Reputation and Quality in Procurement

Field Evidence from of a Vendor Rating System

Riccardo Pacini and Giancarlo Spagnolo

****** VERY PRELIMINARY *******

May 28, 2012

Abstract

Reputational incentives are a powerful mechanism to improve suppliers performance, so strong to possibly start to influence suppliers behavior even before they are put in place. This paper presents field evidence on the effects of announcing the introduction of a reputational mechanism based on past performance when awarding procurement contracts of a large publicly regulated firm. Suppliers appear to react by improving (very) significantly their performance in dimensions to be included in their past performance rating. Awarding prices, on the other hand, did not to follow the strong increase in quality. EU public procurement directives forbid the use of past performance information at the awarding stage. Our results suggest that the loss of procurement quality and of taxpayer's value for money induced by this prohibition may be much larger that previously thought.

Keywords: public procurement, awarding criteria, past performance, reputation, vendor rating

1. Introduction

This paper presents field evidence on the role of reputational mechanisms in procurement. For a number of different reasons, from poor/costly contract enforcement to the complexity of many goods and services, court-enforced contracts are often not sufficient to achieve an effective governance of public procurement exchange.

Since procurement exchanges are rarely occasional, reputational forces may be exploited to improve on what formal contracting allows to achieve. In private procurement reputational considerations are indeed very important, whether they are informal and subjective or formalized in a feedback mechanism/Vendor Rating system (e.g. Bannerjee and Duflo, 2000).

There are several reasons why complementing explicit contracts with reputational mechanisms based on ex-post evaluations of contractor performance may improve the governance of procurement transactions. These are linked to both the inability of explicit contracts to describe (or the courts system to verify) important aspects of the procurement transactions at reasonable cost, but also to the high costs of enforcing explicit contracts through litigation. Many important quality aspects of supplied goods and services, particularly of more complex and valuable ones, are either very difficult to appropriately specify in an explicit contract in a practical and cost effective way or they are impossible to observe or to properly evaluate ex-post for a third party that could enforce the contract, like a court or an arbitrator. Even when a certain qualitative dimension or supplier choice could be specified contractually and verified by the court, the cost of enforcing the contractual remedies through litigation and the effect that this may have on the continuing of buyer-supplier relationship may often prevent an effective contractual governance.[1] (Macauley offer a very nice discussion of the latter costs).

Complementing contractual governance with reputation may be more difficult for public procurement, at least under some current regulations. Public procurement - besides sharing the governance problems of private procurement - it also has to solve the major problem of public governance: how to keep public buyers accountable in the absence of market pressures and with the many layers of agency shielding them from tax-payers' control. The interaction between this regulation and the governance of quality in procurement transactions is all but trivial. In private procurement past performance indicators affect the selection of suppliers and their behaviour because buyers can act upon past performance, refraining from selecting suppliers with a poor track record and favouring those with a good one. In public procurement buyers' this type of 'discretion' is typically limited. The need to prevent favouritism and corruption led lawmakers around the world to ensure that open and transparent auctions where bidders have equal treatment (even when they have a very different track record) are used as often as possible. Open competition is not only seen as an instrument to achieve efficiency and value for taxpayer money, but also to keep public buyers accountable by limiting their discretion in the allocation of public funds.[2]

---

[1] Macaulay (1963) classic study discusses extensively the latter problem and report a purchasing manager saying: "One doesn't run to lawyers if he wants to stay in business because one must behave decently" (p. 61). On the often very high costs of contract enforcement see the discussion in Iossa and Spagnolo (2011) and references therein.

[2] Another way by which lawmakers limit civil servants' discretion is constraining 'discretionary' payments, i.e. monetary transfers not based on observable but non-contractible tasks. Public buyers then tend to recover their

This paper suggests that reputational incentives may be very strong, able to greatly influence suppliers' behaviour already after a first generic announcement that past performance measures will be collected and used in the future for selection purposes.

2. The context of the experiment

The experiment relates to the introduction of a vendor rating system and the announcement on the use of the ratings to award new contracts by one of the largest public multi-utility companies listed on the Italian exchange ("the Firm"). The Firm operates in the sale and distribution of energy, water services and public lighting. In 2010 the Firm had a turnover of 3.6 billion of euro and produced 15.651 GWh of electricity, placing it as the sixth largest operator in Italy. In order to maintain an orderly functioning of its power grid, each year the Firm outsources works worth over 300 million euro. Since the Firm is controlled by a public administration, it has to apply the Italian Code of Public Contracts when selecting contractors and awarding contracts.[3] Being a multi-utilities company, the Firm falls in the "special sectors" which enjoy some flexibility in applying the Code. The Firm then employs a system of suppliers qualification to pre-select vendors. Starting from the second semester of 2007, it introduced a system of vendor rating for its qualified suppliers, with the plan to use its ratings at the awarding stage of the procurement process. The idea was to include vendors past performance regarding quality and security of works performed within the awarding criteria. The plan to introduce such a mechanism was announced to contractors, gradually disclosing details on its functioning and timing, along five main announcement events. This gives the possibility to study the reaction of vendors to the system introduction announcement. Before analyzing and empirically test it, we describe the legal framework and constraints in the use of past performance information in Italian public procurement tender and discuss the design of the vendor rating system.

2.1 Legal limits to consider reputation when awarding a contract

The possibility to introduce reputational elements for the selection of contractors and the award of contracts in public procurement has received wide attention in both academic and jurisprudential studies. This is of particular significance in Europe, where "contracting authorities shall treat economic operators equally and non-discriminatorily and shall act in a transparent way"[4] and competition is of primary importance.[5] Indeed, the use of reputational indicators presents a trade-off between the need to reduce the adverse selection and the moral hazard problems and the enforcement of the competition principle.

Concerning the adverse selection, the public procurement legislation for special sectors (according to Directive 17/2004/CE) is less stringent, since it allows public buyers to institute their own

---

discretion – for the good or for the bad - at the contract management/enforcement stage; see Iossa and Spagnolo (2011) for an analysis of discretional contract enforcement.

[3] The Code is the law that has implemented the European Union public procurement directives 17/2004 and 18/2004.

[4] Art. 2 of DIRECTIVE 2004/18/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 31 March 2004 on the coordination of procedures for the award of public works contracts, public supply contracts and public service contracts.

[5] "Contracts should be awarded on the basis of objective criteria which ensure compliance with the principles of transparency, non-discrimination and equal treatment and which guarantee that tenders are assessed in conditions of effective competition" Recital n.46 of the Directive 18/2004. "Non-discriminatory criteria should be indicated which the contracting authorities may use when selecting Competitors and the means which economic operators may use to prove they have satisfied those criteria." Recital n. 39 dir 2004/18/EC.

qualification system or, in general, to select potential candidates to be awarded, on the basis of their technical and professional skills, discretionally chosen by the contracting authorities. The unique limit in the choice of such criteria is the objectivity: "…Contracting entities which select candidates for restricted or negotiated procedures shall do so according to objective rules and criteria which they have established and which are available to interested economic operators…".[6] To this regard the use of reputation indicators may be fully exploited if built on a system based on measurable parameters, that is verifiable by third parties and agreed by the qualified contractors.

The problem arises in the awarding phase. Since the EU gives special prominence to the free and fair competition principle, the use of reputation as an award criteria in public procurement can constitute an unfair advantage for the incumbents and a not proportionate disadvantage for new entrants: a potential supplier with no past experience cannot enjoy any reputational premium with respect to preexistent competitors. This may reduce entry and competition and violate the general principle of equal treatment. In the phase of awarding contracts, the most economically advantageous tender[7] is the criteria which allow to consider other criteria than the price. To this regard, the European Court of Justice clearly stated that the awarding authorities, when evaluating quality with the most economically advantageous tender should consider the object of the tender and not the bidder's characteristics.[8] Also the Italian Public Procurement Authority reaffirmed the same principle.[9] However, the Firm was experiencing poor performances from its qualified contractors and the inefficacy of penalties to enforce contract provisions on quality.[10] This drove the Firm to introduce reputational criteria when awarding contracts, through exploiting the higher flexibility given by the Code to the contracting authorities belonging to "the special sectors" and awarding contracts under the EU thresholds.

2.2 Designing the vendor rating system

The Firm designed its vendor rating system for the procurement of works in the electricity sector. The system considers a set of 136 parameters linked to the stringent quality and safety regulation of this industry, according to which contractors performances are evaluated and an overall reputation index is calculated. These parameters were collected by a team of (rotating) auditors in a number on site visits. The score given to each parameters is equal to 1 if the value is "regular", to 0 if the value is "irregular" or "n/a" if not possible to be inspected. The set of 136 parameters is divided into two macro-types, Safety (51) and Quality (83), further sub-grouped according to 12 Safety and Quality dimensions (7 for Safety and 5 for Quality) as follow:

---

[6] Art. 54 comma 2, Dir.17/2004/CE.

[7] For awarding criteria specification, see art. 53 Dir.2004/18/EC and art. 55 Dir. 2004/17/EC.

[8] See judgments in Causes C-488/01 or C-31/87.

[9] Resolution n. 30 of 06/02/2007.

[10] Some data and experiences show that penalties are not effective because they are not even applied: a study conducted for Consip, the Italian public procurement agency, on a sample procurement contracts on goods and services, demonstrated that penalties were applied just in the 3.7% of the eligible cases (Albano, Dini, Spagnolo [2008]).

Table 2

Type	Dimension	Number of parameters
Safety	Equipment and machinery	5
	Documentation	9
	Works execution	8
	Personnel	4
	Works site regularity	10
	Works site safety	10
	Works site controls	5
Quality	Works on joints	19
	Customer relationship mgnt	3
	Air works	25
	Underground works	25
	Works on transformer station	13

Before computing a unique reputation aggregated index, each parameter was associated with a relative weight, ranging from 2 to 10. The reputation index (RI) is then calculated as a flat weighted average mean across a predefined time span, according to the following formula:

$$RI = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} v_{ij} p_j}{\sum_{j=1}^{n} p_j} \qquad (1)$$

with v_{ij} indicating the score obtained in each of the n parameter over all the m audits considered, p_j the weight attached. Hence, the reputation index can range from 0 to 1 and be calculated for the overall experiment period, for specific periods, group of contractors, single contractors, Safety and Quality dimensions or mixed criteria.

As we said, the award criteria most suitable to include elements other than price is the most economically advantageous tender, for which the general scoring rule (S) is:

$$S = \sum_{c=1}^{s} w_c f_c \qquad (2)$$

with f_c being the criteria to score the specific element of the bid, usually giving a score between 0 and 1, and w_c the weight to that element as defined in the tender document. Since usually $\sum_{c=1}^{s} w_c = 100$, S can range from 0 to 100. In this experiment, the scoring rule for each bidding contractor was announced to be:

$$S = w_d D + w_r RI \qquad (3)$$

where w_d was the weight attributed to the discount offered, D the discount offered, w_r the weight attributed to the reputation index and RI the reputation index, with $w_d + w_r = 100$, to substitute the current lowest price award criteria. The bidding contractor obtaining the highest S would be the winner of the specific tender.

2.3   Timing of implementation

The Firm defined the parameters in April 2007 and started conducting the audits in October 2007. On the 20th December 2007, the Firm announced to all qualified contractors the rationale behind the audits, namely the introduction of the vendor rating system and the plan to use it in the awarding of contracts in the future. Then there were other four events, on the 4th April 2008, the 10th July 2008, the 21st October 2008 and the 16th January 2009, at which the Firm gave updates on the functioning of the vendor rating system and the results of the audits in terms of RI for each contractors and impact on awarding contracts if the modified awarding rule were in place.
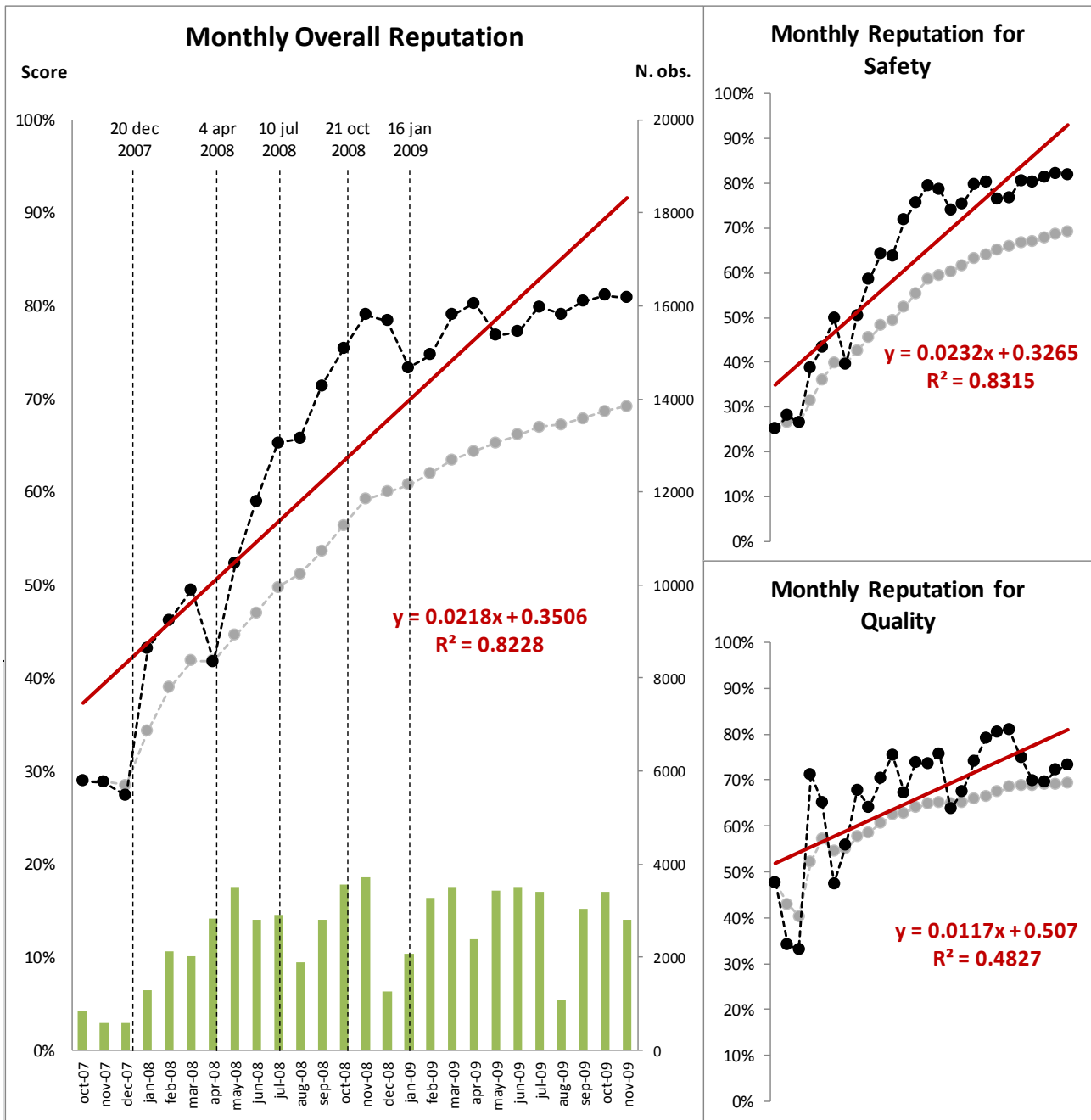
3. Empirical Analysis

The analysis concerns all inspections carried since October 2007 to November 2009 and the reactions by vendors in their performance to the various announcements given since the introduction of the vendor rating.

The Firm gave us access to all the results of inspections in the period between the 16th October 2007 and the 19th November 2009 across 45 different contractors, 222 contracts and 1,952 works sites: the inspections were carried out over the above mentioned 136 parameters that were checked for a total of 64,537 times throughout the sample period. This has generated a time series of 64,537 observations (i.e. inspected parameters). Moreover, we had access to data concerning 120 auctions run to award the contracts, whilst for the remaining 102 contracts the corresponding auction data were missing.

3.1   Descriptive statistics

The left hand side of Figure 1 shows the monthly distribution of the 64,537 observations throughout the sample period (see the green bars) and the progress of the reputation index (RI) calculated both on a monthly and a cumulative basis (the black and gray line respectively) on all the observations with respect to the announcements. The red line shows a significant positive trend in the monthly RI. The right hand side of Figure 1 shows the evolution of RI calculated on the observations relating the Safety and Quality dimensions separately. Figure 2 shows the same monthly distribution of the 64,537 observations in Figure 1 distinguishing the parameters regular (in green) from the irregular ones (in red), and the number of audits/inspections carried out each month. Figure 3 shows the distributions of parameters inspected grouped per the 12 dimensions. The parameters related to the 7 Safety dimensions are the ones most inspected (55,050 times), while those ones related to the 5 Quality dimensions were inspected 9,487 times. This may have a relation with the stronger response from contractors to improve the performance on Safety as shown in Figure 1. Figure 4 shows the progress of the reputation index (RI) calculated on a monthly basis for each of the seven most audited Safety and Quality dimensions.

Figure 1

Monthly Overall Reputation

Score

N. obs.

100% 20000

20 dec 2007   4 apr 2008   10 jul 2008   21 oct 2008   16 jan 2009

y = 0.0218x + 0.3506
R² = 0.8228

Monthly Reputation for Safety

y = 0.0232x + 0.3265
R² = 0.8315

Monthly Reputation for Quality

y = 0.0117x + 0.507
R² = 0.4827

Legend: The black line shows the average score calculated on all parameters inspected in the month of reference. The grey line shows the cumulated average score calculated on all parameters inspected until month of reference. The red line is the trend calculated out of the black line. The green bars are the total number of parameters checked throughout the month of reference. The vertical dashed line identifies each announcement date.
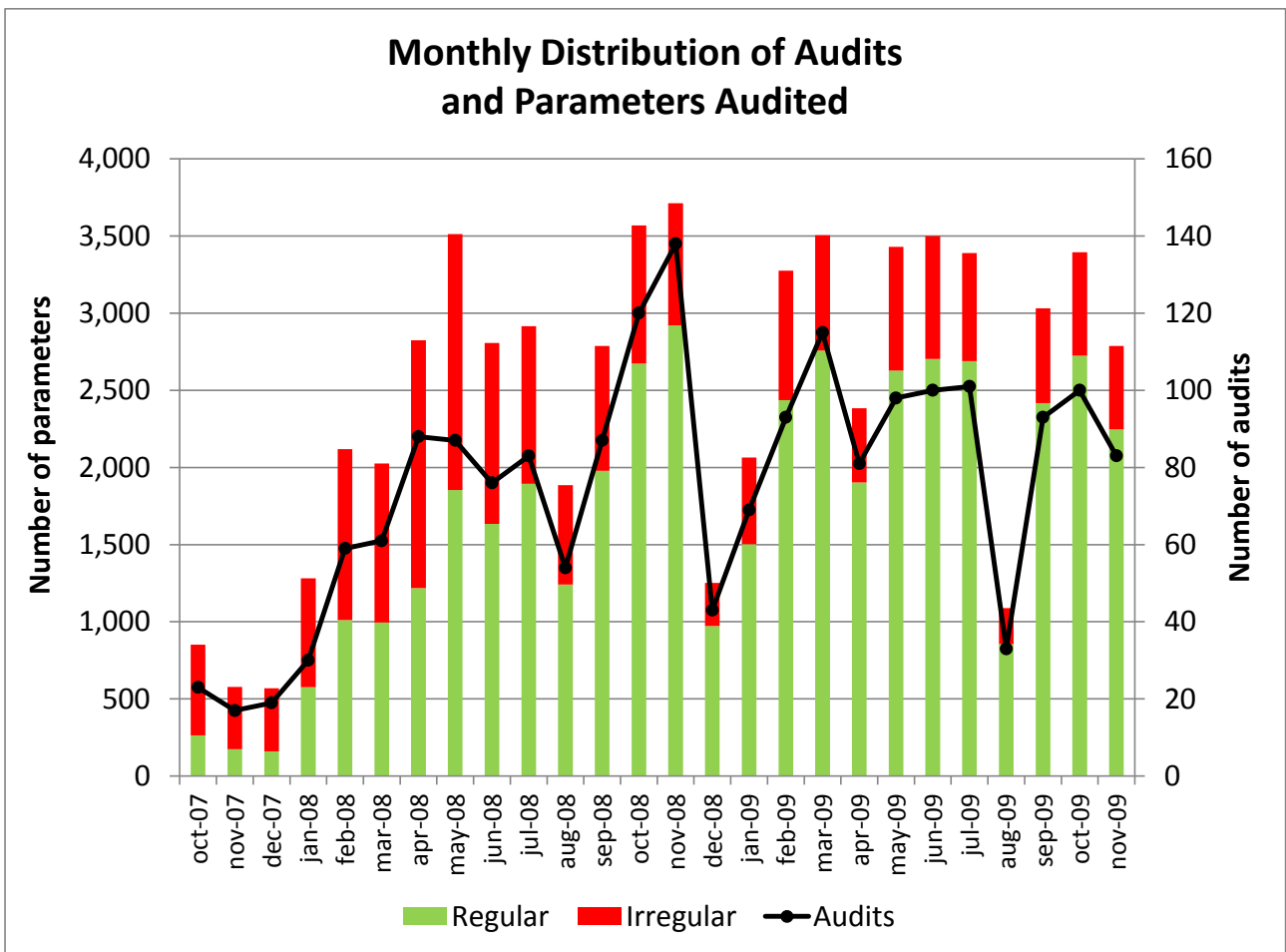
Figure 2

Monthly Distribution of Audits and Parameters Audited

Figure 3

Number of parameters audited per dimension

Figure 4

Monthly Reputation over the most audited dimensions

Legend:
- S - Equipment and machinery
- S - Documentation
- S - Works execution
- S - Personnel
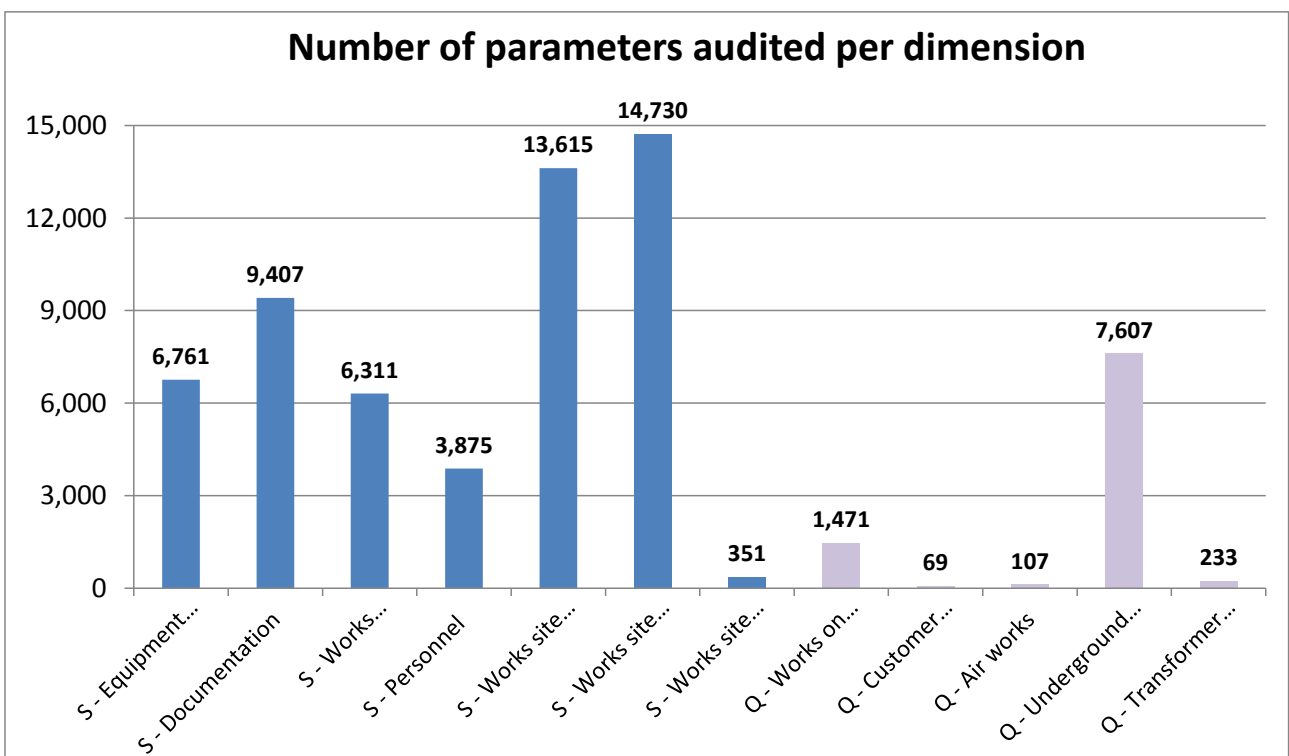- S - Works site regularity
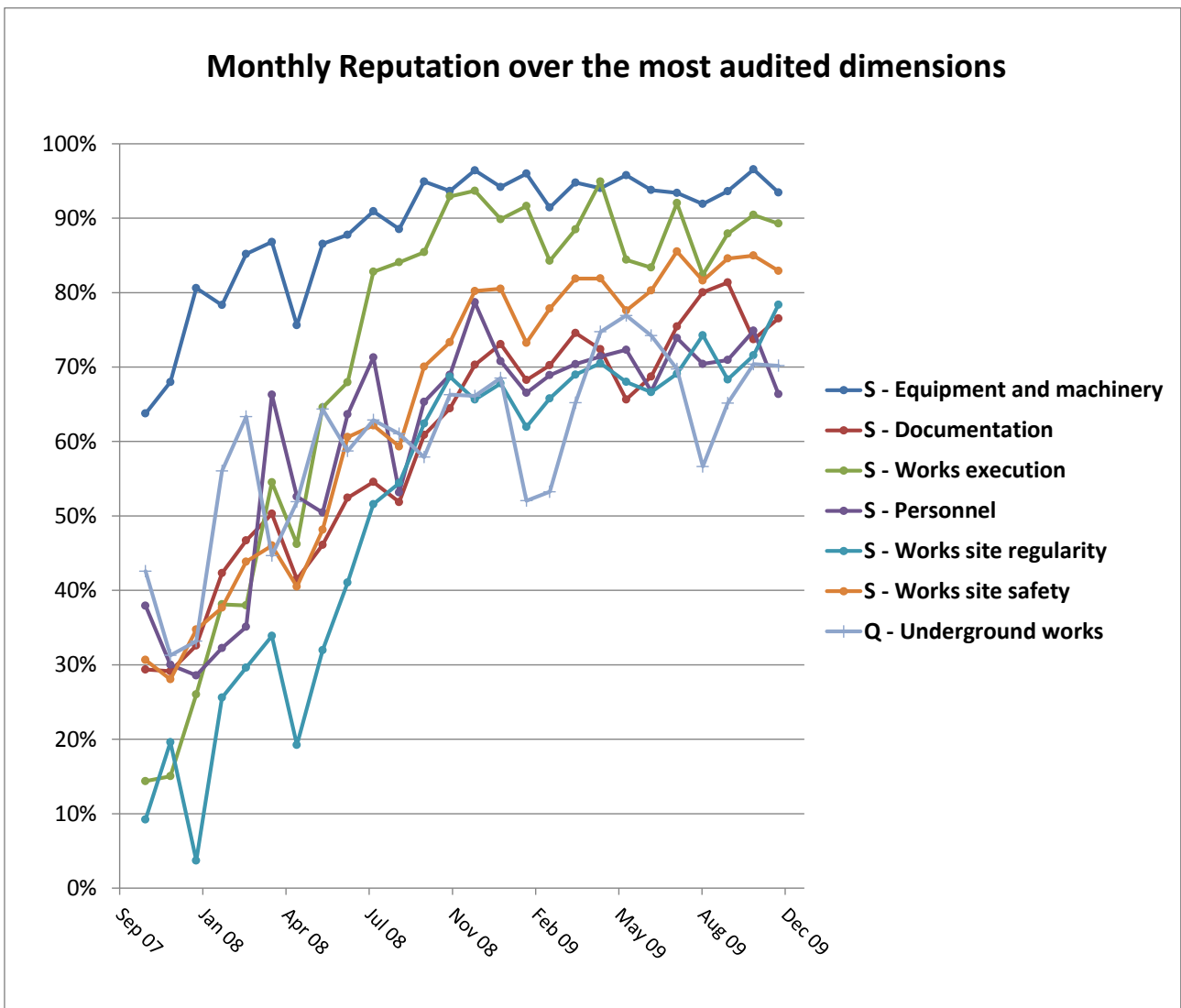- S - Works site safety
- Q - Underground works

Figure 5 shows the distributions of parameters inspected per the 45 contractors. The most inspected contractor was G that had 6,510 parameters inspected over 183 audits. Seven contractors received more than 100 audits and 19 contractors had more than 1,000 parameters inspected. Twelve contractors received less than 10 audits and 11 contractors had less than 200 parameters inspected. Maintaining the same order, Figure 6 gives the corresponding number of contracts awarded by each of the 45 contractors: 9 contractors were awarded 10 or more contracts (with AL as the most awarded contractor with 23 contracts), 10 contractors between 4 and 8 contracts, 15 contractors between 2 or 3 contracts and the remaining 11 contractors only 1 contract. Figure 7 plots the progress of the reputation index (RI), calculated on a monthly basis, for the 4 different grouping according to the number of contracts awarded, showing a common positive trend across the four groups. Figure 8 shows the monthly trend of the reputation index (RI) for each of the 9 contractors belonging to the first group of the most awarded contractors.

Figure 5
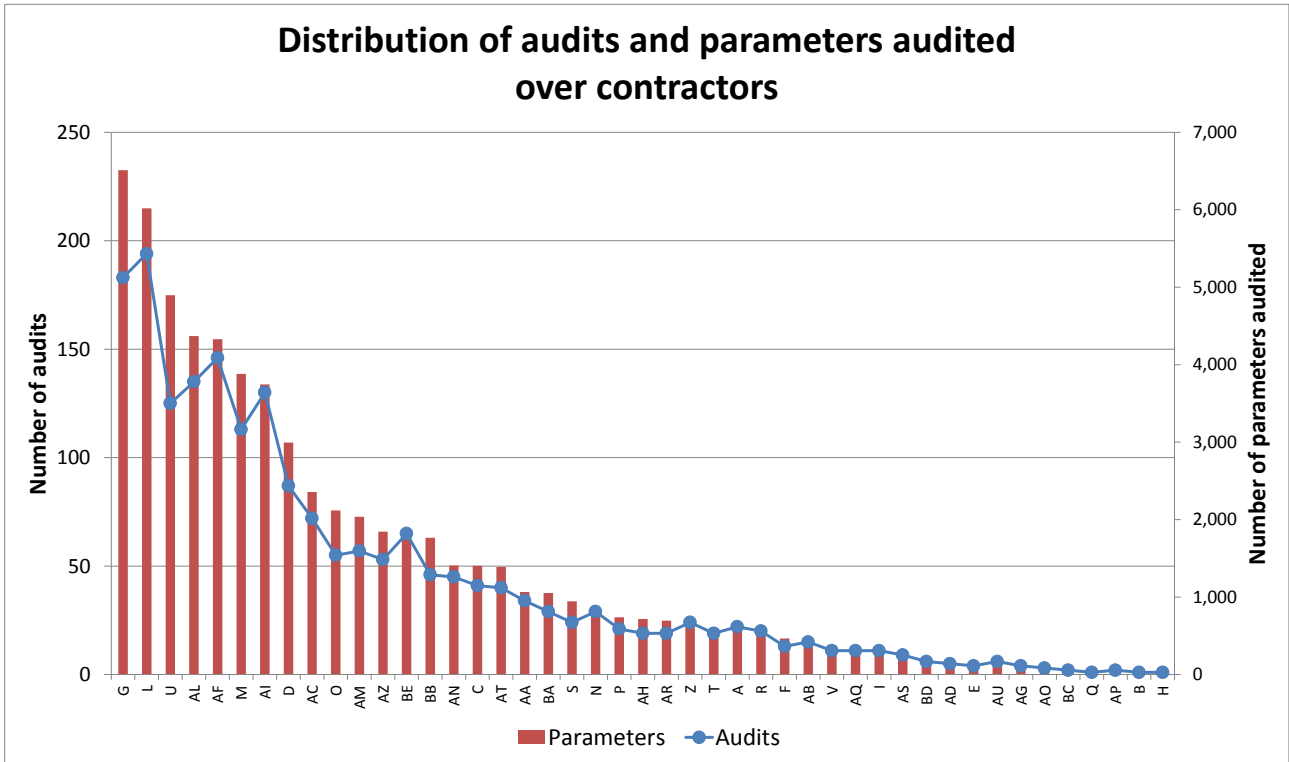

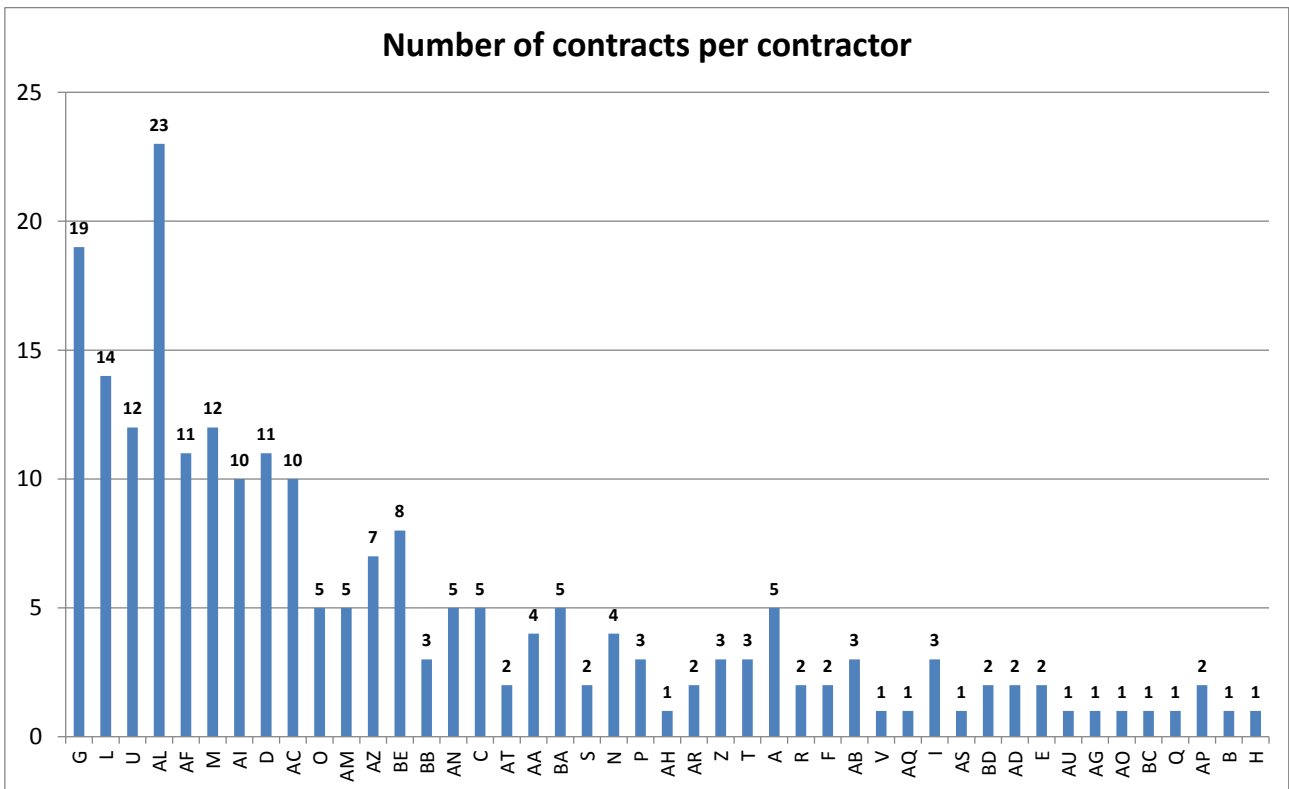Distribution of audits and parameters audited over contractors

Figure 6

Number of contracts per contractor

Figure 7

Firms Monthly Reputation by Group

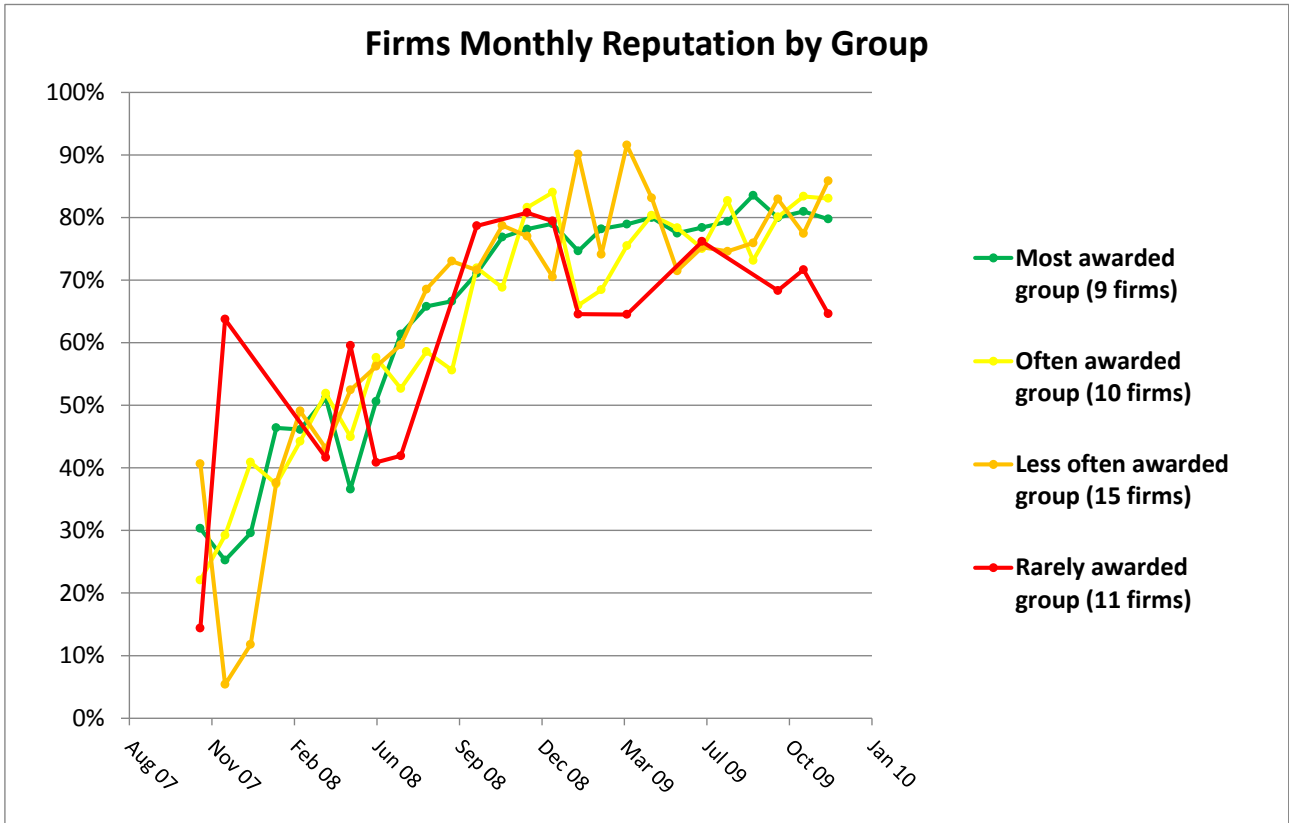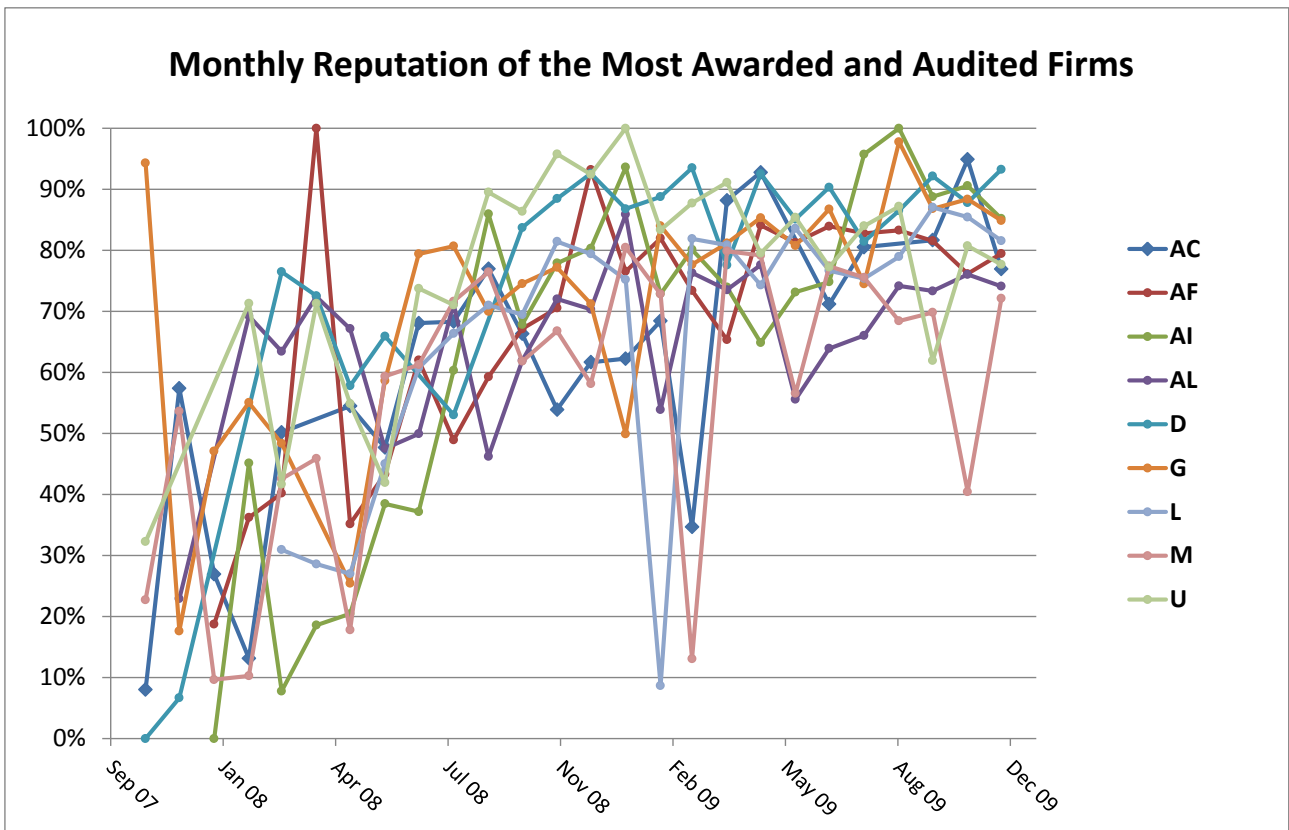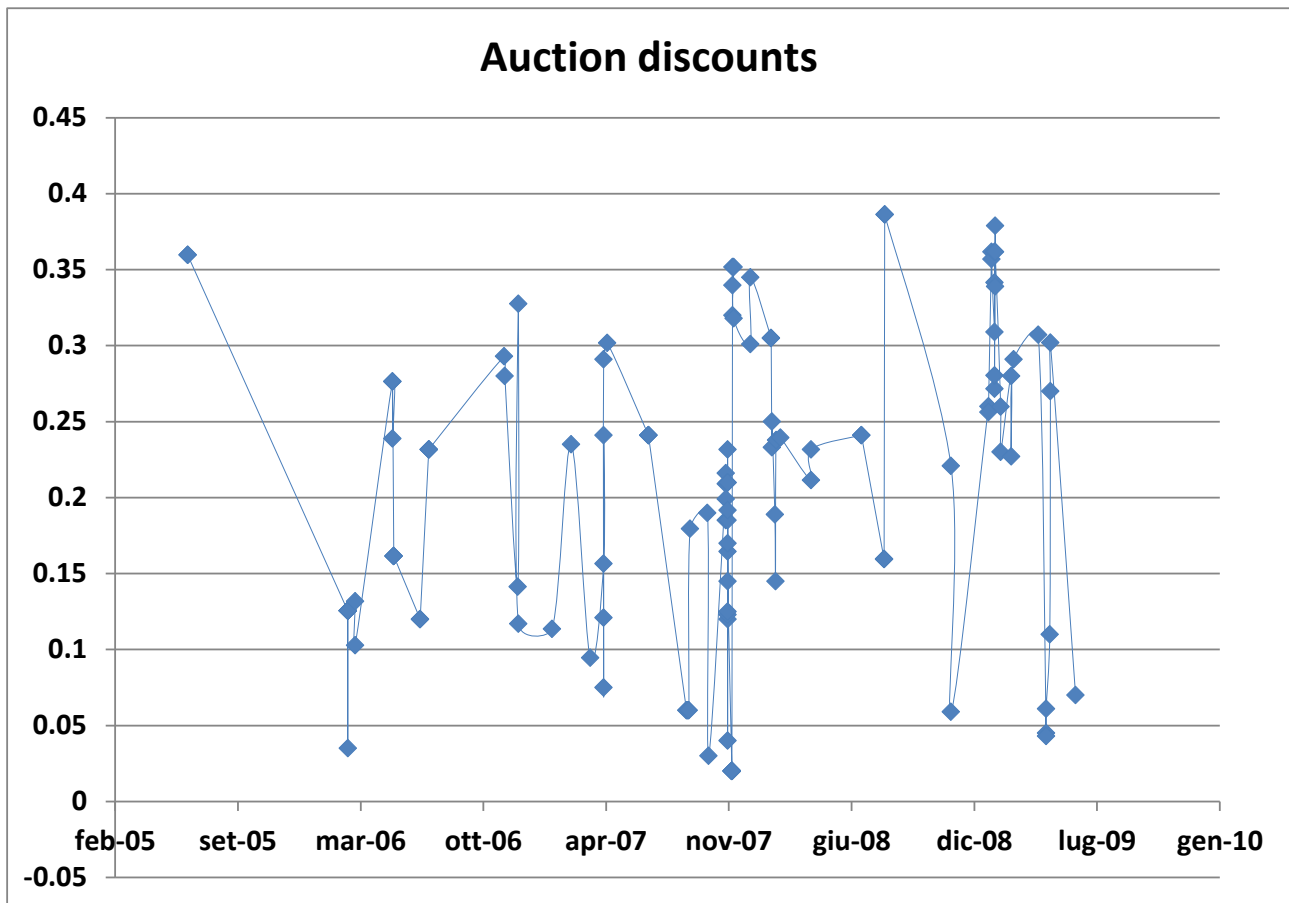**Figure 8**

Monthly Reputation of the Most Awarded and Audited Firms

Finally, Figure 9 shows the auction discounts for the 120 contracts out of the 222 contracts considered in the sample period for the calculation of the reputation index (RI).

Figure 9

3.2 *Empirics*

We carried out three simple statistical tests: i) a series of t-test on the five announcements relating the upcoming introduction of the vendor rating at the awarding stage on the reputation score and auction discount time series; ii) a probit regression on the single parameters scores; and iii) the correlation between reputation score and auction discounts. We used these series to test whether the progressive announcements to vendors that their past performances will be considered in future awarding of the contracts, caused an effect statistically resulting into a structural break.

Table 3 shows the t-test results for the five announcements carried out over 25 time series relating to the reputation for safety, for quality, for the 12 specific dimensions, for the 9 most awarded and audited contractors and for the auction discounts, for a total of 120 t-test (in other 5 cases it was not possible to run the test for lack of observations). Each test is run between the group of all parameters inspected before the specific announcement and the group of all parameters inspected after the specific announcement. In 106 cases the reputation before the announcement resulted significantly lower than after, while only in 3 cases significantly higher. In the remaining 11 cases the reputation was not significantly different. Counter-intuitively, the auction discounts too were positively affected by the announcements relating to the introduction of the vendor rating.

Table 4 reports the probit regression results for the following equation:

$$parameter = \alpha*weight + \beta*announc_{(1to5)} + \gamma*dim_{(1to11)} + \delta*group_{(1to4)} + \varepsilon \qquad (4)$$

where *parameter* is the binary score, 0 or 1, taken by the single parameter audited, *weight* is the weight associated with the parameter, *announc*$_{(1to5)}$ are five dummies which take 0 or 1 if the parameter is audited before or after the specific announcement of reference, *dim*$_{(1to12)}$ are twelve dummies which take 0 or 1 depending on the parameter belonging to one of the twelve specific safety or quality dimension, and *group*$_{(1to3)}$ are three dummies which take 0 or 1 depending on the parameter referring to one of the four grouping of firms according to the number of contracts awarded (see Figure 7). In particular, the regression confirms the positive and significant effect of all the five announcements, along with the significant presence of differences between the safety and quality dimensions and between the groups of firms.

Finally, we studied whether it existed a correlation between the discount offered in the tender procedure and the performance (measured by RI, the reputation index) resulting during the execution of the contract by the awarded contractor. To this purpose we computed RI over all the parameters audited for each contract and we calculated the correlation with the winning discount. Unfortunately, data at our disposition were limited, so we could conduct this analysis just on 120 auction/contracts out of 222. From Figure 10, which plots the discount/reputation combinations, it appears that there is no correlation between discounts and quality/safety of works, notwithstanding both series showed positive and significant breaks at the announcement dates. The correlation is equal to 0.098 and not statistically significant, while using a simple linear regression the R^2 is less than 1% (see Figure 10). Apparently, this means that improvements in quality and safety has come as a free lunch to the Firm.

Table 3

	Announcements				
T-test	**1st** 20 dec 07	**2nd** 4 apr 08	**3rd** 10 jul 08	**4th** 21 oct 08	**5th** 16 jan 09
Overall	–	–	–	–	–
Main Dimensions					
Safety (S)	–	–	–	–	–
Quality (Q)	–	–	–	–	–
Specific Safety and Quality Dimensions					
S - Equipment and machinery	–	–	–	–	–
S – Documentation	–	–	–	–	–
S - Works execution	–	–	–	–	–
Q - Works on joints	+	0	–	–	–
Q - Customer relationship management	+	+	–	–	–
Q - Air works	n.a.	–	–	0	0
Q - Underground works	–	–	–	–	–
Q - Transformer station works	0	0	0	0	0
S - Personnel	–	–	–	–	–
S - Works site regularity	–	–	–	–	–
S - Works site safety	–	–	–	–	–
S - Works site controls	n.a.	n.a.	n.a.	0	0
Most awarded and audited firms					
AC	–	–	–	–	–
AF	–	–	–	–	–
AI	–	–	–	–	–
AL	–	0	–	–	–
D	–	–	–	–	–
G	–	–	–	–	–
L	n.a.	–	–	–	–
M	–	–	–	–	–
U	–	–	–	–	–
Auction discount	–	–	–	–	–

Legend:
–/+ = score before the announcement is significantly (5%) lower/higher than after
0 = score not significantly different
n.a. = not available
Each test is run between the group of all parameters inspected before the specific announcement and the group of all parameters inspected after the specific announcement, relating to the category reported in each row.

Table 4

```
Probit regression                                    Number of obs   =       64537
                                                     Wald chi2(21)   =    14926.65
Log pseudolikelihood = -35158.621                    Prob > chi2     =      0.0000

------------------------------------------------------------------------------
             |               Robust
   parametro |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      weight |   .0085191   .0031194     2.73   0.006     .0024052    .0146329
    announc1 |   .4449855   .0347929    12.79   0.000     .3767926    .5131783
    announc2 |   .1315237   .0215823     6.09   0.000     .0892232    .1738242
    announc3 |   .4341951   .0193842    22.40   0.000     .3962027    .4721875
    announc4 |    .218222   .0223362     9.77   0.000     .1744438    .2620002
    announc5 |    .059171    .018998     3.11   0.002     .0219356    .0964065
        dim1 |   .3535574   .0536464     6.59   0.000     .2484124    .4587024
        dim2 |  -.7230065   .0493303   -14.66   0.000     -.819692   -.6263209
        dim3 |  -.2012378   .0527846    -3.81   0.000    -.3046937   -.0977819
        dim4 |   .6826454   .0755167     9.04   0.000     .5346353    .8306555
        dim5 |   .5061796   .2343647     2.16   0.031     .0468331     .965526
        dim6 |   1.217221   .2937219     4.14   0.000     .6415364    1.792905
        dim7 |  -.5762149   .0479084   -12.03   0.000    -.6701137   -.4823161
        dim8 |   1.840798   .3615874     5.09   0.000     1.132099    2.549496
        dim9 |  -.7522982   .0550745   -13.66   0.000    -.8602422   -.6443542
       dim10 |  -.9344627   .0500092   -18.69   0.000    -1.032479   -.8364465
       dim11 |   -.538382   .0524455   -10.27   0.000    -.6411733   -.4355907
       dim12 |   -.604407   .0868998    -6.96   0.000    -.7747274   -.4340866
      group1 |  -.0034356   .0313599    -0.11   0.913    -.0648999    .0580288
      group2 |  -.0193238   .0328206    -0.59   0.556     -.083651    .0450034
      group3 |  -.0727709   .0337456    -2.16   0.031    -.1389112   -.0066307
------------------------------------------------------------------------------
```

Legend:

dim1	= S - Equipment and machinery	group1	=	9 contractors awarded with 10 or more contracts
dim2	= S - Documentation			
dim3	= S - Works execution	group2	=	10 contractors awarded between 4 and 8 contracts
dim9	= S - Personnel			
dim10	= S - Works site regularity	group3	=	15 contractors awarded 2 or 3 contracts
dim11	= S - Works site safety			
dim12	= S - Works site controls			
dim4	= Q - Works on joints	announc1	= 20/12/2007	
dim5	= Q - Customer relationship mgnt	announc2	= 04/04/2008	
dim6	= Q - Air works	announc3	= 10/07/2008	
dim7	= Q - Underground works	announc4	= 21/10/2008	
dim8	= Q - Transformer station works	announc5	= 16/01/2009	

Figure 10
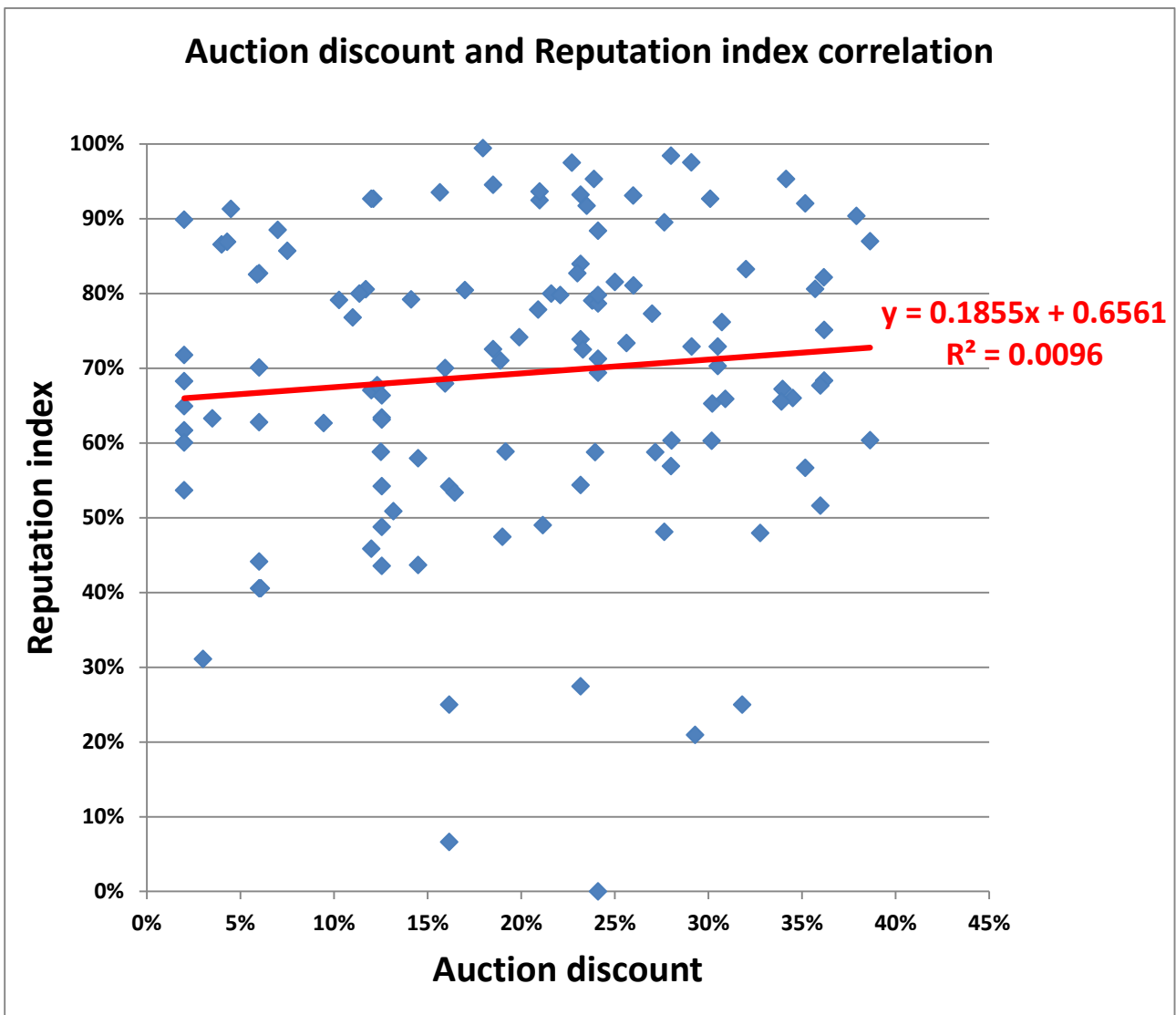
Legend: Each blue dot identify one contract, to which they are associated the auction discount offered by the winning contractor (on the x-axis) and the score calculated on all parameters inspected throughout the same contract life (on the y-axis). The red line is the linear regression line calculated out of the 120 auction discount / reputational score combinations, where the reputational score is the dependent variable and the auction discount is the independent variable. The auction regression coefficient, 0.1855, is not statistically significant (p value = 0.29).

4. Discussion and conclusions

The required performance from contractors could in principle be governed contractually, but contract enforcement is very slow and costly in Italy. Moreover, managers in charge of contract management found it difficult to exercise explicit contractual sanctions without worsening the prospects of long-term cooperation with suppliers. Hence other mechanisms are required.

This paper has studied the merits of using a vendor rating system data at the awarding stage as a mechanism to spur higher efforts from contractors when executing the contract on the basis of an experiment run by an Italian corporation listed on the Italian Exchanges. The results of the

experiment has shown a strong significant upward trend in Safety and Quality performance after the firm has announced the future use of this reputation mechanism to award contracts and after all the subsequent announcements. Interestingly, this was also true for the auction discounts series. However, there was no correlation at all between auction discounts and Safety/Quality improvements, the latter apparently occurring at no costs.

The vendors' average score emerging in the first audit was equal to 0.29, while the last audit we analyzed presented a score equal to 0.81. The first time vendors heard that they would be evaluated for their "quality", they immediately improved their performance, causing a break in the series. To some extent this may recall the well-known *Hawthorne effect.*[11] However, contrary to the Hawthorne effect, the improvement was not short-lived, even if we consider that the contractors could have stop to trust the Firm for the delayed implementation of the new awarding criteria and that it was easier for contractors to improve their score when the starting point was lower than later when the marginal cost to improve became higher. Indeed, some Safety and Quality parameters compliance requires very low investment costs: in particular for the Safety type, it happens that the regularity just calls for a greater level of care in running the works. For instance, the requirement for road signal in the vicinity of the work site (a parameter of the Safety type) is quite costless, and the same happens with other similar parameters. This may also explain part of the lack of correlation between the Safety and Quality improvements and the discounts.

Our results confirm that reputation can represent an important mechanism to exploit for buyers who have to rely on a relatively stable set of contractors. However, the experiment results are different from those ones from the traditional models showing that a reputational equilibrium can be sustained if the level of price is greater than the competitive one and such to guarantee a level of future rents greater than the immediate profits obtained from a cheating behavior. Indeed the Safety and Quality constant improvements occurred over the 2-year sample period were not correlated at all with the prices offered at the awarding stage.

Once the merits of this kind of reputation mechanism to improve contractors' are proven, many aspects remain open and give room for future researches: how to discipline the rating for new entrants, to structure the weights in the awarding criteria, and to opportunely choose the optimal "memory" of the indicator (i.e. how going backward for the calculation of RI).

---

[11] The Hawthorne effect is a form of reactivity whereby subjects improve or modify an aspect of their behavior being experimentally measured simply in response to the fact that they know they are being studied, not in response to any particular experimental manipulation.

References

Akerlof G.A., The market for lemons: quality uncertainty and the market mechanism, (1970), Quarterly Journal of Economics, 84

Allen F., Reputation and product quality, (1984), RAND Journal of Economics, 15

Atkeson A, Hellwig C. Ordonez G., Optimal Regulation in the Presence of the reputation Concerns, (2010), Mimeo.

Baar-Isaac Imperfect Commitment and reputation Commitment, (2005), Economics letters, 89

Bar-Isaac H., Tadelis S., Seller reputation, (2008), Foundations and Trends in microeconomics, 4

Board S., Vehn M., Reputation for quality, (2010), mimeo

Calzolari G., Spagnolo G., Relational contracts and competitive screening, (2009), CEPR discussion paper

Chou S.Y., Shen C.Y., Chang Y.H., Vendor selection in a modified re-buy situation using a strategy aligned fuzzy approach, (2007), Int. J. Production Research, 45

Cook W., Johnston A., Evaluating suppliers of complex systems: a multiple criteria approach, (1992), J. Operational Research Society, 43

Dellarocas C., Dini F., Spagnolo G., Designing reputation mechanisms, (2007), in N. Dimitri, G. Piga and G. Spagnolo: Handbook of procurement, Cambridge University Press

Dellarocas C., Reputation mechanisms, (2005), Handbook on Information Systems and Economics. Elsevier Publishing

Detoni A., Nassimbeni G., A method for the evaluation of suppliers' co-design effort, (2001), Int. J. Production Economics, 72

Doni N., The importance of reputation in awarding public contracts, (2006), Annals of Public and Cooperative Economics, 77

Ekmekci M., Sustainable reputation with rating systems, (2008), mimeo NYU

Ely J., Fudenberg D., Levine D.K., When is reputation bad?, (2008), Games and Economic Behavior, 63

Ely J., Valimaki J., Bad reputation, (2003), Quarterly Journal of Economics, 118

Fishman A., Rob R., Is bigger better? Investing in reputation, (2002), J. Political Economy, 113

Gordon S., (2008) Supplier Evaluation and Performance Excellence: A Guide to Meaningful Metrics and Successful Results, J.Ross Publishing

Heikkila J., From supply to demand chain management: efficiency and customer satisfaction, (2002), J. Operations Management, 20

Holmstrom B., Managerial incentive problems: a Dynamic persective, (1999), The Review of Economic Studies, 66

Horner J., Reputation and competition, (2002), American Economic Review, 92

Iossa E., Ray P., Building reputation for contract renewal: implications for performance dynamics and contract duration, (2009), Discussion paper

Kahraman C., Ruan D., Dogan I., Fuzzy group decision-making for facility location selection, (2003) Information Sciences, 157

Kangas A., Kangas J., Pykalainen J., Outranking methods as tools in strategic natural resources planning, (2001), Silva Fennica, 35

Kaplan R.S., Norton D., The balanced scorecard - measures that drive performance, (1992), Harvard Business Review

Kleijnen J.P.C., Smits M.T., Performance matrix in supply chain management, (2003), J. Operational Research Society, 54

Klein B., Leffler K., The role of market forces in assuring contractual performance, (1981), J. Political Economy, 89

Kranton R., Competition and the incentive to produce high quality, (2003), Economica, 70

Land C.K., Lovell C.A., Thore S., Chance-constrained data envelopment analysis, (1993), Managerial and Decision Economics, 14

Liu F.F., Hai H.L., The voting Analytic hierarchy process method for selecting supplier, (2005), Int. J. Production Economics, 97

Liu Q., Skrzypacz A., Limited records and reputation, (2009), Discussion paper

Mailath G.J., Samuelson L., Who wants a good reputation?, (2001), Review of Economic Studies, 68

Marchi G., Lenti L., La valutazione nei processi di piano - strumenti di trasformazione urbana, (2003), Angeli

Roodhooft F., Konings J., Vendor selection and evaluation: an activity based costing approach, (1996), European Journal of Operational research, 96

Saaty T.L., Decision making with the analytic hierarchy process, (2008), Int. J. Services Sciences, 1

Schmitz J., Platts K.W., Supply logistics performance measurement: indications from a study in the automotive industry, (2004), Int. J. Production Economics, 89

Shapiro C., Premiums for high quality products as returns to reputations, (1983), Quarterly Journal of Economics, 98

Stiglitz, J., (1987), Competition and the Number of Firms in a Market: Are Duopolies more competitive than atomistic markets, Journal of Political Economy, 95

Tadelis S. (1999)What's in a Name? Reputation as a Tradeable Asset, American Economic Review June 1999, 89.

Talluri S., Narasimhaen R., Nair A., Vendor performance with supply risk: a chance-constrained DEA approach, (2006), Int. J. Production Economics, 100

Trent R.J., Monczka R.M., Purchasing and supply management: trends and changes throughout the 1990s,(1998), Int. J. Purchasing and materials

Vanin P., Competition, reputation and compliance, (2009)

Wang J., Li H., Developing a decision model for supplier selection, (2006), Chinese Journal of Management Science

Weber C.A., Current J.R., Benton W.C., Vendor selection criteria and methods, (1991), E. J. Operational Research, 50

Weber C.A., Desai A., Determination of paths to vendor market efficiency using parallel coordinates representation: a negotiation tool for buyers,(1996), E.J. Operational Research, 90

Wu D., Olson T.L., A comparison of stochastic dominance and stochastic DEA for vendor evaluation, (2007), Int. J. Production Research

Yahya S., Kingsman B., Vendor rating for an entrepreneur development program: a case study using analytic hierarchy process method, (1999), J. Operational Research Society, 50