0.088	-0.084	-0.078
	(0.05)	(0.052)	(0.055)	(0.057)	(0.073)	(0.076)	(0.075)	(0.077)
	[0.721]	[0.729]	[0.551]	[0.724]	[0.781]	[0.775]	[0.756]	[0.871]
Log VA/Worker	-0.081	-0.050	-0.109	-0.083	-0.153	-0.143	-0.150	-0.144
	(0.047)	(0.048)	(0.05)	(0.051)	(0.067)	(0.065)	(0.068)	(0.067)
	[0.462]	[0.844]	[0.253]	[0.542]	[0.192]	[0.239]	[0.233]	[0.249]
Log-change VA/Worker	-0.077	-0.089	-0.089	-0.099	-0.108	-0.116	-0.097	-0.099
	(0.05)	(0.051)	(0.057)	(0.058)	(0.077)	(0.079)	(0.078)	(0.08)
	[0.569]	[0.516]	[0.486]	[0.503]	[0.672]	[0.642]	[0.712]	[0.78]
Firm Age	0.261	0.177	0.029	0.029	-0.335	-0.224	-0.333	-0.249
	(0.245)	(0.216)	(0.249)	(0.22)	(0.313)	(0.287)	(0.31)	(0.282)
	[0.721]	[0.887]	[0.991]	[0.989]	[0.786]	[0.919]	[0.756]	[0.871]
Start Up	-0.006	-0.001	-0.006	-0.004	-0.007	-0.006	-0.010	-0.009
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)
	[0.569]	[0.977]	[0.486]	[0.906]	[0.75]	[0.775]	[0.578]	[0.591]

Notes: This table presents the results from a comparison of firm characteristics one year before the call between applicants scoring just above and just below the cutoff. All variables are described in Online Appendix Table B1. Start-up identifies firms in the age class (0-1). The numbers without brackets are the estimated coefficients from RD regressions analogous to (3) in which the dependent variable is the firm characteristic indicated in each row and the main explanatory variable is a dummy equal to one for firms scoring just above the cutoff. The specification in columns (1)-(4) also includes the standardized application score and its interaction with the dummy for applicants above the cutoff, while columns (5)-(8) also include the squared application score and its interaction with the dummy; odd columns include group fixed effects for firms competing in the same ranking; and columns (3)-(4) and (7)-(8) weight observations by a triangular kernel in distance from the cutoff. Standard errors clustered by cell are reported in parenthesis. For each specification, p-values computed controlling the family-wise error rate when performing multiple hypothesis tests (Westfall and Young, 1993) are reported in square brackets.

Table A2: Balance of firm and project characteristics in INPS and Cerved datasets.

Variable Name	INPS	INPS & CERVED	Normalized Difference	Westfall-Young p-value
Skin in the Game	0.508	0.519	0.052	0.808
Job Creation	0.009	0.009	-0.004	1.000
No Waste	1.346	1.279	-0.099	0.982
(log) Size [t-1]	2.820	3.262	0.266	0.001
Employment Growth [t-1]	0.184	0.222	0.063	0.960
(log) Funds Requested	5.659	5.846	0.150	0.960
Age	11.222	10.705	-0.050	0.663
South-based	0.774	0.714	-0.138	0.747

Notes: The second and third columns of the table report averages of firms' and firms' projects' characteristics for applicants matched only in the INPS dataset and in the INPS and Cerved dataset, respectively. The fourth and fifth columns report the normalized difference (see Imbens and Rubin, 2015, Section 14.2) of the second and third columns, and the Westfall and Young (1993) p-values computed controlling the family-wise error rate for multiple hypothesis tests. Only auctions concerning industry are included. The typical rule of thumb to detect imbalances for (the absolute value of) the normalized difference is 0.25 (see Imbens and Wooldridge, 2009, p.24).

Table A3: Conditional independence tests

Variable	Left of cutoff	Right of cutoff
Conditional on $X^{\star}$:		
Running variable	0.0012	-0.0029
t-statistic	0.313	0.334
p-value	0.754	0.734
Unconditional:		
Running variable	0.0388	0.0145
t-statistic	5.155	1.265
p-value	0.000	0.206
Obs	16,007	11,045

Notes: The table reports regression-based tests of the conditional independence assumption in equation (4). We regressed employment growth in the six years after the award of L488/92 subsidies on the running variable (i.e., the application score) separately for the sub-samples of applicants above and below the cutoff. The top panel shows the estimated coefficients when controlling for cell fixed effects and for the vector of covariates $X^{\star}$, while the bottom panel reports the estimated coefficients when controlling only for cell fixed effects. Results are robust to including a quadratic polynomial in the running variable. The covariates included in $X^{\star}$ are listed at the beginning of Section 6.

Table A4: Applicants' characteristics before and after the change in the selection rule

Variable Name	Before Discretion	After Discretion	Normalized Difference	Westfall-Young p-value
Skin in the Game	0.458	0.453	-0.019	0.915
Job Creation	0.006	0.007	0.098	0.915
No Waste	1.288	1.359	0.233	0.459
(log) Size [t-1]	3.463	3.095	-0.247	0.414
Age	11.297	10.840	-0.048	0.999
(log) Wage	7.765	7.803	0.094	0.972
Share Blue Collar	0.749	0.751	0.005	0.993
Has Managers	0.207	0.153	-0.143	0.934
Has Apprentices	0.345	0.344	-0.004	0.873
Employment Growth [t-1]	0.136	0.245	0.183	0.998
(log) Funds Requested	5.903	5.700	-0.158	0.998
Investment Rate	0.113	0.120	0.059	0.365

Notes: This table reports averages of applicants' characteristics in the auctions right before and right after the introduction of political discretion, their normalized difference (see Imbens and Rubin, 2015, Section 14.2), and the Westfall and Young (1993) p-values computed controlling the family-wise error rate for multiple hypothesis tests. Only auctions concerning industry are included. The typical rule of thumb to detect imbalances for (the absolute value of) the normalized difference is 0.25 (see Imbens and Wooldridge, 2009, p.24).

Table A5: Cost of new jobs and investment matching firms in the South

	(1)	(2)	(3)	(4)	(5)	(6)
Cost measure:	cost per new job (thousands of €'s)		cost per worker-year (thousands of €'s)		cost of new investment (cost per €1 of investment)	
$X^{\star}$:	manual	data-driven	manual	data-driven	manual	data-driven
all regions	115	106	30	30	0,63	0,49
	[87; 220]	[80; 206]	[27; 38]	[27; 36]	[0.47; 1.10]	[0.36; 0.69]
south	140	122	39	35	0,78	0,63
	[107; 238]	[95; 205]	[33; 51]	[31; 44]	[0.59; 1.44]	[0.47; 0.93]
north-center	68	70	19	25	0.35	0.25
	[41; 211]	[44; 214]	[17; 24]	[22; 29]	[0.24; 0.59]	[0.18; 0.34]

Notes: This table shows the cost of new jobs and of new investment generated by the L488/92 subsidies over a six-year period. All amounts are expressed in thousand € at constant 2010 prices. Differently from Table 4 in the main text, costs are calculated on a subsample of Southern firms matched (1-to-1) to Northern firms based on a set of observables (age and industry, size and employment composition, average wage, and past employment growth). The estimates in columns labelled as "manual" employ the set of covariates listed at the beginning of Section 6, while the estimates in columns labelled as "data-driven" employ the set of covariates selected by the algorithm described in detail in Section S3 of the Supplementary Materials. 90% confidence intervals are reported in brackets and are computed using 1,000 draws of a non-parametric cluster Efron bootstrap, where clusters are defined at the cell-level.

Table A6: Policy invariance test, difference-in-differences

Variable Name	Skin in the Game	Job Creation	No Waste	(log) Size $[t-1]$	Age	(log) Wage
$POST1998 \times DISCRETION$	-0.032	-0.000	0.082	-0.108	0.138	0.052
	[0.027]	[0.001]	[0.086]	[0.073]	[0.280]	[0.014]
Obs	38,367	38,367	38,367	38,367	38,367	34,747
Adjusted R^2	0.109	0.100	0.685	0.121	0.045	0.092
WY p-value	0.826	0.932	0.870	0.746	0.932	0.142
	Share of Blue Collar	**Has Managers**	**Has Apprentices**	**Employment Growth $[t-1]$**	**(log) Funds Requested**	**Investment Rate**
$POST1998 \times DISCRETION$	-0.004	-0.005	-0.065	-0.018	-0.030	0.008
	[0.011]	[0.016]	[0.012]	[0.012]	[0.062]	[0.007]
Obs	34,747	38,367	38,367	34,819	38,367	15,104
Adjusted R^2	0.020	0.070	0.053	0.006	0.230	0.009
WY p-value	0.932	0.932	0.104	0.736	0.932	0.826

Notes: This table shows the results of difference-in-differences regressions comparing project and applicant characteristics between regions attributing and not attributing discretionary points, before and after the introduction of discretion. In particular, we estimate the specification $Y_{irt} = \phi(POST1998_t \times DISCRETION_r) + FE_r + FE_t + \nu irt$, where $POST1998_t = 1$ for the period after 1998 and $POST1998_t = 0$ otherwise, $DISCRETION_r = 1$ in regions attributing discretionary points and $DISCRETION_r = 0$ otherwise, and FE_r and FE_t are region and year fixed effects, respectively. Robust standard errors clustered at the region-year level and reported in brackets. The last row reports Westfall and Young (1993) p-values corrected for multiple-hypothesis tests.

Table A7: Cost of new jobs and investment generated by L488 subsidies

	(1)	(2)	(3)	(4)	(5)	(6)
Cost Measure	cost per new job (thousand of €'s)		cost per worker-year (thousand of €'s)		cost of new investment (thousand of €'s)	
$X^{\star}$	manual	data-driven	manual	data-driven	manual	data-driven
all firms	178	159	54	56	0.81	0.63
	[133; 299]	[118; 260]	[47; 62]	[51; 62]	[0.59; 1.25]	[0.48; 0.87]
large	78	78	24	29	0.4	0.29
	[47; 222]	[52; 174]	[19; 30]	[25; 35]	[0.27; 0.74]	[0.21; 0.41]
small	253	209	80	74	1.08	0.87
	[203; 349]	[162; 296]	[73; 90]	[68; 81]	[0.83; 1.49]	[0.68; 1.15]

Notes: This table shows the cost of new jobs and investment generated by the L488 subsidies over a six-year period. All amounts are expressed in thousand € at constant 2010 prices. The estimates in columns labelled as "manual" employ the set of covariates listed at the beginning of Section 6, while the estimates in columns labelled as "data-driven" employ the set of covariates selected by the algorithm described in detail in Section S3 of the Supplementary Materials. 90% confidence intervals are reported in brackets and are computed using 1,000 draws of a non-parametric cluster Efron bootstrap, where clusters are defined at the cell-level.

Table A8: Point estimates and confidence intervals of treatment effect and average cost per new job, rules vs. discretion

Panel A: Treatment Effect

Quintiles of sub-score SD		Quintiles of sub-score SR: 1	2	3	4	5
	5	0.158	0.188	0.172	0.180	0.190
		[.14; .175]	[.163; .217]	[.151; .195]	[.159; .201]	[.163; .214]
	4	0.126	0.155	0.163	0.167	0.149
		[.099; .156]	[.131; .179]	[.141; .191]	[.137; .199]	[.129; .172]
	3	0.164	0.163	0.171	0.181	0.176
		[.126; .193]	[.139; .185]	[.148; .189]	[.154; .203]	[.152; .197]
	2	0.122	0.151	0.148	0.166	0.171
		[.09; .154]	[.127; .171]	[.122; .168]	[.133; .194]	[.142; .192]
	1	0.112	0.107	0.132	0.135	0.157
		[.088; .144]	[.084; .136]	[.107; .156]	[.114; .155]	[.138; .176]

Panel B: Cost Effectiveness

Quintiles of sub-score SD		Quintiles of sub-score SR: 1	2	3	4	5
	5	0.254	0.340	0.385	0.465	0.652
		[.203; .33]	[.269; .468]	[.286; .519]	[.291; .695]	[.48; .882]
	4	0.359	0.491	0.505	0.682	1.157
		[.261; .502]	[.306; .852]	[.348; .754]	[.466; .984]	[.751; 1.739]
	3	0.487	0.428	0.526	0.681	1.054
		[.388; .677]	[.337; .577]	[.39; .772]	[.493; .991]	[.611; 1.648]
	2	0.466	0.434	0.444	0.610	1.149
		[.324; .741]	[.343; .589]	[.32; .754]	[.472; .837]	[.588; 1.935]
	1	0.373	0.703	0.753	0.816	1.284
		[.172; .745]	[.505; .987]	[.595; .999]	[.621; 1.124]	[.827; 1.979]

Notes: This table reports the heterogeneity in treatment effects on firm employment growth (Panel A) and the cost effectiveness of subsidies (Panel B), by quintiles of the sub-scores for objective rules (SR) and political discretion (SD). In Panel A, the treatment effect for each bin ($SR = r, SD = d$) is estimated as $\mathbb{E}\left[Y^1 - Y^0 \mid SR = r, SD = d\right] = (\beta_1 - \beta_0) \cdot \mathbb{E}\left[X^\star \mid SR = r, SD = d\right]$. The covariates included in $X^\star$ are listed at the beginning of Section 6. In Panel B, cost effectiveness is measured by the number of newly created per €100,000 of subsidies received by the firm. The number of newly created jobs in each bin is computed by multiplying the size of each firm by the treatment effect for its respective bin, as reported in Panel A, and aggregating across all firms in that bin. 90% confidence intervals are reported in brackets and computed using 1,000 draws of a non-parametric cluster Efron bootstrap, where clusters are defined at the cell-level.

2 Data Description

The analysis leverages on three main sources of micro data:

1. The administrative data on all applications for L488/92 subsidies (1996-2007), sourced from the Italian Ministry of Economic Development, DG Firm subsidies;

2. The National Social Security Institute (INPS) firm archive (called DM10M) covering the universe of Italian firms with at least one dependent worker, available starting from 1986;

3. The Cerved database containing balance sheet information on Italian limited liability companies, available starting from 1993.

The L488/92 archive contains administrative data on 74,584 applications for L488/92 subsidies, submitted by 49,082 firms. It cover nearly the universe of rankings (only some smaller auctions are excluded) for which it contains all submitted applications.

The data contain:

(i) information on the project: a unique project identifier; the three (five) sub-indexes measuring project quality; the score (the forcing variable) obtained aggregating the sub-indexes standardized at the auction and region level; the position in the ranking; an indicator for winning projects; the amount of funds requested in the application and that of funds actually transferred - separate by each of three instalments; whether financed on EU or Italian sources.

(ii) information on the auction (number, region and sector of destination of funds, date of issuance, date of closure, dates of each of the three instalments).

(iii) information on the firm (fiscal identifier or individual fiscal code (in case of sole proprietorships), legal form, address, municipality.

Additional information on the auction, recovered from the Official Journal of the Italian Republic (*Gazzetta Ufficiale*), associate each project to "cells" identified by the following dimensions: firm size (Large/Medium/mall), sector (Industry, energy, Tourism, Trade, Services), eligibility for co-financing (Yes/No), and geographical area (Region). This additional information allowed to exactly allocate projects to the several sub-rankings within the same call, region, and (possibly) special category of applicants (see Section S1 of the Supplementary Materials).

The firm archive is assembled by INPS sourcing on a master dataset collecting all social security payments made every month by legal entities for any employee with open-ended, fixed-term and apprenticeship contract. The archive covers therefore the universe of firms with at least one employee at some point during a given calendar year. The data is available between 1986 and 2015. For each firm it reports the fiscal code; monthly information on the number of employees; and yearly information on the number employees and their total wage bill by qualification (manager, blue collar; white collar; apprentices; others); date of birth and cessation of activity; detailed geographical (municipality) and industry (3-digit) data; and an identified for firms belonging to groups.

Information on firms' balance sheet and income statement comes from a proprietary database assembled by the Cerved Group S.p.a. The Cerved Firm Registry, which is the Italian source of the Orbis database, covers the universe of limited liability firms in the private non-financial sector and is available since 1993.

Further data used in the paper include: (*i*) the administrative registries of local politicians and local elections, available from the Italian Ministry of Interior, `https://dati.interno.gov.it/`; (*ii*) a classification of local governments' ideologies, sourced from the Local Opportunities Lab, `https://www.localopportunitieslab.it/`; and (*iii*) data various economic variables related to labor market participation, unemployment, employment rates, education, and other demographic and economic indicators at the municipality level, obtained from the Census, `https://ottomilacensus.istat.it/`.

Table B1: Description and source of all the variables used in the analysis.

Variable	Description	Source
	Main variables from L488/92 data	
Info on Auctions	Date, region, and result of the auction. Complementary information from the Official Journal includes, for each project, all the details required to recover the rankings within each auction-region cell, as explained in Section S1 of the Supplementary Materials	MinEcDev
Score	Project quality obtained combining the 3(5) indicators below, once standardized within each call-region	MinEcDev
Skin in the Game	Ratio of the applicant's own investment in the project relative to the amount requested	MinEcDev
Job Creation	Number of jobs created by the project	MinEcDev
No Waste	Proportion of funds requested in relation to an ad-hoc benchmark set by the EU Commission	MinEcDev
Political Discretion	Score attributed on the basis of priorities indicated by the regional government	MinEcDev
Environmental Responsibility	Compliance with the requirements of an environmental management system, e.g. ISO 14001 or EMAS	MinEcDev
Funds Requested	Amount of subsidies requested in application	MinEcDev
Funds Paid	Amount of subsidies disbursed to winners, in three instalments	MinEcDev
	Main variables from INPS	
Size	Number of employees, available at monthly frequency	INPS
Growth	Employment growth rate between two dates. Computed over different horizons starting and ending in the month of the auction	INPS
Age	Firm age at any given year	INPS
Wage	Average wage of employees. Obtained aggregating yearly data on wage bill and employees by qualification (managers, blue collar; white collar; apprentices; others)	INPS
Share of blue collars	Ratio between blue collar employees and total employees, computed from the same data	INPS
Manager	Dummy for presence of (middle) managers in workforce	INPS
Apprentices	Dummy for presence of apprentices in workforce	INPS
Survival	Dummy for whether firm is alive at any given future horizon	INPS
Area	Headquarter municipality	INPS
Industry	3-digits NACE Rev. 2 industry codes	INPS
	Main variables from CERVED	
Revenues	Firm revenues (sales) (thousands of €)	CERVED
VA	Firm value added(thousands of €)	CERVED
Total Assets	Total assets (thousands of €)	CERVED
Investment	Investment in tangible and intangible fixed assets (thousands of €)	CERVED

Description and source of all the variables used in the analysis (CONT.).

Variable	Description	Source
Political proximity and other predictors of the discretionary score (SD)		
Political alignment	Dummy for the same party (right, centre, left, civic) ruling both the Region and the municipality the firm is located	Ministry of Interior and Local Opportunities Lab
Margin of victory	Dummy for the margin of victory in the last elections municipality the firm is located	Ministry of Interior
Birthplace of Regional president	Dummy for the president of the Regional government being born in the municipality the firm is located	Ministry of Interior
Birthplace of Regional counsellor	Dummy for at least one counsellor in the Regional government being born in the municipality the firm is located	Ministry of Interior
Birthplace of Regional alderman	Dummy for at least one alderman in the Regional government being born in the municipality the firm is located	Ministry of Interior
Human capital of Regional president / municipality mayor	Dummy for level of schooling of Regional president / municipality mayor (primary, lower sec., high school, university degree)	Ministry of Interior
Local unemployment	Unemployment rate at province level (ISTAT)	ISTAT
Credit constraints	Spread between loan and deposit rates in provinces	Guiso et al. (2013)
Main variables from Census Data		
Participation Rate	Labor market participation (males, females), ratio of active and inactive young people	Census data
NEET	Incidence of young people aged 15-29 not studying, not working, and outside the labor market and education	Census data
Unemployment	Male, female, and youth unemployment rate	Census data
Employment	Male, female, and youth employment rate, employment turnover index, incidence of employment in the agricultural, industrial, tertiary (excluded trade), and trade sector	Census data
Education	Early exit from the education and training system, incidence of adults with a diploma or higher, incidence of adults with a middle school diploma	Census data
Socio-economic	Population density, housing usage potential in urban centers, incidence of families at risk of economic hardship	Census data

3 RD estimates at the cutoff: Additional results

Two important issues could affect the interpretation of our RD estimates in Section 5. First, applicants in a given call may re-apply (and obtain funds) in subsequent calls. We deal with this issue in Supplementary Appendix, Section S2. Second, the effects on funded firms may spill over to other, non-funded firms.

Spillovers Employment increases by subsidized firms may affect other, non-subsidized firms. The sign of these effects is also unclear a priori. The growth of subsidized firms may benefit upstream and downstream producers in the same market, or it may erode the market share of competitors – possibly including firms in the control group. In the latter case, estimates in section 5 would overstate the effects of the policy.

To address this possibility, we look across Italian Local Labor Markets (LLM) comparing the employment dynamics of non-subsidized firms in subsidized LLMs to those of firms in non-subsidized LLMs; spillover effects should affect more (or exclusively) employment in the former group. We focus on the following specification:

$$\ln L_{m,t+k} - \ln L_{m,t} = \theta_k D_{m,t} + \alpha \ln L_{m,t} + FE_m + FE_t + \varepsilon_{m,t} \tag{1}$$

where $L_{m,t+k}$ and $L_{m,t}$ are the total employment of non-subsidized firms in the m-th LLM in year $t+k$ and t, taken from the INPS administrative data on the universe of workers in (non-agricultural) firms; $D_{m,t}$ is a dummy equal to 1 when at least one firm in LLM m received funds in year t; FE_m and FE_t are LLM- and year-specific fixed effects; and $\varepsilon_{m,t}$ is a residual summarizing the effect of other factors. The coefficient of main interest, θ_k, captures the differential employment response, after k years, of non-subsidized firms within the same LLM as a subsidized firm relative to non-subsidized firms in other LLMs.

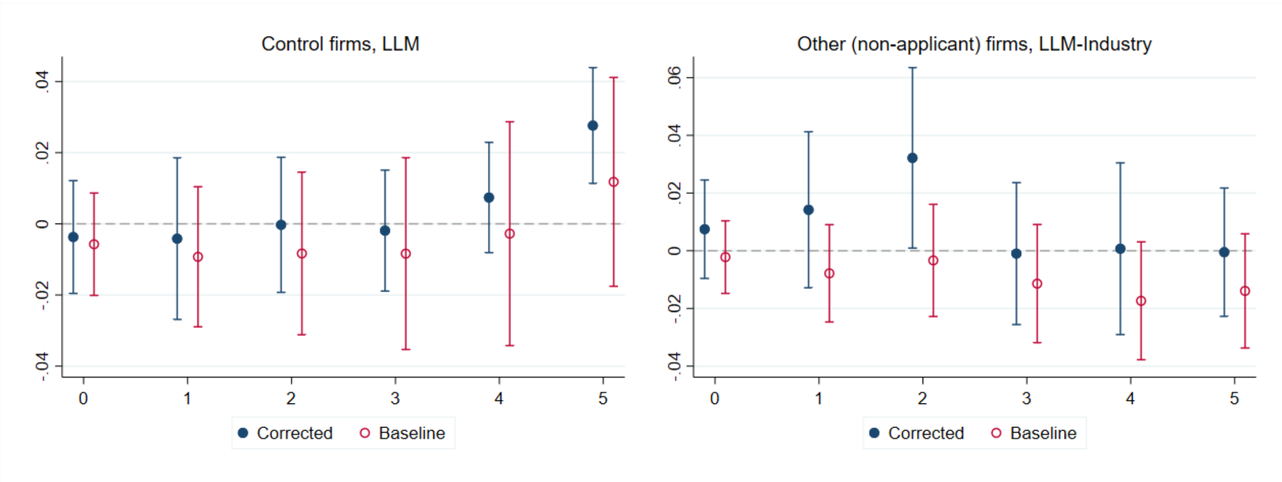
Figure C1i plots the estimated coefficients θ_k's for two different subsets of non-subsidized firms – respectively, applicant firms not obtaining the subsidy (left graph) and non-applicant firms in the same LLM-industry cell as subsidized firms.[1] Both graphs present baseline difference-in-differences estimates as well as "corrected" estimates accounting for the staggered research design, using the approach suggested by de Chaisemartin and D'Haultfœuille (2020). Overall, there is no evidence of significant spillover effects; the same is true when replacing the binary indicator $D_{m,t}$ in equation 1 with the (log of) funds actually paid to subsidized firms in each LLM or LLM-industry, see Figure C1ii.

These results imply that higher employment among subsidized firms reflects a net increase in aggregate employment, rather than a mere reallocation of jobs from non-subsidized to subsidized firms. Cerqua and Pellegrini (2022) reach the same conclusion by decomposing worker flows towards subsidized firms. Using worker-level data, they show that the majority of recruits come from new entrants in the labor market, and conclude that L488/92 subsidies generate few displacement effects across firms, if at all.

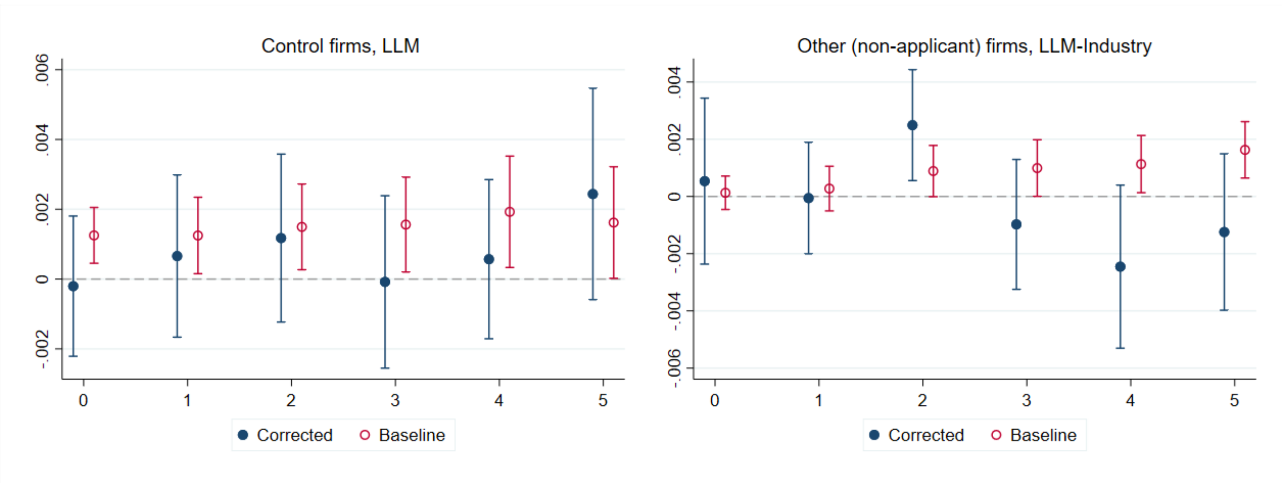
[1] Industry is defined at the 3-digit level.

Fig. C1: Spillover effects on other firms in the same labor market

(i) Using a binary indicator for presence of subsidized firms in LLM



(ii) Using the log of total subsidies in LLM



Notes: The graphs show the estimated spillover effects of the subsidy on local employment at different time horizons, indicated on the horizontal axis, and associated confidence intervals (at the 90% significance level). The left panel plots the aggregate employment response of control firms located in the same LLM as treated firms. The right panel focuses on non-participating firms in the LLM and (3-digit) industry as treated firms. The treatment variable is the log of funds received by treated firm in a LLM (or LLM-industry cell). "Baseline point" estimates and confidence intervals are obtained from specification 1 in the main text, clustering heteroskedasticity-robust standard errors by LMM. "Corrected" coefficients are obtained using the estimator proposed by de Chaisemartin and D'Haultfœuille (2020) to account for biases arising if group-time treatment effects are averaged with negative weights.

4 Politicians' response to (expected) objective scores

As explained in Section 2.2, local politicians attribute a discretionary score SD to projects depending the municipality within the region and industrial sectors in which they are realized, and the type of investment to be implemented. In addition, SD must be set ex-ante and communicated to the Ministry of Economic Development by October 30th of the year before each call was issued, and it is not circulated publicly.

When allocating the discretionary points by municipality-industry-type of project (SD), politicians may in principle take into account the *expected* score received by projects on objective criteria, call it SR^e – the actual score will only be revealed a few months later. In particular, they should attribute more points to projects that they favor and, at the same time, they expect to be scoring lower on objective criteria.

Letting Z denote the triple of projects' operating municipality, industry, and type of investment, and $\mathcal{Z}$ the set of possible values for such triple, politicians allocate SD across Z's taking into account the scores received on the objective criteria:

$$SD(Z, SR^e(Z)).$$

Therefore, SD depends on Z both directly and indirectly through SR^e, and we expect that SR^e enters negatively SD. In order to estimate such effect, we need to impose assumptions on how politicians form expectations SR^e, since our data allow us to observe ex-post realizations of SR in each call t but not the ex-ante expectations. In particular, we consider two alternative hypotheses.

(i) "Adaptive expectations": politicians form such expectations based on the average realizations of SR within each group $z \in \mathcal{Z}$ in the previous call $t-1$, $SR^e_{z,t} = \sum_{g(j)=z,t-1} SR_{j,t-1}/N^z_{t-1}$, where N^z_{t-1} denotes the number of projects in each group $z \in \mathcal{Z}$ in call $t-1$ and $g : \{1, 2, \ldots, N\} \rightarrow \mathcal{Z}$ is a function mapping rankings into groups.

(ii) "Perfect foresight": politicians are able to correctly predict the average realization of SR in group z in the call at time t ($SR^e_{z,t} = \sum_{g(j)=z,t} SR_{j,t}/N^z_t$).

In addition, we must impose that politicians' utility is stable over time. Under these assumptions, we can identify the effect of SR^e on SD across groups defined by Z leveraging longitudinal variation over subsequent calls and controlling non-parametrically for differences in projects' characteristics across groups through a full set of fixed effects:

$$SD_{z,t} = \beta SR^e_{z,t} + FE_z + FE_t + \varepsilon_{z,t}, \tag{2}$$

where FE_z and FE_t are fixed effects by group z and call, respectively. We weight the regression across Z's by the number of projects in each Z triple.

Estimates of equation (2) are reported in Table D1. When assuming adaptive expectations (columns 1-4), the estimated coefficient β is essentially zero; in particular, we can reject (with 95% confidence) effects as small as a -0.03 standard deviation changes in the discretional score for a one standard deviation increase in the objective score. When assuming perfect foresight (columns 5-8), the coefficient is more precisely estimated but remains extremely small in magnitude – we can reject effects larger than a -0.04 standard deviation changes in the discretional score for a one standard deviation increase in the objective score. The results are very similar when assuming that politicians predict the median – rather than the mean – of SR (columns 3 and 7); when running on the unweighted regression across Z's (columns 2 and 6); and when considering a quadratic specification (columns 4 and 8).

In the Supplementary Materials, we presents additional analysis, which allows for a flexible relationship between SD and SR^e. These additional results corroborate the evidence that SR^e is independent of SD.

Table D1: SD **response to** SR

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	expectations based on avg. SR in $t-1$				expectations based on avg. SR in t			
SR^e	-0.004	0.001	-0.005	-0.002	-0.024***	-0.023***	-0.025***	-0.026***
	(0.012)	(0.013)	(0.012)	(0.012)	(0.008)	(0.009)	(0.008)	(0.008)
SR^e, squared				-0.001				0.002
				(0.002)				(0.002)
Observations	3,083	3,083	3,083	3,083	8,471	8,471	8,471	8,471
N. Z-triples	1,284	1,284	1,284	1,284	3,518	3,518	3,518	3,518
Weighted	Y	N	Y	Y	Y	N	Y	Y
Statistic	average	average	median	average	average	average	median	average
Adjusted R^2	0.417	0.395	0.417	0.417	0.385	0.374	0.386	0.386

Notes: This table reports the results of regressions on equation (2) across groups of projects defined by triples of municipality-industry-type of project. The dependent variable is the sub-score SD attributed by politicians to each group. The main explanatory variable are politicians' expectations on the average score SR in each group $z \in \mathcal{Z}$. In columns (1)-(4) we assume that such expectations are based on the average realizations of SR within each group in the previous call $t-1$ ($SR^e_{z,t} = \sum_{g(j)=z,t-1} SR_{j,t-1}/N^z_{t-1}$, where N^z_{t-1} denotes the number of projects in group $z \in \mathcal{Z}$ during the call $t-1$ and $g : \{1, 2, \ldots, N\} \to \mathcal{Z}$ is a function mapping rankings into groups), while in columns (5)-(8) we assume that they are based on the average realizations of SR within each group in the call at time t ($SR^e_{z,t} = \sum_{g(j)=z,t} SR_{j,t}/N^z_t$). All regressions include fixed effects by group and call. Standard errors clustered by group are reported in parenthesis.

References

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma, "Simple local polynomial density estimators," *Journal of the American Statistical Association*, 2020, *115* (531), 1449–1455.

_, Richard K Crump, Max H Farrell, and Yingjie Feng, "On binscatter," *arXiv preprint arXiv:1902.09608*, 2022.

Cerqua, Augusto and Guido Pellegrini, "Decomposing the employment effects of investment subsidies," *Journal of Urban Economics*, 2022, *128*, 103408.

de Chaisemartin, Clément and Xavier D'Haultfœuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, September 2020, *110* (9), 2964–96.

Guiso, Luigi, Luigi Pistaferri, and Fabiano Schivardi, "Credit within the Firm," *Review of Economic Studies*, 2013, *80* (1), 211–247.

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

_ and Jeffrey M Wooldridge, "Recent developments in the econometrics of program evaluation," *Journal of economic literature*, 2009, *47* (1), 5–86.

Westfall, Peter H and S Stanley Young, *Resampling-based multiple testing: Examples and methods for p-value adjustment*, Vol. 279, John Wiley & Sons, 1993.