Pellegrini, Guido; Cerqua, Augusto

Conference Paper

Are the subsidies to private capital useful? A Multiple Regression Discontinuity Design Approach

51st Congress of the European Regional Science Association: "New Challenges for European Regions and Urban Areas in a Globalised World", 30 August - 3 September 2011, Barcelona, Spain

Provided in Cooperation with:
European Regional Science Association (ERSA)

Suggested Citation: Pellegrini, Guido; Cerqua, Augusto (2011) : Are the subsidies to private capital useful? A Multiple Regression Discontinuity Design Approach, 51st Congress of the European Regional Science Association: "New Challenges for European Regions and Urban Areas in a Globalised World", 30 August - 3 September 2011, Barcelona, Spain

This Version is available at:
http://hdl.handle.net/10419/120241
Are the subsidies to private capital useful? A Multiple Regression Discontinuity Design Approach

Abstract: There is still little consensus on the effectiveness of business support policies. The empirical evaluation is complicated by the difficulty in achieving reliable identification. We analyse the impact of Law 488/92, the main Italian regional policy. We propose a new approach, named multiple regression discontinuity design that exploits the sharp discontinuities in the L488 rankings and extends the RDD approach to a context where the treatment is assigned by multiple rankings with different cut-off points. We find that the impact of L488 on investment and production of the financed firms is positive and statistically significant.

Keywords: multiple regression discontinuity design, policy evaluation, public subsidies, regional policy.

JEL: R38, C14, H71

1. Introduction

Business support programs are popular industrial policies used by most governments in EU and US, in order to foster competitiveness, self-sustaining growth and employment, particularly in disadvantaged areas. A huge amount of funds are spent each year on regional policies and subsidies or “State Aid”. Moreover, aid granted by governments in response to the recent crisis contributed to the big increase in State aid to industry and services in the last year: in 2009, the share of State aid for industry and services as percentage of GDP in the EU-15 economies was 3.7%, 0.5% excluding crisis measures. Not surprisingly, several studies have evaluated the extent of the economic payoff of these kinds of incentives (see, inter alia, Schalk and United, 2000; Roper and Hewitt-Dundas, 2001; Bondonio, 2004; Rodrik, 2007; Criscuolo et al., 2009). Despite this vast literature, there is little consensus among economists on the effectiveness of investment incentives. As well argued in Bondonio (2009), evaluating the impact of business incentive programs is still a challenging task. The problems are not only related to limitations in data availability, but (and mainly) due to difficulties faced when isolating the effects of subsidies from the confounding effects induced by other factors. Endogenous participation and high selection bias in the programs require implementing an accurate micro-econometric evaluation of their causal effects, but this is rare primarily because of the difficulty of achieving credible identification (Criscuolo et al., 2009).

In this paper we present a robust econometric analysis of the causal effect of capital subsidies to private firms, by exploiting an unusual characteristic of an important regional policy in Italy. We analyse the impact of subsidies distributed by Law 488/92 (henceforth L488), which has been the main policy instrument for reducing territorial disparities in Italy, and represented the 30% of total

---

1 We are grateful to Raffaello Bronzini, Alessio D’Ignazio, and Guido de Blasio from the Bank of Italy for useful comments and suggestions. We also thank Sergio Gison from the Ministry of Industry for making available the administrative L488 data set.

2 UE COMMISSION (2010).
financial aid to firms in Italy in the period 1996-2006. This law has been characterized by rigorous and transparent selection procedure. Each year, subsidies are allocated to a broad range of investment projects through regional “calls for tenders”, which mimic an auction mechanism. In each regional “call for tender”, the investment projects are ranked on the basis of a score that depends on a number of (known) characteristics of both the project and the firm. Projects receive subsidies according to their position in the ranking until the financial resources granted to each region are exhausted.

The presence of sharp discontinuities in the L488 rankings allows us to use a quasi-experimental method deriving from a regression discontinuity design (henceforth RDD) approach, capable of overcoming the referred difficulties. The RDD technique compares the economic situation arising under policy interventions with a scenario representing the hypothetical situation where L488 does not exist. With a RDD approach, we can estimate the L488 funds effect if the treatment is determined by whether an observed forcing variable (the score determining the position of the project in the regional ranking) exceeds a known cut-off point. The critical assumption of this method consists in the faculty of the forcing variable to determine the selection process by itself. In the case of L488, this assumption in completely plausible, in effect the forcing variable (the score) is the only variable that determines the allocation of the subsidies. Nonetheless, the standard RDD approach is not appropriate in evaluating the effectiveness of L488, because of the particular subdivision in regional “call-for-tenders”. Indeed, the L488 selection mechanism does not determine a single ranking with one cut-off point, but every region in every year has got his specific ranking with a cut-off point of the forcing variable depending on the number of applicants and the amount of aid requested by the firms.

This peculiar configuration of L488 requires modifying the standard RDD technique. We propose a novel approach, called multiple regression discontinuity design (henceforth MRDD), that extends the RDD to a context where the treatment is assigned by multiple rankings with different cut-off points.

Our analysis is based by and large on the same data sets used by Bronzini and de Blasio (2006), and it covers also the same period (1995-2001). In order to strengthen the relevance of the aggregated estimates, we compare our results with the ones obtained by pooling the data sets and using parametric and nonparametric estimators. Moreover, different robustness tests are presented as suggested by Imbens and Lemieux (2008).

Using the MRDD approach, the impact of L488 on the growth of the financed firms is positive and statistically significant: the investment increases from 6.5 to 7.7 percentage points higher every year in the subsidized firms than in non-subsidized ones, the turnover increases yearly from 7.5 to 10.5 percentage points higher in favour of the subsidized firms, over the period 1995-2001.

The paper has been organised in the following way: the next section summarizes the literature and available empirical evidence on the effects of L488. Section 3 describes L488 procedures in more detail. The evaluation method is discussed in Section 4, followed by a presentation of the data set in Section 5. The results of the empirical analysis are discussed in section 6, while section 7 assesses the robustness of the results. Finally, we briefly conclude and define some path for future research.

3The aggregate data set is a pooled version of the different datasets by regions and years, in which the observations of each ranking under analysis have been first normalized, and then added up to produce a single ranking with only one cut-off point (see Section 4).
2. The previous literature

Recently, researchers have shown an increased interest in the problem of evaluating the impact of public subsidies to private firms. Some analyses suggest that regional capital incentives can induce additional investment in subsidized firms (Daly et al., 1993; Faini and Schiantarelli, 1987; Harris, 1991; Schalk and United, 2000; Criscuolo et al., 2009); others argue that intertemporal substitution effects prevail (Bronzini and De Blasio, 2006). Also, the employment impact of capital subsidies is doubtful (Gabe and Kraybill, 2002). Finally, the effect of subsidies on efficiency and productivity seems negligible or negative (Lee, 1996; Bergstrom, 2000; Harris and Trainor, 2005; Criscuolo et al., 2009).

Concerning the effectiveness of L488, after more than ten years of policy intervention, empirical evidence remains mixed and contradictory. A positive effect of L488 on investment is found in the Ministero dell’Industria (2000). Pellegrini and Carlucci (2001) and Carlucci and Pellegrini (2003) present empirical evidence of a positive employment effect, using different parametric and non-parametric approaches. Adorno, Bernini and Pellegrini (2007) highlight a positive, but U-reversed relationship between the amount of subsidies and employment and product. Bernini and Pellegrini (forthcoming) evidence higher growth in output, employment and fixed assets in subsidized firms but a lesser increase in Total Factor Productivity than in unsubsidized firms. The presence of a modest spatial crowding out, where subsidized regions attract employment and firms from neighboring areas, is tested in De Castris and Pellegrini (2011). Bronzini and De Blasio (2006) investigate the presence of cross-sectional substitution (subsidized firms may receive some of the investment opportunities that non-subsidized firms would have otherwise had in absence of the incentives) and intertemporal substitution (firms may have brought forward investment projects originally planned for the post-intervention period in order to take advantage of the incentives). The Authors find that financed firms have substantially increased their investments when compared with the pool of firms whose applications have been rejected. They also show that these firms have significantly tapered their investment activity in the years following the program, confirming the L488 time substitution effect (see also Bernini and Pellegrini, forthcoming).

3. Law 488/92

Italy is among the European countries with the highest inequality in the distribution of the wealth between different areas. In the recent years, different instruments have been introduced in order to aid firms located in lagging areas. In 1996 the Ministry of Industry has introduced L488, which has been the main policy instrument for reducing territorial disparities in Italy. In the period 1996-2006, roughly 44,000 projects (over €23 billion) have been financed by L488. Most of this funding has targeted Mezzogiorno, which comprises the least developed regions of Italy.⁴

L488 has been introduced in order to replace the ‘extraordinary intervention’ which was directed to all investing firms located in the Southern regions. As a result, in Italy until 1995, subsidized projects have never been selected basing on economic parameters and criteria. L488 has changed the selection procedure, allocating subsidies through a rationing system based on “calls for

⁴ L488 operates in the less developed areas of Italy. These areas are either designed as Objective 1, 2 or 5b for the purpose of EU Structural Funds or subject to exemptions from the ban on state subsidies. Objective 1 refers to the regions suffering from general underdevelopment, as reflected in GDP per capita that is less than 75% of the EU average. Objective 2 is related to regions suffering from a concentration of declining industries, as reflected in higher average unemployment, higher dependency on industrial employment and observable job losses in specific industries. Objective 5b includes predominantly peripheral rural regions, as reflected in a high share of agricultural employment and a low level of agricultural income.
tender”, which mimic an auction mechanism, which guarantees compatibility of demand and supply incentives.

L488 makes available grants on capital account for projects designed to build new productive units in less developed areas\(^5\), or to increase production capacity and employment, increase productivity or improve ecological conditions associated with productive processes, technological updates, restructuring, relocation and reactivation.

The Italian Ministry of Industry stands over the whole selection process and determines the specific deadline of each auction. After receiving the application form including the technical report and business plan, the relevant authority performs a preliminary screening, evaluating the funding eligibility of the project. Within four months of the deadline, the Ministry of Industry publishes the rankings. The law requires that firms awarded assistance receive the first annual instalment within two months. The amounts awarded are paid out in three equal instalments\(^6\). The second and the third instalments are paid on the same date in subsequent years.

Incentives are allocated on the basis of regional competitive auctions. In each auction the investment projects are ranked on the basis of five objectives and predetermined criteria:

1) the share of owners’ funds on total investment;
2) the new job creation by unity of investment;
3) the ratio between the subsidy requested by the firm and the higher subsidy applicable;
4) a score related to the priorities of the region in relation to location, project type and sector;
5) a score related to the environmental impact of the project.

The criteria 4 and 5 were introduced starting from the 3\(^{rd}\) auction. The five criteria carry equal weight: the values related to each criteria are normalized, standardized and added-up to produce a single score that determines the place of the project in the regional ranking (this normalized score is the forcing variable used in the following analysis). The rankings are drawn up in decreasing order of the score awarded to each project and the subsidies are allocated to projects until funding granted to each region is exhausted\(^7\).

Yearly checks are made to control if subsidizes firms respect their targets (in terms of new employment faced by the financed investment). If the firm doesn’t reach its goals, the financed investment would be revoked and the firm has to return back the total amount received by the L488. A tolerance threshold (30% of the total new employment) has been introduced to provide a slight flexibility to firm investment.

L488 auctions have been issued on a yearly basis. Our analysis refers to the period 1995-2001 and focuses on three of the four L488 auctions that were concluded by 2001. For these auctions the treatment started and finished within the time span provided by the financial-statement data. The timing of the assistance by auction is presented in Table 3.1:

---

\(^5\) As will be explained in Section 5, we did not consider startups in our analysis.

\(^6\) Only two instalments if the project is completed within 2 years.

\(^7\) There are also special rankings for large projects and reserved lists for small and medium-sized firms.
Tab. 3.1 – Timing of the assistance

<table>
<thead>
<tr>
<th>Auction</th>
<th>Application deadline</th>
<th>Presumed time of the 1st instalment</th>
<th>Presumed time of the 2nd instalment</th>
<th>Presumed time of the 3rd (last) instalment</th>
</tr>
</thead>
</table>

Source: Bronzini and de Blasio (2006)

However, in several cases administrative complications and technical and economic problems have increased the time span of the project (estimated in 3.6 years by Bernini and Pellegrini (forthcoming)).

The data relative to the auctions derive from two different data sets: the administrative L488 data set of the Ministry of Industry and a financial-statement data set.

The estimation results that we present below are based on the assumption that there are no other governmental programs correlated with the allocation of L488 funding. Actually, a feature of L488 minimizes the extent of this bias, requiring that firms applying for the incentives have to give up to other public subsidies, without any guarantee of receiving the L488 funds.

4. The econometric evaluation procedure

If the L488 funds had been allocated by a random process, an optimal way to evaluate the impact of the subsidies would have been a simple difference between the outcomes of treated and not treated firms for each auction. Unfortunately the assumption of random assignment is not credible when the policy instrument (such as L488) determines a deliberate selection process. If a support programme selects firms in a non-random way, the participation is endogenous and the projects are heavily selected. In order to avoid the selection bias and the omitted variable bias, the policy effect should be measured as the difference between the result of a group of firms composed by financed firms and the result of the same group in case the policy would not be existed. Obviously the data relative to the latter group of firms are not directly available; therefore the challenge is to find a valid control group.

In L488 case, the data regarding the firms applied for the incentives but not financed since they scored low in the L488 ranking are available. These non-treated firms are firms willing to invest which have a valid investment project checked by a preliminary screening. As a consequence, in each ranking we can consider these firms as the best control group available; in fact, as suggested by Brown et al. (1995), they show a propensity for investment very similar to that of subsidized firms.

The particular configuration of the L488 dataset - for each auction, there are as many rankings as the number of regions involved, and each ranking has a different cut-off point - is similar to the one faced by Smith et al. (2007) in analysing the reemployment services system in Kentucky, and the one addressed by Gamse et al. (2008) in evaluating the impact of Reading First, which is a US federal education program. In these papers and in the empirical literature in general, there exist two different approaches for exploiting the RDD in order to estimate the treatment effect across different rankings. One consists in two different steps: first, estimating the treatment effect for each ranking; second, pooling the treatment effects in order to get the global treatment effect of
the policy under analysis. The other approach pools observations from different rankings into a single dataset re-centering and standardizing the forcing variable. Both methods compare the economic scenario arising under policy interventions with what would have happened if the policy was not implemented. We have chosen the first one because it exploits all the information available from the dataset, increasing efficiency, but it relies on weaker assumptions than the second approach\(^8\). However, we exploit the pooling approach as a robustness test.

The methodological approach we use can be named multiple rankings regression discontinuity design (MRDD). As we have a particular configuration of the L488 data set - for each auction, there are as many rankings as the number of regions involved - we extend the RDD to a context where the treatment is assigned by multiple rankings with different cut-off points. The main assumption is that, in each ranking, the best control group for the units just above the cut-off point is represented by the firms ranked just below the cut-off point (the ones that are not treated). Only the not treated firms in the same ranking are the appropriate counterfactual. However, we also assume that the ATT does not depend on the level of the forcing variables, i.e. on the differences in the cut-off points\(^9\). In this case, aggregating the disaggregated estimates, this method exploits all the available observations of the L488 merged data set; this feature of the MRDD improves the efficiency of the estimation process, making the resulting ATT more reliable.

The MRDD consists in two different steps; firstly, we apply a sharp RDD\(^10\) in each ranking, exploiting the sharp discontinuity determined by the forcing variable; in this way, we obtain a nonparametric estimation of the local ATT in each ranking. Secondly, we aggregate the different first step estimates by a weight structure, where the weights are based on the share of treated units in each ranking. The MRDD estimator mimics a matching estimator in a RDD context: we can define each ranking as an “homogeneous strata”, determine the best matching in each ranking using a RDD approach and computing outcome differences within “strata” using a specific cut-off point in each ranking, and finally integrate such differences over the distribution of ranking in the treatment population to retrieve the global ATT.

This method is perfectly operable, because each cut-off point is not known a priori by the firms applied for the L488 incentives. As a consequence, there is not any degree of sorting and, as suggested by Lee (2008), the actual allocation of subsidies among the group of applicant firms is randomised in a neighbourhood of each cut-off.

Obviously, we can apply the MRDD on the L488 data set only if the forcing variable (the score equal to the sum of the normalized indicators) is available. In this procedure the indicators are the selection variables, i.e. the indicators can explain most part of the differences between the group of subsidized firms and the group of non-subsidized firms; thereby the forcing variable is available, and we can reconstruct the selection process, estimating the selection effect in the control group.

The use of MRDD for assessing the impact of L488 is however complex, because of the limited number of observations near each cut-off point, which creates a trade-off between the interval extension around the cut-off point and the statistical precision of estimates. Another potential

---

\(^8\) Pooling observations from different rankings could cause biases in the ATT due to the remarkable differences that are possible between treated firms of different rankings.

\(^9\) In order to test this assumption, we ran a regression to evaluate the potential dependence of the nonparametric estimates of each ranking to the different cut-off points. The outcome of the regression is not statistically significant at the 10% level, strengthening the reliability of this assumption.

\(^10\) In the sharp design, the treatment assignment is a deterministic function of the forcing variable. See Lee and Lemieux (2010).
problem is the correlation among the auctions. However, the construction of the merged data set has been done in the perspective of minimizing the interaction among the auctions, hence this key assumption can be considered quite reliable.

It could be possible that one firm which wins a project in one auction has some kind of positive externality on other firms (e.g. increased supply) which then increases the choice probability of related firms in the subsequent auction. In our analysis, the auctions considered, were very close in time with respect to the average project time span. Therefore, even if supply or demand spillovers are theoretically possible, from an empirical point of view they can be negligible. We have also some empirical evidence: a recent paper\textsuperscript{11} has explored the presence of spillovers generated by subsidized firms considering the main regional development policies in the Southern regions of Italy (L488 and Contratti di Programma\textsuperscript{12}). Results show that spillovers are small and negative across areas (suggesting the presence of a modest spatial crowding out, where subsidized regions attract employment and firms from neighbouring areas).

In this paper, we use the MRDD approach for estimating the ATT of L488 on the economic growth of the subsidized firms. In particular we use two different dependent variables:

1) Yearly growth rate of the share of investment on turnover;

2) Yearly growth rate of turnover.

By comparing the average yearly growth rate of the share of investment on turnover and the average yearly growth rate of turnover of firms receiving the L488 funds and non-beneficiaries at the margin, we can control for confounding factors and identify the ATT locally at each threshold.

with depending on the

the sensitivity of the results to the estimator and the bandwidth in the nonparametric case, we will compare the MRDD results with the ones obtained through other two different analyses, in order to evaluate the robustness of our conclusions.

These additional analyses concern a joined version of the dataset, in which the observations have been first normalized, and then added up. This aggregation has been carried out re-centering and standardizing the forcing variable to create a dataset with a unique cut-off point (where every cut-off point is equal to zero), in order to analyse the general ATT of the subsidies with a simple sharp RDD. We analyse this aggregated dataset both with a parametric and a nonparametric method. The resulting estimates are meaningful under the assumption of a random allocation of the subsidies to the firms ranked around the unique cut-off point. The latter one is a very strong assumption; thereby we use these estimates merely to strengthen the relevance of the main analysis.

Despite the use of two different methods (the sharp MRDD applied to the disaggregated data set and the sharp RDD applied to the aggregated dataset), inference is still complex. Here, we use local linear regressions with standard errors computed with the bootstrap\textsuperscript{13} method in nonparametric analyses, and the OLS estimator with robust standard errors in parametric

\textsuperscript{11} De Castris and Pellegrini (2011).
\textsuperscript{12} Regional policy designed for foreign firms with large projects.
\textsuperscript{13} Bootstrapping is the practice of estimating properties of an estimator (such as mean or variance) by measuring those properties when sampling from an approximating distribution.
regressions, as suggested by Imbens and Lemieux (2008). Finally, we use the following robustness tests (Imbens and Lemieux, 2008):

1) the presence of a discontinuity in the density function of each ranking at the relative cut-off point which could signal the existence of manipulations in the forcing variable by the firms;

2) the presence of other discontinuities in the forcing variables of each ranking, which will make weaker the assumption that each discontinuity is an effect of the subsidies;

3) the presence of exogenous covariates with discontinuities at each cut-off point, which could determine the discontinuity of each outcome.

5. Data and methodological issues

Our econometric analysis is based on the integration of two different data sets: the administrative L488 data set of the Ministry of Industry and a financial-statement data set, collecting data from AIDA and other sources of financial information. The first one records all the firms applied for at least a L488 auction, both financed and non-financed. It provides us with information that is important for our analysis, such as the firm ranking at the regional level and the timing of the instalments. Unfortunately this data set lacks of financial and economic information, such as investment and turnover; therefore we also need to use a financial-statement data set, which collects financial statements only for corporations. The integration between these two data sets has required a complex process of cleaning and merging. Combining these data sets permits us to compare the change in the participating firms’ performance to a control group of firms applied for the incentives, but not financed.

The financial-statement data set used in our analysis extends from 1995 to 2001, allowing us to study the impact of the program over a period that includes pre-intervention as well as post-intervention years of the first four auctions of L488. For these auctions the treatment started (with the 1st instalment) and finished (with the 3rd instalment) within the time-window provided by the financial-statement data. We focus below on the 2nd, the 3rd, and the 4th auctions. These auctions are ideal for our purposes since they occurred within the time span under analysis, thus providing us with pre and post-intervention observations. The 1st auction has been excluded because it included a transitory clause which allowed firms not eligible under the 1st auction of L488 to be financed as well. We use data referring to the period 1995-2001 in order to compare our estimates with the literature on L488.

By linking the L488 data set with the financial-statement data set, we reconstruct a merged data set over the period 1995–2001 for 2,044 firms applied for “call of tenders” in the South of Italy. The time horizon is 6 years and it ranges from 1995 to 2001, in order to verify the growth of the dependent variables in a wide period of time. We have also decided to use this six-year time span to verify Bronzini and de Blasio’s hypothesis (2006) about the presence of intertemporal substitution.

Because of technical and logical motivations we have established that, in order to carry out the econometrical analysis for each ranking, the presence of at least ten subsidized and ten non-

---

14 For this reason, it is skewed toward larger firms.
15 We merged the data sets using the fiscal and number of commerce codes as firm identifiers.
16 It is possible that subsidized firms don’t make additional investments, but they just bring forward investment projects originally planned for future periods.
subsidized observations\textsuperscript{17} is a pre-requisite. This criterion has caused the impossibility to evaluate Basilicata and Molise; additionally it has been impossible to analyse Sicily in the 2\textsuperscript{nd} auction, Abruzzi, Calabria, and Sardinia in the 4\textsuperscript{th} auction\textsuperscript{18}, consequently the rankings under analysis are fourteen (five for the 2\textsuperscript{nd} auction, six for the 3\textsuperscript{rd} auction, and three for the 4\textsuperscript{th} auction).

After verifying that the cleaning and integration procedures do not have a different impact on financed projects and control group, the attention focused on the final data set on which the evaluation model has been implemented\textsuperscript{19}. It consists of 519 financed projects and 1525 non-financed projects over the period 1996-1999.

6. Results

The estimation procedure starts with some graphical evidence. A simple way to evaluate the effect of L488 is to plot the relation between each outcome variable (yearly growth rate of the share of investment on turnover and yearly growth rate of turnover) and the forcing variable (sum of the indicators normalized) by firms on either sides of the cut-off point. With the MRDD have been analysed fourteen different rankings and it is impractical to graphically represent each and all; thereby in this section we present two of the most representative rankings, Campania and Puglia in the 2\textsuperscript{nd} auction. Fig. 6.1 and 6.2 plot the respective rankings for both the dependent variables in the period 1995-2001 by subsidized firms against the non-subsidized ones. In each graph, the cut-off line sharply separates treated and not treated firms. Each figure superimposes the fit of a nonparametric flexible polynomial regression model, together with the 95\% confidence bands.

\textbf{Fig. 6.1 – Yearly growth rate of the share of investments on turnover and yearly growth rate of turnover in Campania for the 2\textsuperscript{nd} auction with the 95\% confidence bands}

\textsuperscript{17}Such criterion allows us to limit the problems due to sample of firms too restricted in order to be analysed by an econometric method.

\textsuperscript{18}In order to verify the entity of the loss of these observations, we tested if the integration of these observations into the aggregated data set modifies significantly the estimates of the ATT of both the dependent variables. The results obtained from this widened aggregated data set support the hypothesis that the loss of these observations is not critical; in fact they statistically do not differ from the ones obtained on the aggregated data set used in the analysis.

\textsuperscript{19}We consider two different data sets (one per each dependent variable) and in each one, we trim the sample at the 5 and 95 percentiles respectively to the relative dependent variable.
Fig. 6.2 – Yearly growth rate of the share of investments on turnover and yearly growth rate of turnover in Puglia for the 2nd auction with the 95% confidence bands

Fig. 6.1 and 6.2 clearly show that, on average, the subsidized firms grow more than the others, and the effect is statistically significant at the 5% level. As a matter of fact, the presence of a discontinuity at the cut-off point is supported by every graph. Each nonparametric regression line shows a positive jump moving from the non-financed firms to the financed ones. The descriptive evidence, using a graphical representation, suggests that there are discontinuities in the investments and turnover growth between treated and not treated firms.

In order to provide some graphical evidence for the entire merged data set, we present other two graphs relative to the aggregated data set, in which we exploit all the available observations. Higher the number of observations, higher the precision of the estimates; so, in the following graphs we expect a clear distinction between the confidence interval of the non-subsidized firms and the confidence interval of the subsidized ones, in case the financed firms made additional investment because of L488. Fig. 6.3 presents the yearly growth rate of the share of investment on turnover, whilst in Fig. 6.4 is represented the yearly growth rate of turnover.

Fig. 6.3 – Yearly growth rate of the share of investments on turnover for subsidized and non-subsidized firms in the aggregated sample with the 95% confidence bands
The graphical analysis displays a possible and systematic difference between financed and non-financed firms. To evaluate the extent of these differences, we use the MRDD on the disaggregated sample. The standard approach is to use a local linear regression, which minimizes the bias (Fan and Gijbels, 1996).

There are two key issues in implementing the estimation by a local linear regression: the choice of the kernel and the choice of the bandwidth.

Different kinds of kernel are available.\textsuperscript{21} We present our results using three different kernel (triangular, Gaussian, and Epanechnikov).

A very delicate part of the analysis is the choice of the bandwidth. In a nonparametric RDD estimation, it involves finding an optimal balance between precision (more observations are available to estimate the regression) and bias (larger the bandwidth, larger the differences between treated and not treated firms). Smaller bandwidths are feasible if the number of observation is reasonably high. There are several rule-of-thumb bandwidth choosers, but none is completely reliable. A recent contribution of Imbens and Kalyanaraman (2010) presents a data-dependent method for choosing an asymptotically optimal bandwidth in the case of a RDD.

However, different bandwidth choices are likely to produce different estimates. We decided to report three estimates as an informal sensitivity test: the first one using Imbens and Kalyanaraman formula (the optimal bandwidth), the other ones reducing of 25% and increasing of 50% the optimal bandwidth. The standard errors are estimated by a bootstrap procedure.\textsuperscript{22}

\textsuperscript{20} Each point is an average of the dependent variable for each interval, which is 0.05 wide.
\textsuperscript{21} It has been shown in the statistics literature that a triangular kernel is optimal for estimating local linear regressions at the boundary (Fan and Gijbels, 1996), and therefore has good properties in the RD context. However, while other kernels (Gaussian, Epanechnikov, etc.) could also be used, Lee and Lemieux (2010) argue that the choice of kernel typically has little impact in practice (see also Imbens and Lemieux, 2009). The statement is basically true also in our case.
\textsuperscript{22} In the bootstrapping procedure, for the rankings with more than 50 observations we used 100 repetitions, while for the rankings with less than 50 observations we used 200 repetitions.
The results of the application of the method we propose (the sharp MRDD) are presented in Tab. 6.1 and 6.2. As exposed in Section 4, this method consists in a disaggregated analysis of the datasets, which provides nonparametric estimates of the dependent variables for every region in each auction. In the first step we carry out every analysis for each ranking; in the second step we apply the MRDD, which aggregates the first step estimates by a weight structure, where the weights are based on the share of treated units in each ranking, in order to retrieve the global ATT of the L488 funds, and the global results are shown in the tables below.

The optimal bandwidths derive from a weighted procedure of the optimal bandwidths computed for each ranking, and for this reason they are not numerically reported.

### Tab. 6.1 – MRDD applied on the yearly growth rate of the share of investment on turnover (Local Linear Regression of the differences between treated and not treated firms).

<table>
<thead>
<tr>
<th>Bandwidth</th>
<th>Triangle Kernel</th>
<th>Epanechnikov Kernel</th>
<th>Gaussian Kernel</th>
</tr>
</thead>
<tbody>
<tr>
<td>75% opt. bw.</td>
<td>0.0823</td>
<td>0.0743</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>(0.022)***</td>
<td>(0.016)***</td>
<td>(0.015)***</td>
</tr>
<tr>
<td>opt. bw.</td>
<td>0.0771</td>
<td>0.0656</td>
<td>0.0653</td>
</tr>
<tr>
<td></td>
<td>(0.0184)***</td>
<td>(0.014)***</td>
<td>(0.015)***</td>
</tr>
<tr>
<td>150% opt. bw.</td>
<td>0.0876</td>
<td>0.0589</td>
<td>0.0557</td>
</tr>
<tr>
<td></td>
<td>(0.0163)***</td>
<td>(0.0124)***</td>
<td>(0.0132)***</td>
</tr>
</tbody>
</table>

Note: Bootstrapped Standard Errors in parentheses. *, **, *** = significant at 10%, 5%, 1% respectively.

### Tab. 6.2 – MRDD applied on the yearly growth rate of turnover (Local Linear Regression of the differences between treated and not treated firms).

<table>
<thead>
<tr>
<th>Bandwidth</th>
<th>Triangle Kernel</th>
<th>Epanechnikov Kernel</th>
<th>Gaussian Kernel</th>
</tr>
</thead>
<tbody>
<tr>
<td>75% opt. bw.</td>
<td>0.1405</td>
<td>0.1298</td>
<td>0.1128</td>
</tr>
<tr>
<td></td>
<td>(0.046)***</td>
<td>(0.035)***</td>
<td>(0.028)***</td>
</tr>
<tr>
<td>opt. bw.</td>
<td>0.1051</td>
<td>0.0882</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>(0.028)***</td>
<td>(0.025)***</td>
<td>(0.022)***</td>
</tr>
<tr>
<td>150% opt. bw.</td>
<td>0.1032</td>
<td>0.0921</td>
<td>0.0796</td>
</tr>
<tr>
<td></td>
<td>(0.027)***</td>
<td>(0.025)***</td>
<td>(0.023)***</td>
</tr>
</tbody>
</table>

Note: Bootstrapped Standard Errors in parentheses. *, **, *** = significant at 10%, 5%, 1% respectively.

The results obtained carrying out the sharp MRDD show that for each kernel and for every bandwidth used, the effect of the L488 policy is positive and statistically significant at the 1% level for both the dependent variables. Using the optimal bandwidth, the investment increases from 6.5 to 7.7 percentage points higher every year in the subsidized firms than in non-subsidized ones, and the turnover increases yearly from 7.5 to 10.5 percentage points higher in favour of the subsidized firms, over the period 1995-2001.

The aggregation process of the sharp MRDD has made possible the computation of the effect of L488 on investment and turnover growth also for each single auction. Concerning the yearly growth rate of the share of investment on turnover, the estimates show a positive effect of L488
for every auction, but this effect is statistically significant strictly for the 2\textsuperscript{nd} and the 4\textsuperscript{th} auction. With reference to the turnover, the estimates still show a positive effect of L488 for each auction, nevertheless this effect is statistically significant at the 1% level only for the 2\textsuperscript{nd} auction. Possibly the lack of statistical significance in some single auctions is due to the smaller number of firms (in presence of high variability, as common in firm’s performance analysis), which affects the statistical significance of the estimates.

Given the sensitivity of the results to the estimator and the bandwidth in the nonparametric case, we will compare the MRDD results with the ones obtained through two additional analyses (both exploiting the sharp RDD) of the aggregated sample: a nonparametric one and a parametric one.

The results obtained carrying out the sharp RDD with the aggregated sample are rather similar to the ones obtained with the sharp MRDD. In fact using the optimal bandwidth, the difference between the two groups of firms is 6-7 percentage points for the yearly growth rate of the share of investment on turnover and 8-10 percentage points for yearly growth rate of turnover. The other bandwidths show similar results for each dependent variable.

In these analyses we have used the whole merged data set, composed by 2,044 observations\textsuperscript{23}. Given the nature of the RDD, could be appropriated to shrink the sample in order to compare exclusively the observation next to the cut-off point, both for financed and non-financed firms. We have carried out other two analyses: the first one on a sample reduced by 50% of the observations; the second one on a sample reduced by 75% of the observations. In both analyses, the proportion between treated and not treated firms has been kept unchanged.

Also with these analyses, the estimates have been computed for three different bandwidths and three different kinds of kernel for each dependent variable.

The estimates with the reduced samples, as well as the previous estimates, show a positive and statistically significant effect of L488 on subsidized firms, for both the dependent variables. Respective to the results obtained with the MRDD, these estimates are, on average, somewhat higher both for investment and turnover.

We have tried to consolidate the significance of our estimates also using a parametric analysis of the aggregated sample. The parametric approach can integrate the nonparametric one, assessing the robustness of the RDD (in our case MRDD) estimates of the treatment effect.\textsuperscript{24} The choice of the order of the polynomial can be assessed using some goodness-of fit criteria, like the well-known Akaike information criterion (AIC) of model selection or the Bayesian information criterion (BIC), where the penalty for additional parameters is stronger than that of the AIC.\textsuperscript{25} The adoption of these criteria corresponds to use a generalized cross-validation procedure.

The results of OLS estimates with heteroskedasticity-robust standard errors on the full sample, adding different polynomials, Show that the effect of the subsidies is positive, statistically significant at the 1\% level, equal to 5 percentage points per year for both dependent variables, smaller than in the MRDD.

\textsuperscript{23} As pointed out in Section 5, this sample of firms has been trimmed at the 5 and 95 percentiles respectively to the relative dependent variable.

\textsuperscript{24} Lee and Lemieux (2010) argue that, in the case of polynomial regressions, the equivalent to bandwidth choice in the nonparametric regression is the choice of the order of the polynomial regressions. Therefore it is advisable to try and report a number of specifications to see to what extent the results are sensitive to the order of the polynomial.

\textsuperscript{25} Schwarz (1978).
7. Robustness Proofs

Following Imbens and Lemieux (2008), we assess the robustness of our results adopting various specification tests:

- Testing for possible discontinuities in the conditional density of each forcing variable (sum of the indicators normalized) at the relative cut-off point;
- Testing whether the outcomes are discontinuous not only at the cut-off point but also at other values of the forcing variable for each dependent variable;
- Testing for possible jumps in the value of other exogenous covariates at each cut-off point.

Testing for discontinuities in the conditional density of each forcing variable at the relative cut-off point is related to the possibility of manipulations of the forcing variable. If firms can manipulate the forcing variable in order to obtain desirable treatment assignments (that is, in our case, they have a great deal of control on some of the indicators), one would certainly expect firms on one side of the cut-off to be systematically different from those on the other side. However, Lee (2008) shows that if individuals do not have precise control over the forcing variable, variation in treatment status will be randomized in a neighbourhood of the cut-off. In this case the RDD (in our case the MRDD) can be considered “as good as” a local randomly assignment. In the allocation of the L488 funds, the selection process leads to a high degree of uncertainty over the assignment results. The cut-off point is not fixed, and it depends on the region and the auction under analysis. This cut-off point is known only after the availability of the data referring to all firms. In addition, the Ministry of Industry has a strict control over the procedure estimating the indicators.

The evidence of a jump in the conditional density of each forcing variable can be a test of the imprecision of control over the forcing variable, as suggested in McCrary (2008): if there is some degree of sorting of the firms around the threshold, the appropriateness of the RDD (in our case MRDD) in this contest is dubious. In Figure 7.1 we present a formal test of manipulation related to continuity of the forcing variable density function proposed by McCrary (2008). We report the graphs for the rankings of the second auction, but the other rankings analyzed show similar results. For each ranking, the weak discontinuities around the cut-off point are not statistically significant.
Fig. 7.1 – Estimated densities of each forcing variable at the relative cut-off (2nd auction)

Another robustness test verifies that there are no extra jumps in the levels of the outcomes, where no hypothesized cut-off exists. The approach used here consists of testing for a nil effect in different points of the forcing variables.

Testing the effects of the subsidies using the triangular kernel and the optimal bandwidth for both the dependent variables we found that most of the statistically significant discontinuities are relative to the hypothesized cut-off point (i.e. zero), for both the dependent variables.

An important test regarding the assumptions underlying the sharp MRDD consists in verifying that there are no jumps at the cut-off point in variables that should not be affected by the treatment. The absence of discontinuities around the threshold supports the causality relation between the jump in the outcome variable and the treatment. We look at possible jumps in the value of two exogenous covariates - a financial one, and a covariate related to the labour market - at the cut-off point, using a nonparametric local linear regression with the triangular kernel. The results of this robustness test show only three statistically significant discontinuities out of twenty-eight tests.
So, also the third robustness proof confirms that the L488 selection process has been carried out without significant irregularities.

8. Conclusions

The aim of this paper is to develop a novel approach to evaluate causal effects of L488 on two different outcome variables – yearly growth rate of the share of investment on turnover and yearly growth rate of turnover - in a RDD framework. We take advantage of the quasi-experimental procedure indicated by the RDD, in order to exploit the discontinuity points of the forcing variable, created by the auction mechanism of L488. After an accurate and complex process of merging, we use the sharp RDD for each region in each one of the auctions analysed, afterwards the estimates are aggregated and weighted through the innovative sharp MRDD (see Section 4).

The results show a positive and statistically significant effect of the L488 funds on the yearly growth rate of the share of investment on turnover and on the yearly growth rate of turnover. The investment increases from 6.5 to 7.7 percentage points higher in the subsidized firms than in non-subsidized ones, whereas the turnover increases from 7.5 to 10.5 percentage points higher in favour of the subsidized firms, over the period 1995-2001. These estimates are statistically significant and robust to different bandwidths and kernels. The parametric and the nonparametric analyses on the aggregated data set, confirm the extra investment caused by L488. The results show that L488 has achieved the targets selected by the policy makers. The subsidized firms have invested more than they usually would, and they have increased the production more than the non-subsidized firms. The positive impact of subsidies on firm growth is also consistent with most of the literature. The effect on turnover is higher to the estimates presented in Bernini and Pellegrini (forthcoming) (2-3% on annual base), using a DID matching, whereas the effect on capital accumulation are similar (around 11%). The effect on capital accumulation is larger than what is reported in Bronzini and Di Blasio (2006). They show an almost nil cumulative effect of subsidized investment on assets after five years, whereas in our study the average increase in assets is positive and significant, even if measured over a shorter period. Unfortunately, effects of L488 over a longer period cannot be estimated in our dataset, because of the financial-statement data set end in 2001. As a consequence it may not be feasible to fully disentangle the substitution effects; in addition, this limited time span has not allowed us to analyse the auctions started after 1999. Even the analysis shows a strong evidence of additional investment and product induced by the policy, the long–run effects on firms’ competitiveness cannot be assessed. Some empirical evidence suggests that this positive effect on firm output is mainly in the short term (Bernini and Pellegrini, forthcoming).

Concerning the validity of our results, the methodological approach we used and the L488 characteristics limited the external validity\(^ {26}\); on the other hand the subsidies allocation mechanism is a source of high internal validity for our results.

There are two aspects that are left out for future research: a methodological one, related to the use of a hierarchical model able to control for the correlation among different auctions; an

\(^ {26}\) The methodological implication in using a MRDD is that new firms’ performances cannot be evaluated. Moreover, our sample is restricted only to firms with a meaningful balance sheet (we use only corporate enterprises already active at least one year before the auction) that applied to an auction in the southern regions. Therefore our results can hardly be extended to very small firms and to policy interventions in other (more developed) areas, like the northern regions of Italy.
empirical one, in which are analysed the macro effects of L488 in the long run, in order to verify if L488 has been able to activate growth processes self-propulsive on the Italian lagging areas.

References


Approach”, w. paper N. 6, Dipartimento di Teoria Economica e Metodi Quantitativi per le Scelte Politiche, Università di Roma “La Sapienza”.


Imbens G., Kalyanaraman K. (2010). "Optimal bandwidth choice for the regression discontinuity estimator," CeMMAP w. paper CWP05/10, Centre for Microdata Methods and Practice, Institute for Fiscal Studies


