The role of active monitoring in reducing the execution times of public works. Evidence from Italy

Giuseppe Francesco Gori *, giuseppe.gori@irpet.it
Patrizia Lattarulo *, patrizia.lattarulo@irpet.it
Marco Mariani *, marco.mariani@irpet.it

* IRPET – Tuscany's Regional Institute for Economic Planning, Villa La Quiete alle Montalve - Via Pietro Dazzi, 1 · 50141 Florence, Italy

Abstract
The article investigates the role of active monitoring with respect to the execution times of public works. Exploiting the data from an Italian regional case study (Tuscany), where a local law imposes more active monitoring than usual for projects passing a minimum ‘financial’ size and benefiting from co-financing by the regional government above a certain threshold, the article contributes to the literature by estimating the causal effect of increased monitoring within a regression discontinuity approach and by means of survival analysis techniques. The presence of two variables jointly determining the assignment of projects to increased monitoring makes this estimation problem particularly challenging. Results suggest that, at the threshold values of the two assignment variables, the causal effect of active monitoring on time-to-completion is positive on very short projects that would have not lasted long anyway and, more interestingly, on the subset of projects that are very persistent in time.

1. Introduction
Public procurement accounts for a sizeable share of economic activity in most countries, represents an important device for economic development and an essential tool of economic policy (e.g. Pickernell et al., 2011; Loader, 2013). As such, it receives remarkable and increasing attention from scholars in different areas of economics, public administration, policy management and planning.

In the economic literature, the buyer's choice of the auction format and the resulting selection of contractors, as well as the degree of contracts' completeness are often deemed determinant in explaining the final cost and the time-to-completion of public works (Dimitri et al., 2006; Bajari et al., 2009; Lewis and Bajari, 2011; Bucciol et al., 2013; Decarolis, 2014). However, some recent contributions have started to shed light on some further systemic and institutional aspects which potentially affect procurement performances after the negotiation process has ended, i.e. during the execution phase. These aspects include court efficiency (Iossa and Spagnolo, 2011; D’Alpaos et al., 2013; Coviello et al., 2013), corruption and collusion (Flyvbjerg, 2007; Iossa and Martimort, 2011).

Despite these recent developments, many are the issues that still remain in the shade. One of these, i.e. the role of public monitoring on the works' execution phase, has been emphasized in the public administration, planning, policy and project management literatures (e.g. Brown and Potoski, 2003; Thay, 2008; Pryke and Smyth, 2012), while it has quite surprisingly been neglected by empirical economists, possibly due to lack of suitable data. The obvious idea is that public buyers should combat delays and cost escalations by performing tighter post-contract monitoring, accompanied by timely actions against bottlenecks (henceforth active monitoring).
Some contributions have also highlighted that public buyers at different ‘governance’ levels may behave very differently not only in terms of achieving advantageous contractual procurement costs but also in the subsequent management of the execution process: while central governments seem to be less able than peripheral ones to close advantageous contracts (Bandiera et al., 2009), peripheral governments and agencies seem to be less able than central bodies to fight execution delays (Guccio et al., 2014). This heterogeneity can be ascribed to a variety of reasons that are hardly ‘captured’ in empirical analyses, however, it represents an undesirable phenomenon in a public perspective, where the interest lies in raising general post-contract monitoring effort.

For example, in the Italian context that will be the focus of our analysis, the most prominent example of monitoring activity is the one carried on by the National Anti-Corruption Authority (ANAC). Although the Authority collects information about the tendering and the execution of every public procurement contract, it mainly focuses its monitoring activity on the formal legality of the tendering phase but, possibly due to the enormous amount of contracts in the country, it exerts much less active monitoring on executions (AVCP, 2012).

The opportunity to analyze the role of active monitoring on executions in greater depth is offered to us by an Italian case study, where it is highly encouraged by local regulation. In fact, Tuscany’s Regional Law 35/2011 (henceforth the Law) aims at supporting local buyers in the contracts’ enforcement for those public works that are considered as strategic from the regional government standpoint. As such, these projects must pass a minimum ‘financial’ size and benefit from co-financing by the regional government above a certain threshold.

In addition to imposing a more careful definition of the timing of the contracts’ phases and to requiring buyers to report frequent progress updates to governmental offices, the law confers ultimate substitutive powers to the regional offices in case of extremely inactive local buyers.

Substitutive powers have so far been exerted in a very limited number of cases. More than providing ad hoc treatments for pathological cases, this regulatory scheme has the purpose of giving a systematic boost to the practice of active monitoring by buyers, with particular regard to execution phases that would otherwise remain outside of the attention of the national supervision authority.

Our main interest lies in estimating the contribution of this boost, if any, in terms of reduction of execution times. The analysis is based on a rich administrative dataset of public works implemented in Tuscany from 2008 to 2014, where projects are subject to mandatory active monitoring provided that they meet the already mentioned size and co-financing conditions.

In order to identify a proper causal effect of active monitoring, we exploit the exogeneity of these conditions and adopt a sharp regression discontinuity (RD) approach (Lee and Lemieux, 2010), extending this identification strategy to the case of multiple assignment variables. To this regard, note that our situation is substantially different from the ones analysed in the very recent literature on multiple assignment variables in regression discontinuity designs (e.g. Papay et al., 2011; Wong et al., 2013), where assignment to treatment depends on meeting just one of multiple conditions. In our case, in fact, multiple conditions must be jointly met in order to assign a project to monitoring.

Moreover, instead of focusing, as most of previous contributions have done, on completed projects only, our approach comprises a higher number of public works whose execution can be either completed or still on-going at the time of the analysis. In order to handle this data structure and address the potential right-censoring (i.e. the incompleteness) of execution spells we estimate the discontinuity in execution times by means of survival analysis techniques (e.g. Caliendo et al., 2013).

We find that the causal effect of monitoring on time-to-completion is positive at the threshold values of the two assignment variables. In particular, active monitoring seems to be effective on very short projects that would have not lasted long anyway and, more interestingly, on the subset of projects that are very persistent in time.

The remaining of the paper proceeds as follows. Section 2 describes the database, while section 3 present the empirical strategy and the econometric model. Section 4 describes the results while Section 5 concludes.
2. Data

Our dataset integrates the SIMOG dataset from National Anti-Corruption Authority (ANAC), with the SITAT dataset from the Observatory on public contracts of the Tuscany Region, the latter dataset being more frequently updated for what concerns the execution phase of public works. The resulting integrated dataset includes 8,000 records, 690 of which are monitored according to the Law. However, each record of the dataset corresponds to a single lot, which can represent a project or a only a portion of a multi-lot project. Since the monitoring provided for by the Law targets projects, we have aggregated the lots, where necessary, in order to obtain a dataset where each record represents a single project instead of a lot.

Moreover the SIMOG-SITAT dataset presented a certain number of records with missing information about the sources of financing of the project or the starting date of its execution phase. Since the share of regional co-financing is a key variable for the assignment of a project to the monitoring, and the observed duration of the execution phase is determinant for the estimation of the effect of the monitoring, we choose to further reduce the number of observations considered for our analysis by discarding these records. The resulting dataset, includes 1,896 projects, where 74 works are subject to the monitoring and 1,822 are not. Below (Table 1) we provide a short description of the characteristics of the works belonging to the two groups.

| Table 1 – Characteristics of the monitored and not monitored projects in the dataset |
|-------------------------------------------------|----------------|----------------|
| Share of regional co-financing                  | 25%            | 80%            |
| Total funding (Mln Euros)                       | 0.9            | 2.7            |
| Average Duration (Months)                       | 15             | 30             |
| Percentage of censored projects                 | 59%            | 31%            |
| Average duration of concluded projects(Months)  | 21             | 25             |

3. The empirical strategy

According to the criteria established by the Law, the share of regional co-financing and the financial dimension of the projects are the two key variables that determine whether or not a project is subject to the active monitoring. In particular, projects are monitored if they meet two conditions at the same time: a financial dimension equal to or greater than the threshold value of 500 thousand euros and a share regional of co-financing higher than the 50%.

Considering that the process of assignment to monitoring is a deterministic function of two specific variables, the financial size of the project and the share of regional co-financing, we choose to adopt a sharp regression discontinuity design (RDD, see Lee and Lemieux, 2010).

The idea behind this approach is that in the immediate vicinity of the threshold values that determine the assignment to monitoring, the projects on the right and left of the threshold are sufficiently similar - both in terms of observable and unobservable characteristics – so as to estimate correctly a local causal effect of the monitoring.

This is true if the level of the thresholds is determined exogenously, i.e. regardless of the potential outcomes of the projects (Imbens and Wooldridge, 2008). In our case, exogeneity holds since the financial dimension of the project and the share of regional co-financing are set even before the work is put out to tender. Also note that, once identified, the executing company can in no way intervene on variables that have already resulted in the assignment of the work to the monitoring and, most likely, is not even aware of the fact that the project is receiving additional monitoring than usual.

Under these conditions, and thanks to few other mild assumptions (see Lee and Lemieux, 2010, for details) the execution times of the close-to-threshold projects that were not assigned to monitoring can approximate what the execution times of close-to-threshold monitored projects would have been, if the latter were not monitored.

3
In our dataset, durations are represented as a continuous variable (expressed in days) and it is potentially right-censored at the end of the observation period (31-12-2014). A possible approach is that of limiting the analysis to the uncensored observations. However, with data like ours, where censorship is not negligible, this would imply a drastic reduction of the sample size and potentially lead to a biased estimation of the causal effect of active monitoring. Accordingly, a proper way to proceed in the presence of right-censored data is to use survival analysis techniques. However, for estimation convenience, we discretise execution times into semesters, so as to handle issues of hazard non proportionality in a more straightforward way than what would be possible, for example, with the popular semi-parametric Cox model. Discretisation of right-censored spells in a single assignment variable regression discontinuity setting was recently undertaken by Caliendo et al. (2013). This requires the dataset to be re-structured so that the logit model typically used for estimation accounts for the “population at risk” in each time period, which makes it a discrete-time survival model (Klein et al., 2013).

The most novel aspect of our empirical strategy relates to how the discontinuity is estimated in the presence of two assignment variables that projects must fulfil in order to receive monitoring. To this regard, note that our situation is substantially different from the ones analysed in the very recent literature on multiple assignment variables in regression discontinuity designs (e.g. Papay et al., 2011; Wong et al., 2013), where assignment to treatment depends on meeting just one of multiple conditions. Following Choi and Lee (2014), the problem can be viewed as follows.

Figure 1 – Monitored and Non-monitored (control) groups

Monitored projects are in the region 1 in Figure 1, corresponding to the region above both threshold values (C1 and C2 in the figure, which stand for “cutoff”). The other regions host non-monitored projects that meet only one of the two conditions (2 and 4) or none of them (3). As already mentioned, we are interested in estimating the jump in the outcome variable involving only those observations that are in a close proximity of the intersection of the two thresholds, for example those in the are surrounded by the dotted red line in Figure 1. This region is identified by the intersection of two bandwidths: \( h_1 \) and \( h_2 \), where \( h_1 \) relates to the first running variable (the share of regional co-financing) and \( h_2 \) relates to the second running variable (the financial dimension).

According to the most recent trends in the regression discontinuity literature, to this purpose, it is preferable to discard observations too far away from the thresholds so that the estimation can be carried out by means of a simply specified local regression, instead of resorting to complex polynomial specifications to be run on all available observations (Imbens and Lemieux, 2009; Gelman and Imbens, 2014).

The extent to which observation are to be included in the analysis should be decided using a bandwidth selection procedure that was recently put forward in the literature (Imbens and Lemieux, 2008, Imbens and Kalyanaraman, 2012, Calonico et al., 2014). These procedures allow to establish a bandwidth \( h \) where the estimation can be carried out, a bandwidth that offers a good balance between bias-reduction and precision issues. However, all these procedures are conceived with respect to a single assignment variable, and need to be extended to our multiple assignment variable setting. The most intuitive way to perform this extension relies on the cross-validation procedure set out in Imbens and Lemieux (2008).
Typical cross-validation compares model fitting for all possible alternative bandwidth of a single assignment variable and chooses the bandwidth where this fitting is maximised. Our extension compares model fitting for each possible alternative combinations of the two forcing variables.

Based on these arguments, considering that the literature also favours flexible specifications, allowing coefficients to take different values above and below the thresholds (Imbens and Lemieux, 2008), and considering that our outcome variable is binary and observed in each time period, we specify the following logit model:

\[
Pr(Y_{it} = 1|\text{at the thresholds}) = \frac{\exp(\cdot)}{1 + \exp(\cdot)}
\]

where \(Y_{it} = 1\) if the project \(i\) is completed at time \(t\), and 0 otherwise, and the argument of the exponential function in (1) writes as follows:

\[
(\cdot) = \sum_{t=1}^{T} \beta_0 t + \beta_1 M_i^h + \sum_{t=1}^{T} \beta_2 t M_i^h(t) + \beta_3 T_1^h + \beta_5 T_2^h + \sum_{t=1}^{T} \beta_6 t T_2^h(t) + \sum_{t=1}^{T} \beta_7 d_{11i} + \beta_8 d_{12i} + \beta_9 d_{13i} + \beta_10 d_{14i} + \beta_11 d_{21i} + \beta_12 d_{22i} + \beta_13 d_{23i} + \beta_14 d_{24i} + \sum_{i=1}^{2} \sum_{k=1}^{4} \sum_{t=1}^{T} \beta_{15,j,k,t} d_{jk}(t)
\]

In particular, \(M\) is our treatment variable (\(M_i = 1\) indicates that the project is monitored, \(M_i = 0\) otherwise). \(T_1\) and \(T_2\) are binary variables which assume value 1 if the value of the corresponding scoring variable is above the threshold, and 0 otherwise. In particular, the scoring variable 1 (\(S_1\)) is the share of regional co-financing, while (\(S_2\)) is the financial size of the project. The coefficients associated to the dummies \(T_1\) and \(T_2\) can be viewed as the partial fixed effects due to meeting one of the two criteria but not the other one (Choi and Lee, 2014). In addition to these dummies, the assignment variables are also inserted in the model with their continuous values.

However, the insertion of these assignment variables in the model requires that they are centred on their threshold value. In so doing, we account for the “distance” of each observation from the two thresholds. This entails the terms \(d_{jk}\) where \(j = 1,2\) indicates the scoring variable’s threshold with respect to which the distance is calculated, while \(k = 1,2,3,4\) index the region of the co-financing/size space.

Note that \(\cdot\) also includes time dummies in order to account for duration dependence and that, in order to allow for possible non proportionalities depending on the fact that the project fulfils only one or two assignment criteria, these dummies are interacted with all the variables of the model.

The quantity of primary interest, corresponding to the discontinuity in execution times ascribable to the active monitoring is not expressed by a single coefficient, but need to be reconstructed, for each time period, using the coefficients \(\beta_1\) and \(\beta_2\). The model, as previously argued, has to be estimated limited to the projects included in the optimal bandwidth \(h\). Figure 2 shows the projects in the space defined by the two assignment variables.
Figure 2 – The projects under analysis with respect to the two assignment variables and the bandwidth

Note that, in the proximity of the thresholds, the density of the projects is quite different in the four regions, which, in our view, suggests allowing the bandwidth to be non-symmetrical around the thresholds. The cross-validation procedure that we have applied suggests the following optimal bandwidth configuration (Table 3; see also the dashed lines in Figure 2). After setting these bandwidths, estimation is carried out on 257 (56 monitored, 201 not monitored) out of 1.896 projects initially taken into consideration.

Table 3 – Non-symmetrical bandwidths obtained through leave-one-out cross validation extended to the case of multiple assignment variables

<table>
<thead>
<tr>
<th>Region</th>
<th>$S_1$(%)</th>
<th>$S_2$ (Euros)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>976.000</td>
</tr>
<tr>
<td>2</td>
<td>12.1</td>
<td>803.000</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>280.000</td>
</tr>
<tr>
<td>4</td>
<td>81.7</td>
<td>270.000</td>
</tr>
</tbody>
</table>

4 Results

The following table presents the results of the estimation. As previously recalled, the coefficient of interest are $\beta_1$ and $\beta_{2t}$ and, as a result of the model we specified that fully allows for non proportionality, the magnitude of the discontinuity varies with time.

Table 4 – Estimation results

|          | Odds Ratio | Std. Err. | z      | P>|z|  | [95% Conf. Interval] |
|----------|------------|-----------|--------|-----|----------------------|
| $\beta_{01}$ | 0.2615499  | 0.2051292 | -1.71  | 0.087** | 0.0562304 - 1.216572 |
| $\beta_{02}$ | 0.7809157  | 0.5524593 | -0.35  | 0.727 | 0.1951737 - 3.124548 |
| $\beta_{03}$ | 2.275928    | 2.280553  | 0.82   | 0.412 | 0.3193208 - 16.22146 |
| $\beta_{1}$  | 9.84E+13    | 3.87E+14  | 8.2    | 0*** | 4.44E+10 - 2.18E+17  |
| $\beta_{2t}$: base $\beta_2$ | | | | | |
| $\beta_{21}$ | 4.70E-13    | 2.64E-12  | -5.06  | 0*** | 7.92E-18 - 2.79E-08  |
| $\beta_{22}$ | 9.36E-19    | 6.01E-18  | -6.46  | 0*** | 3.19E-24 - 2.75E-13  |
| $\beta_3$   | 5.40E-38    | 1.56E-37  | -29.66 | 0*** | 1.86E-40 - 1.57E-35  |
\[ \begin{align*}
\beta_{4}: \text{base } \beta_{4} \\
\beta_{41} & \quad 1.51E+36 & 6.32E+36 & 19.89 & 0^{***} & 4.11E+32 & 5.54E+39 \\
\beta_{42} & \quad 2.19E+40 & 1.03E+41 & 19.71 & 0^{***} & 2.13E+36 & 2.26E+44 \\
\beta_{5} & \quad 7.06E+22 & 2.29E+23 & 16.2 & 0^{***} & 1.21E+20 & 4.10E+25 \\
\beta_{63}: \text{base } \beta_{63} \\
\beta_{61} & \quad 7.99E-24 & 3.16E-23 & -13.43 & 0^{***} & 3.41E-27 & 1.87E-20 \\
\beta_{62} & \quad 1.44E-22 & 6.85E-22 & -10.58 & 0^{***} & 1.29E-26 & 1.61E-18 \\
\beta_{7} & \quad 0.4695351 & 0.5425952 & -0.65 & 0.513 & 0.0487549 & 4.521864 \\
\beta_{8} & \quad 4.87E+90 & 4.33E+91 & 23.49 & 0^{***} & 1.32E+83 & 1.80E+98 \\
\beta_{9} & \quad 55.36857 & 126.3509 & 1.76 & 0.079^* & 0.6321561 & 4849.56 \\
\beta_{10} & \quad 4.06E+62 & 1.97E+63 & 29.75 & 0^{***} & 3.05E+58 & 5.41E+66 \\
\beta_{15,1,2r}: \text{base } \beta_{15,1,2r} \\
\beta_{15,1,2,1} & \quad 5.40E-91 & 5.04E-90 & -22.26 & 0^{***} & 6.09E-99 & 4.79E-83 \\
\beta_{15,1,2,2} & \quad 3.67E-80 & 7.23E-79 & -9.28 & 0^{***} & 6.07E-97 & 2.22E-63 \\
\beta_{15,1,3r}: \text{base } \beta_{15,1,3r} \\
\beta_{15,1,3,1} & \quad 0.0064147 & 0.0177452 & -1.83 & 0.068^* & 0.0000283 & 1.451679 \\
\beta_{15,1,3,2} & \quad 0.0244515 & 0.0652682 & -1.39 & 0.164 & 0.0001307 & 4.575209 \\
\beta_{15,1,4r}: \text{base } \beta_{15,1,4r} \\
\beta_{15,1,4,1} & \quad 4.67E-65 & 6.68E-64 & -10.34 & 0^{***} & 2.99E-77 & 7.29E-53 \\
\beta_{15,1,4,2} & \quad 1.54E-69 & 1.42E-68 & -17.18 & 0^{***} & 2.17E-77 & 1.09E-61 \\
\beta_{15,2,1r}: \text{base } \beta_{15,2,1r} \\
\beta_{15,2,1,1} & \quad 0.9999925 & 4.63E-06 & -1.61 & 0.107 & 0.9999835 & 1.000002 \\
\beta_{15,2,1,2} & \quad 0.9999964 & 3.82E-06 & -0.93 & 0.352 & 0.9999889 & 1.000004 \\
\beta_{15,2,2r}: \text{base } \beta_{15,2,2r} \\
\beta_{15,2,2,1} & \quad 1.000214 & 0.0000168 & 12.75 & 0 & 1.000181 & 1.000247 \\
\beta_{15,2,2,2} & \quad 1.000218 & 0.0000211 & 10.35 & 0 & 1.000177 & 1.00026 \\
\beta_{15,2,4r}: \text{base } \beta_{15,2,4r} \\
\beta_{15,2,4,1} & \quad 1.000266 & 0.000013 & 20.5 & 0^{***} & 1.000241 & 1.000292 \\
\beta_{15,2,4,2} & \quad 1.000317 & 0.0000235 & 13.5 & 0^{***} & 1.000271 & 1.000363 \\
\beta_{11} & \quad 0.9999978 & 2.59E-06 & -0.86 & 0.391 & 0.9999927 & 1.000003 \\
\beta_{12} & \quad 0.9997828 & 0.0000111 & -19.56 & 0^{***} & 0.9997611 & 0.9998046 \\
\beta_{13} & \quad 1.000002 & 1.96E-06 & 1.05 & 0.295 & 0.999982 & 1.000006 \\
\beta_{14} & \quad 0.9997212 & 0.0000114 & -24.5 & 0^{***} & 0.9996988 & 0.9997435 \\
\end{align*}\]

(*) Significant at 10%; (**) Significant at 5%; (***) Significance 1%. Standard errors are cluster-robust.
Note in Table 4 that the coefficient of interest are statistically significant. Based on these results we can estimate the probability of completion for monitored and non-monitored projects, at the thresholds, for each time period $t$. This probability, which is estimated on the population still at risk in $t$, is the main result of our survival analysis.

Table 5 – Predicted completion probabilities and discontinuity at the intersection of thresholds

<table>
<thead>
<tr>
<th>Period</th>
<th>if monitored</th>
<th>if not monitored</th>
<th>discontinuity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.92</td>
<td>0.20</td>
<td>0.72</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0.43</td>
<td>-0.43</td>
</tr>
<tr>
<td>3</td>
<td>0.98</td>
<td>0.69</td>
<td>0.29</td>
</tr>
</tbody>
</table>

These preliminary results suggest that the effect of monitoring can be appreciated in the very short and in the long run, much less in the medium term. In our view, these results could be explained as follows: active monitoring boosts the execution speed of “short” projects. As a result, the population at risk is soon constituted by a mixture of short, but unmonitored, projects and of long and persistent projects of both kinds. In the second semester, short but unmonitored project tend to come to an end, which explains the negative sign of the discontinuity. In a longer time horizon, we have only more persistent project waiting for completion. Here, monitoring turns to work again in reducing execution times, although it has to be noted that the magnitude of the positive discontinuity is much smaller than previously.

6 Concluding remarks

The article has investigated the role of active monitoring in reducing the execution times of public works. Exploiting the data from an Italian regional case study (Tuscany), where a local law imposes more active monitoring than usual for projects passing a minimum ‘financial’ size and benefiting from co-financing by the regional government above a certain threshold, we have checked whether increased monitoring works in reducing time-to-completion. Our estimation strategy relies on a sharp regression discontinuity design and is suitable to identify a causal relationship. Our estimation problem is made particularly challenging by the facts that i) the outcome of interest is potentially right censored, which calls for the adoption of survival analysis techniques that have so far been unusual with regression discontinuity designs and ii) the assignment of project to increased, active monitoring is jointly determined by two exogenous assignment variables, which requires the design of a very novel adaptation of regression discontinuity techniques to this particular setting. The results suggest that, at the threshold values of the two assignment variables, the causal effect of active monitoring on time-to-completion is positive on very short projects that would have not lasted long anyway and, more interestingly, on the subset of projects that are very persistent in time. To the best of our knowledge, this kind of result is completely new in the public procurement literature and suggests that there is room for public buyers to increase effort and improve their action during the execution stage of projects.

Bibliography


Choi, J., M., Lee (2014) Regression Discontinuity with Multiple Running Variables for One Treatment. Korea University, mimeo.


