Buyer Quality and Procurement Outcomes: Explorative Evidence From the US[*]

Francesco Decarolis[†], Leonardo Giuffrida,[‡]

Elisabetta Iossa[§], Vincenzo Mollisi,[¶] Giancarlo Spagnolo[||]

October 26, 2016

Abstract

We explore empirically the impact of buyer quality on public procurement outcomes. Using purchases data (Federal Procurement Data System) and survey data (Federal Employee Viewpoint Survey) from US federal agencies, we find that procurement quality is highly heterogeneous across different agencies and persistent over time. The qualitative aspect that better predicts procurement performance is the perceived degree of cooperation within the unit, followed by the presence of appropriate incentives. We then asses the main channels through which public procurer quality affects outcomes. We find that buyer quality improves the selection of suppliers and strengthens the association between the use of cost plus contracts and the negotiation procedure.

Keywords: Buyer Quality, Competence, Cost and Time Overrun, Procurement.

JEL Classification: D44, H11, H57.

I Introduction

In a very influential paper, Bandiera, Prat and Valletti (2009) analyze purchases of standardized goods by Italian public bodies and find that some public bodies pay systematically more than others for equivalent goods. The variation in prices is principally due to variation in what they call "passive waste" (which does not involve an utility for the public buyer) rather than "active waste" (which involves an utility for the public buyer). Stated differently, and contrary to the common wisdom, even in a country where corruption is considered a major problem it is inefficiency the major source of excessive spending.[1]

A direct implication of this is that more attention needs to be posed on the quality of contracting authorities, i.e. on the ability of public buyers to perform their job effectively. Focusing on the optimal incentives for suppliers, as the Principal Agent literature does, or on optimal auction design, as part of the procurement literature does, will remain of second order importance if procurement units do not appropriately choose and manage the mechanisms to screen suppliers and the contracts to incentivize them.

Confirming these concerns, Saussier and Tirole (2015) report about a recent study by the *Union des groupements d'achats publics* (UGAP, French Public Procurement Grouping Union) revealing that 63% of French public buyers do not have a legal profile and 61% of public buyers joined a purchasing department following a period of internal mobility, with no prior experience in the field. Only 39% of public buyers undertook some form of course or training resulting in qualification in the field of purchasing. Finally, the study shows that over two-thirds of buyers acknowledge the fact that they are not very familiar with the economic and industrial fabric and nearly half admit that they do not monitor economic or technological developments.

The issue is of paramount importance for policy makers because of the sheer size of public procurement markets, estimated around 15% of world GDP (e.g. World Bank, 2015[2]), and

---

[1]In their estimates, passive waste accounts for over 83 percent of total estimated waste, and if all public bodies were to pay the same price as the ones at the 10th percentile, expenditure would fall by 21%, saving 1.6-2.1% of Italian GDP.

[2]In high-income countries public procurement averages 12% GDP and about 29% of total general government expenditure. In fiscal year 2013 alone, federal procurement spending in the United States topped

because procurement policy is increasingly seen as a strategic tool to foster other important long-term government objectives like innovation, social inclusion and environmental sustainability. And indeed, the heterogeneous quality of contracting authorities and their different ability to perform their task effectively is now gaining increasing attention among policy makers.[3]

Following this debate and policy changes, in this paper we undertake an explorative empirical analysis of the impact of public buyers' quality on public procurement outcomes in the US. We consider a large dataset of purchases by all US federal agencies contained in the Federal Procurement Data System (FPDS), a system tracking every new awarded federal contract worth more than $2,500, as well as every follow-on contracting action, such as a renewal or modifications. These data are publicly available and include records from fiscal year 2000 to present.[4] They contain information on about 40,000,000 actions and on several variables related to each transaction, including amount/value, purchaser, contract, and contractor information, the agent place of performance, the service information, record, and competition information, contractor characteristics, legislative mandates. We supplement this dataset with those from the Federal Employee Viewpoint Survey (FEVS), that measures government employees' perceptions on several characteristics of the agency they are employed by and of the specific unit/bureau they belong to. This large survey has been administered on average every two years between 2002 and 2015 covering almost every US government agency, and includes questions divided into five different sections: my work experience, my work unit, my agency, my satisfaction and work/life.

We first exploit the time dimension of the FPDS data to construct measures of agency and bureau procurement quality based on their past ability to effectively obtain the procured

US$460 billion, according to the Office of Management and Budget, a number that roughly equates to the combined GDPs of Chile and the Czech Republic. In developing countries the fraction of public procurement to GDP averages 15-20% and public procurement expenditure may reach up to 60-70% of all government expenditure (World Bank 2015).

[3]In Europe, for example, the latest procurement Directives, the law regulating purchases by national governments of goods, works or services above a certain threshold, has been followed by national transpositions which pose procurer competency at the forefront of procurement guidelines. For example, in Italy, a qualification system for public buyers is being introduced. It requires procurers to have (increasing) adequate skills, experience and expertise to be allowed tender increasingly complex contracts.

[4]See https://www.usaspending.gov. The reporting threshold for individual transactions was $25,000 before 2004. We employ the version of FPDS dated September 30, 2015.

goods and service within the initially contracted upon costs and time limits. We name these buyer quality measures Past Cost Performance (PCP) and Past Time Performance (PTP). Then we perform a first case study analysis of these buyer past performance measures and of some of the most relevant questions on the perceived measures from the FEVS.

We find that both the performance-based quality measures and the answers to closely related FEVS questions are highly heterogeneous across different agencies and quite persistent (though not invariant) in time. They do no correlate as well as one could have expected, but this is understandable given that the first measures are endogenous but "objective" as they are based on measured cost and time effectiveness, while the second ones are exogenous but reflect "subjective" perceptions of the employees.

We also find that past-performance measures vary substantially across different bureaus of the same agency. This suggests that the right organizational unit to look at is the bureau rather than the agency, so in the reminder of the analysis we focus on the more disaggregated bureau level performance measures and correspondingly on the survey questions relative to the bureau rather than the whole agency.

We then proceed to perform our econometric analysis trying first to quantify the effects of overall public buyers' quality, as measured by our past performance measures, on procurement outcomes. Then, we try to evaluate the role of the more specific, perceived bureau quality captured by the FEVS questions. Finally, we try to shed light on the importance of different channels through which buyers' quality may be affecting procurement outcomes, such as a correct choice of the award procedure in different procurements (e.g. auction vs negotiation) and the relative importance of selecting good quality contractors versus managing the post-selection contractual relationship with them (selection vs moral hazard).

We find that our past-performance based quality measures, both relative to costs and time of delivery, are significantly related to procurement outcomes. When deriving this result we control, among other things, for buyer experience, the size/complexity of the project, and the type of procurement contract (cost-plus vs fixed price), which have been shown to be important in previous empirical and theoretical work. According to our estimate, a standard deviation increase in PCP is associated with an increase of about 30 percent of a standard

deviation in CP. The result is similar, almost 40 percent, for the effect of PTP on TP. These results are robust across specifications.

We then move to analyze the more detailed FEVS perceived quality measures. We first validate the overall survey data by verifying the consistency of the general perception measures on overall buyer quality on procurement outcome. We find that indeed, using a question of the survey that captures the perception on the general quality of work in the unit we obtain results broadly in line with those obtained with our buyer past performance measures. Specifically, the sign and significance are essentially analogous to those obtained with PCP and PTP, but the magnitude is somewhat smaller, amounting to one third of the effect of PCP and PTP. We then use other questions of the survey to try to unpack different determinants of the perceived quality of the bureau. The FEVS questions we focus on ask about the intensity of incentives/rewards for good performance, of the degree of cooperation within the unit, and the ability to attract/retain highly skilled individuals.

We find that the qualitative aspects that contribute the most on predicting procurement performance are those related to the perceived degree of cooperation within the unit. This may not be all that surprising for procurement experts, given the very strong complementarities between legal, economic and engineering/market technical skills that exist in the procurement activity, but it is an interesting finding that points at the importance of having highly skilled managers running procurement units. A cooperative environment in the unit is a main objective/result of competent management, so this result suggests that more attention should be given to prevalent management practices in the public sectors. The presence of appropriate incentives also contributes significantly to predict procurement outcomes, though considerably less that the measure for cooperation, again pointing at the importance of management practices. These two channels are followed by the ability of the unit to attract/retain employees with the necessary skills, whose effect seems to be diluted by the previous two measures.

While we have to keep in mind that these are survey measures of perceived quality, and should therefore be considered as such, we believe that the consistency in the first two results and their robustness across specifications should be taken to indicate that improving

the quality of a purchasing unit requires not only selecting lower ranked employees with the right skills, but also, more importantly, ensuring that the unit is run by a highly skilled procurement managers, able to adopt appropriate procurement management practices that ensure the necessary degree of cooperation among different specialities (law, economics, engineering/management) and the motivation in the unit, a result that seems in line with recent findings on the importance of management practices in other branches of the public sector.

We then conclude trying to asses the main channels through which public procurer quality affects procurement outcomes. We mentioned that the effect of procurement quality on outcome could be driven by a competent choice of the awarding mechanism for different projects. We indeed find that buyer quality tends to strengthen the association between the cost plus nature of the reimbursement rule and the use of negotiation. Finally, to shed light on how much the effects of buyers quality on outcomes depend on a higher ability to select the "good contractor" (selection), or by a higher ability to monitor and enforce contractors (management), we construct a past-performance based measure of the "quality" of contractors appearing repeatedly in our data. We then estimate how much of the improved procurement outcome coming from buyer quality can be attributed to a more frequent selection of high quality suppliers rather than a better performance of all suppliers. We find that accounting for suppliers' ability in the regression substantially improves the fit of the model. The suppliers' ability measure is positively and significantly associated with better performance, in terms of both cost and time. Selecting a good seller thus appears to be a particularly important element of procurement performance, although some of the unexplained variation is also likely to be driven by how a bureau is able to interact with its contractors, regardless of their ability.

The rest of the paper unfolds as follows. In Section 2 we survey related literatures. Section 3 describes the two databases, summary statistics and the preliminary case-study analysis. Section 4 contains the main empirical analysis. Section 5 concludes.

II   Related literature

As mentioned earlier, our paper is most closely related to Bandiera, Prat and Valletti (2009) where the introduction of a central procurement agency is exploited as a natural experiment to identify the amount and the sources of public waste in Italian public procurement. They find that inefficiency is by large the most important dimension in explaining public waste, that there is enormous variance in waste across different public buyers, and that the public buyers that perform better (both in terms of active and passive waste) are those enjoying more autonomy/discretion. Our paper is in the same spirit, as we also try to quantify inefficiencies induced by public buyers of different quality, but in the US and for a wider range of services. Though we do not have a natural experiment to study, we exploit the long time dimension of our data to create past performance indicators reflecting buyers' quality and estimate the different performance they lead to. We also use the rich survey on US agencies and bureaus to try understand which characteristics of buyers' quality matter more.

Our work is also related to a strand of recent studies documenting the heterogeneity of managerial practices in the public sector and their importance for explaining the large differences in the efficiency of public management and public goods provision, both within and across countries. In particular, Bloom *et al.* (2015) extend their methodology developed for manufacturing firms (in Bloom and Van Reenen, 2007) and collect data on management practices in over 1,800 high schools in eight countries showing that higher management quality is strongly associated with better educational outcomes. Analogously, Bloom *et al.* (2014) collect data on management practices in 2,000 hospitals around the world, documenting large variations in performance linked to presence of better management practices. Most importantly for us, both studies document substantial within country variation in management practices and associated outcomes. Our findings are related and consistent with these findings, as we document substantial variation in the quality of US public buyers as measured by past performance and the self assessment survey, a significant association between higher quality public buyers and public procurement outcomes, and a tendency of the quality of management practice to be the most relevant features of public buyers quality.

6

Our work also contributes to the current discussion on competencies and professionalization in public procurement. In the management literature, employee competence refers to a person's underlying characteristics that are directly influencing related to job performance.[5] For specific jobs or organizations, scholars have developed competency models which define the competencies needed. In the case of procurement agencies, Tassabehji and Moorhouse (2008) and references therein found that technical skills have a direct impact on the ability of procurement professionals to fulfill their role proficiently, but the degree of organizational support and internal acknowledgement of the role's importance are a key ingredient for the development of procurement professionals. Our results appear to reinforce this view. Among the competencies we can relate our results to, those relative to "work coordination", encompassing those skills and abilities required to organize and prioritize work and work cooperatively with a group of people with different backgrounds appear the most important for procurement outcomes, followed by incentives and recognition.

Economists studying procurement have more or less explicitly recognized that more competent, higher quality procurers should choose better designed tenders, more efficient contract design, engage in more effective oversight or reduce red tapes, but a systematic analysis of these effects has not been yet undertaken. Two exceptions are Bajari, McMillan and Tadelis (2009) and Decarolis (2014). The former study auctions versus negotiation. Employing a dataset of private sector building contracts awarded in Northern California during the years 1995-2000, they find that project characteristics affect the choice of the award mechanism and that more experienced buyers use auctions more often.[6]The latter studies procurement outcomes in terms of ex post contract renegotiations and shows that they depend on the choice of the procurement mechanism and on the level of bid screening undertaken by the buyer. Large buyers, who are the most experienced, are better able to

---

[5]It is defined as the combination of knowledge, personal traits, skills and abilities, which affect the capability to successfully perform critical work tasks, specific functions, or operate in a given role or position. These factors are acquired through pre-service education, in-service training, hands-on experience, and the assistance of mentors and preceptors, and are affected by internal and external constraints, environments, relationships related to the job or occupation, motivations and perceptions of the work and ones self (see e.g. McClelland 1973 and Boyatzis 1982).

[6]They proxy expertise by public buyers with three variables: (i) a "cumulative experience buyer dummy", equal to 1 if the buyer has appeared in the dataset at least one previous project; (ii) the credit rating and (iii) organizational size.

screen offers, as shown by the better outcomes in terms of time and cost renegotiations for given contract choices. Our work is in line with these findings - we are able replicate their findings of Bajari et al. (2009) in our larger data set - and extends further our knowledge by studying additional quality features of public buyers besides those captures by experience, and the channels through which these may affect procurement outcomes.

More broadly, our paper is related to a strand of research on the determinant of public procurement outcomes. Other papers in this strand have studied the choice of procurement mechanism and procurement contracts without analyzing the role of buyer quality. For example, Lewis and Bajari (2011) studied the impact of scoring auctions with quality concerns in terms of time incentives in highway repair projects awarded by the California Department of Transportation between 2003 and 2008. Lewis and Bajari (2014) investigate how higher penalties for delay in delivery can induce greater effort, but can also increase the agent's risk in performing the contract. Guccio, Pignataro and Izzo (2014) employ data of public contracts for Italian Cultural Heritage conservation between 2000 and 2005 to test the effect of buyer specialization (contracting authorities with a greater percentage of in-house projects) on performances in terms of time and cost overruns. Other studies on the choice of procurement mechanism consider the level of transparency (Coviello and Mariniello, 2014), the degree of discretion (Coviello, Guglielmo and Spagnolo 2016), and the role of supplier past performance (Duflo and Banerjee 2000, Decarolis, Pacini and Spagnolo 2016).[7]

---

[7]At more general level, in the economic literature, the concept of competence has been studied theoretically in the context of political selection (Besley, 2005), leadership in firms (Bolton, Brunnermeier, and Veldkamp, 2008), decision making by judges (Levi, 2005), incentives to gather information (Dur and Swank, 2005), and information transmission and revelation (Burajade and Jullien, 2011). Empirically, the effect of policy maker characteristics on outcomes has been studied looking, for instance, at the impact on growth of the quality of leaders (Jones and Olken, 2005) or of their education (Besley Montalvo and Reynal-Querol, 2011), and the impact of age and years in office on regulators decisions (Grout, Zalewska and Garside, 2013). A channel, through which competence may improve performance is via its impact on oversight. More oversight may result in more red tapes, which paradoxically make an increase in oversight lead to worse rather than better outcome. In a recent paper (2016) Calvo, Cui and Serpa (2016) analyze the impact of oversight on cost and time overruns in the US, using the same dataset as ours. They find that decreasing the level of oversight significantly improves the efficiency of public projects; in particular, a project's delay time and cost overrun decrease by an average of 53.5 days (i.e., 15% of the total project duration) and $3,890 (i.e., 2.6% of the total project budget). Further, their results show that when there is high operational oversight, (i) experienced suppliers significantly outperform inexperienced suppliers and (ii) projects that offer performance-based contracts significantly outperform those that do not offer these incentives (i.e., fixed-fee contracts). Conversely, when there is low operational oversight, (i) inexperienced suppliers perform almost as well as experienced suppliers, and (ii) fixed-fee contracts perform almost as well as performance-based contracts. They do not however test for the effect of procurer competency. Recently, Giuffrida and Rovigatti

III Data

This section presents the two datasets that we combine in our empirical analysis. We first illustrate the main variables and then discuss the case of four Agencies to illustrate features of our measures.

A. Federal procurement data: The FPDS-NG dataset

The Federal Procurement Data System (FPDS) is a publicly available dataset that tracks all federal contracts awarded in the US. The data covers all contracts awarded from year 2000 to the present that are worth more than $2,500. The information contained regard is of two kinds: *a)* data concerning the awarding stage, like the identity of the bureau and Agency awarding the contract, the identity of the winner, the winning price, the contractual duration of the contract, and several details concerning the awarding procedure and *b)* data concerning the subsequent life of the contract, like each modification, the final price and the final delivery time. These data have already been used to research key features of the US public procurement system by several studies, including Goldman, Rocholl and So (2013) and Liebman and Mahoney (2013)[8]. Although the data contain contracts for both products and services, the former typically do not exhibit any ex post variation in price or delivery time. Thus, for our analysis we focus exclusively on the procurement of services.[9] The key dependent variables that we construct from these data concern the cost and time renegotiations characterizing the contract. We describe first these two dependent variables and then illustrate the main independent variables.

Cost Performance and Time Performance: Since we can observe the whole history of each contract, from the first signature to the end of the job, including each renegotiation, we can construct two performance variables that compare the final cost and time with the ones

---

(2016) investigate the oversight of public procurement contracts and distinguish between public and private checks. Their results are in line with Calvo et al. (2016) as regards the detrimental impact of public oversight and argue that this is the case if and only if the monitoring is exerted by low-competence contracting offices. Conversely, private oversight enhances contract outcomes and, following Calveras, Ganuza and Hauk (2004), the authors show that this effect stems from the screening mechanism exerted by the surety companies underwriting public projects.

[8]The version of FPDS employed dates back to September, 30 2015.

[9]For further details upon the sample selection, please refer to the corresponding subsection of Giuffrida and Rovigatti (2016).

initially declared at the awarding stage. Specifically, *Time Performance* (TP) for contract i is defined as:

$$TP_i = \frac{Expected\ time\ of\ completion\ of\ i}{Expected\ time\ of\ completion\ of\ i\ +\ Time\ delay\ of\ i}$$

By construction it maps the couple ($Expected\ time\ of\ completion_i, Time\ delay_i$) to the interval $[0, 1]$, with an increasing performance approaching 1 (i.e. in the case of no renegotiations on completion time). The *Cost Performance* (CP) index has the same structure and the same rationale applies:

$$CP_i = \frac{Expected\ cost\ of\ completion\ of\ i}{Expected\ cost\ of\ completion\ of\ i\ +\ Cost\ overrun\ of\ i}$$

These two performance measure are highly correlated, with a coefficient of linear correlation of 0.802. Nevertheless, as their scatterplot in Figure 1 clearly shows, these two proxies of contractual performance are not always moving together. In particular, the high density of observations concentrated where CP equals one indicate that for a substantial number of contracts completed with no cost overruns there are, nevertheless, large time delays. Indeed, the density of contracts located where TP equals one is not very high, thus suggesting that time delays are a more prevalent phenomenon relative to cost overruns. The figure also confirms the strong correlation between the two measures (most observations are clustered on the 45 degree line) and the presence of some threshold effects. The latter can be seen in the form of what appear like multiple vertical and horizontal lines. For instance, there is a clear vertical line at CP=0.5 and, analogously, a clear horizontal line at TP=0.5.

Regarding the independent variables, we are going to exploit data from both the FPDS and the other dataset described below. The variables that we construct form the FPDS are as follows. First, we are interested in measuring the persistency of performance over time, within the same purchasing organization and for the same category of procured good or service. Thus, for each contract awarded we measure the weighted average outcome (in terms of CP and TP, separately) of the contracts for the same service group awarded by the

Figure 1: Scatterplot of Cost Performance and Time Performance
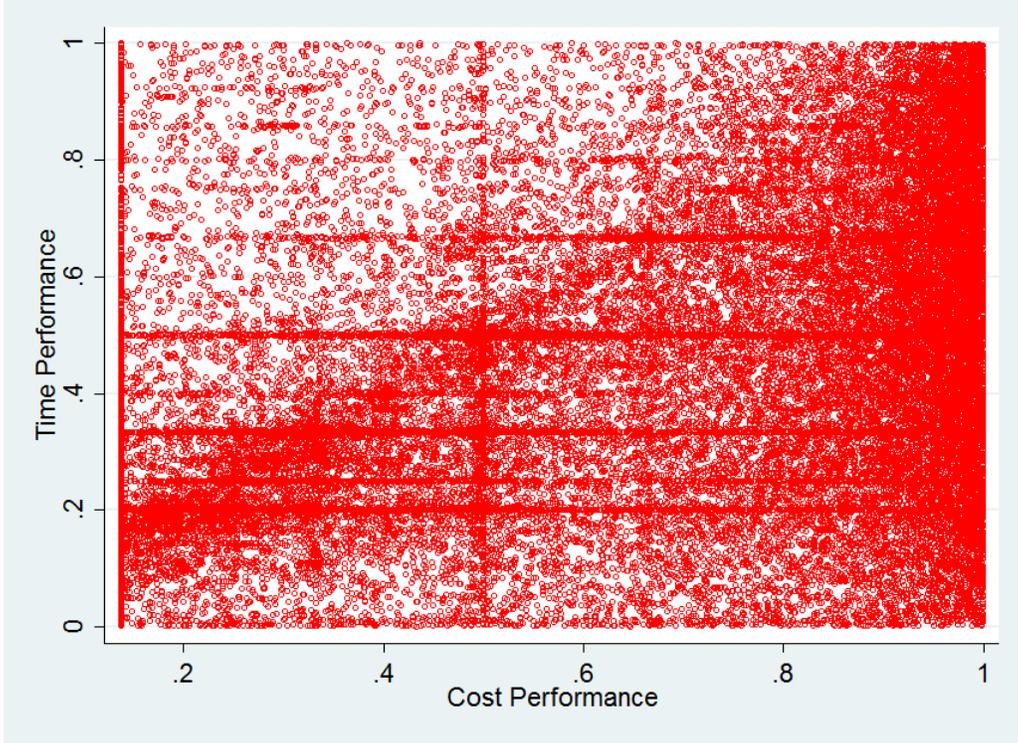
same bureau in the past. We differentiate by category not to homogenize the competence on different sectors within the same federal office. Therefore, we define *Past Cost Performance* (PCP) and of *Past Time Performance* (PTP) for contract i, belonging to contract category k and awarded at time j as follows:

$$PTP_{ijk} = \frac{\sum\limits_{j_{ik} < J_{ik}} w_{ijk} * TP_{ijk}}{\sum\limits_{j_{ik}} w'_{ijk}}$$

$$PCP_{ijk} = \frac{\sum\limits_{j_{ik} < J_{ik}} w_{ijk} * CP_{ijk}}{\sum\limits_{j_{ik}} w'_{ijk}}$$

Where both summations are take over contracts awarded by the same bureau awarding contract i and where, in order to give more importance to contracts that are closer in time, we have made use of the following Bartlett weights: $w_{ijk} = (1 - \frac{j_{ik}}{J_{ik}})$, for $j_{ik} = 1...J_{ik}$. To

11

make the weights sum up to 1, we normalize them such that:[10]

$$w'_{ijk} = \frac{(1 - \frac{j_{ik}}{J_{ik}})}{\sum\limits_{j_{ik}}(1 - \frac{j_{ik}}{J_{ik}})}$$

In addition to the two procurer performance measures (PTP and PCP), the other relevant variables that we obtained from the FPDS data all concern the awarding procedures. The main ones are the following four. *Negotiation*: This dummy variable reports whether contract uses negotiated procedures, i.e. contract is awarded on the basis of a direct agreement with a contractor, without going through the competitive bidding process. *Cost plus*: This dummy variable equals one if the contract is cost plus (i.e., the supplier is entitled to obtain compensation in proportion to the amount of inputs that it delivers) and zero if it is fixed price (i.e., the supplier is paid a fixed price, regardless of its actual delivered inputs levels). *Amount*: the (log of the) awarding price. *Buyer Experience*: the (log of the) number of times a buyer has appeared in the FPDS data set in the same service category. Following Bajari et al. (2008), the latter two variables are proxy respectively for a bureau's reputation (a cumulative owner experience measure) and contract complexity.

B. Federal Employee Viewpoint Survey: The FEVS dataset

We supplement our dataset of federal procurement contracts with the Federal Employee Viewpoint Survey (FEVS). This is a large survey aimed at measuring government employees' perceptions of whether, and to what extent, conditions characterizing successful organizations are present in their bureaus and agencies. Employees are surveyed in 2002, 2004, 2006, 2008, 2010, 2011, 2012, 2013, and 2015. The agencies that are invited to participate account for 97% of the executive branch workforce. Of the 839,788 employees who received the FEVS in 2015, 392,752 completed the survey, resulting in a final response rate of 46.8 percent. The FEVS consists of 85 questions divided into five different sections: my work experience, my work unit, my agency, my satisfaction and work/life.

---

[10]Indeed, it is the case that: $\sum\limits_{j} w'_{ijk} x_{ijk} = \frac{\sum\limits_{j} w_{ijk} x_{ijk}}{\sum\limits_{j} w_{ijk}}$.

In principle, the survey contains a multiplicity of questions that might be interesting to relate to procurement performance. First of all, there are questions at the level of the bureau and at the level of the agency. Since our procurement level is at the level of the bureau, it seems natural to focus on this finer level of aggregation and, indeed, the results discussed next will confirm the greater relevance of the bureau-level survey responses. Second, within the set of questions asked regarding the bureau, we can roughly divide them in four types. There are questions related to the degree of cooperation (e.g., Q20 asks "The people I work with cooperate to get the job done"), to the skill level (e.g., Q21 asks: "My work unit is able to recruit people with the right skills"), to incentives (e.g., Q23 asks "In my work unit, steps are taken to deal with a poor performer who cannot or will not improve") and to the overall performance (e.g., Q28 "How would you rate the overall quality of work done by your work unit?"). For the first three categories, the survey contains other questions that could be used, but we focus on the four questions (Q- 20, 21, 23 and 28) based their greater statistical relevance, as explained in detail below. Q28, instead, is particularly interesting as this is the only question in the survey that can proxy for an ex-post, self-evaluation of the overall work conducted by each single work unit of each agency. Employees may respond in five ordered different levels: very poor, poor, fair, good, very good. The share of unanswered questions is very low, on average 0.50%, and the median response is "good". Finally, to distinguish bureau vs agency features we will also use Q29 from the section "my agency" which asks "The workforce has the job-relevant knowledge and skills necessary to accomplish organizational goals". This agency-level question is more related to the pool of competence and skill retained by each agency and thus we consider it as a proxy for the overall, self-perceived quality of the agency. Employees can respond either "do not know" or in 5 ordered levels: strongly disagree, disagree, neither agree nor disagree, agree, strongly agree. The answer's median is "agree". The percentage of "do not know" is 1.5% in average and the missing percentage is not greater than 2.2%. The correlation between the bureau and agency level outcome proxies (Q28 and Q29) is high 0.765, but far away from 1. We will first describe the summary statistics of the analysis sample and then return to the issue of how these outcome variables are related to each other and to the ones created from the FPDS.

To better understand the content of this survey and how it relates to the FPDS data, we explore the case of four agencies that are at the extremes of the performance as measured by Q28. The top two agencies in terms of Q28 are NASA, with an average Q28 equal to 0.86, and P&T (Patents and Trademarks), with average Q28 equal to 0.84). The worst two are CBP (Custom and Border Protection), with an average Q28 equal to 0.75, and ICE (Immigration and Customs Enforcement), with an average Q28 equal to 0.77. In Figure 2, we plot for each of these four agencies the evolution of the four "quality" proxies: Q28, Q29, PTP and PCP. The evidence reveals that not for all agencies these four proxies deliver a consistent result. For instance, while ICE is the worse (or the second worse) along all measures, the CBP has a self assess, bureau-level performance (Q28) that is very low but a cost performance that is superior to that of all other three agencies. This difference could result from a multiplicity of factors, like the difference between self-assess and objective performance measures, or by the different role that procurement performance has for employees to judge the degree of "overall quality of work done by your work unit."

Figure 2: Dynamics of the performance measures
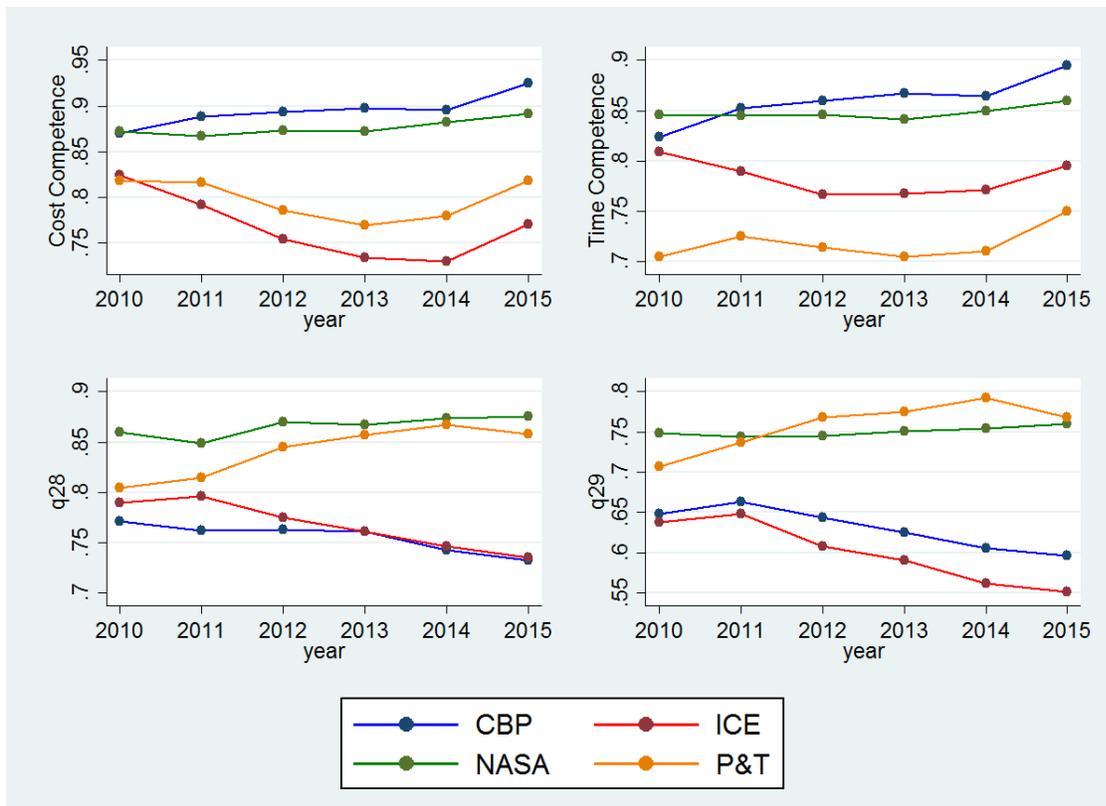
C. Summary statistics

Our analysis sample combines FPDS and FEVS datasets. For our analysis, we will focus on the years where the FEVS has an yearly frequency and where the two datasets, FEVS and FPDS, overlap. Thus, we focus on the years between 2010 and 2015. However, to construct our two measures of past performance from the FPDS (PTP and PCP), we go back in time until the very beginning of the FPDS in year 2000. In Table 1, we report selected descriptive statistics of the service subsample of contracts which amounts to 529764 observations. Each observation in our database refers to a single project. There are 25,516 projects.[11] Auctions represents the 12% of these contracts. The average award per contract and its duration are respectively 196.4 thousand dollars and 175 days. Both characteristics have a lower median, respectively 17.9 thousand dollars and 95 days. Cost plus is only employed for a small fraction of contracts, 6%. The past performance measures from the FPDS have, respectively a mean of 88% out of 1 and 80% out of 1. On the other hand, the performance measures from the FEVS have relative lower mean and median 0.81 for question Q28 and 0.69 for question Q29. Comparing the latter four variables, which we can all consider as different proxies of the procurer "quality," we find the former more dispersed with a sd of 0.24 and 0.32. Regarding the other variable in Table 1 we observe, for instance, that the average number of contracts undertaken by each contracting authority before to sign a new contract is 15948. Moreover, 56% of sellers in the sample are small business.[12]

We can further supplement the descriptive analysis with a series of scatterplots of the main variables. Figure 3 illustrates how the past performance measures (aggregated at the level of the agency) comove with the overall performance measures from the survey. Since past experience is also sometimes considered as a proxy for quality of the administration (the idea being that procurers holding more contracts accumulate greater experience and this helps to improve performance), we also report scatterplots of this variable (aggregated at the level of the agency) with Q28 and Q29. Overall, all these scatterplots are not indicative

---

[11]These are the projects that respect the definition of Negotiated and Auctioned contracts proposed by BMT.

[12]Small Business Authority (SBA) define firm as small based on the particular service category, which the contract belongs to, and sellers' characteristics like revenues and the number of employees. SBA revises this list every five years.

Table 1: SUMMARY STATISTICS: SERVICES

	(1) Mean	Median	S.D.	Min.	Max.	N
Negotiation	0.82	1	0.39	0	1	21620
Cost Performance	0.88	1	0.24	0.14	1	515502
Time Performance	0.80	1	0.32	0.00027	1	515502
Past Cost Performa e	0.89	0.91	0.083	0.14	1	515502
Past Time Performa e	0.82	0.82	0.093	0.0036	1	515502
Cum. Category Expe e	8.73	8.90	1.69	0.69	11.6	515502
Perceived Quality	0.81	0.81	0.027	0.73	0.88	515502
Cooperation	0.71	0.72	0.036	0.61	0.89	515623
Skill	0.54	0.55	0.039	0.38	0.83	515623
Incentive	0.46	0.45	0.048	0.29	0.67	515623
Project Value	214.5	18.5	1999.4	3	479000	515502
Cost Plus	0.044	0	0.20	0	1	515502
Small Business	0.58	1	0.49	0	1	515491
Limited Company	0.13	0	0.33	0	1	515502
Expected Length	185.7	114	185.5	1	1368	515502

of a particularly strong relationship, at the agency-level, between the different measures. This confirms the correlation that we derived from exploring the four-agency case reported above.

Indeed, a first takeaway from this initial analysis of the data is that the most adequate level at which we shall link the FPDS and FEVS datasets is that of the bureau, and not that of the agency. This is most clearly illustrated by Figure 4 where we report the distribution of PCP and PTP across all bureaus and agencies. For each agency, we report the performance of all bureaus with which the agency appears in the FPDS. The length of the horizontal lines measure the performance of Past Cost Performance (left) and Past Time Performance (right). It is clear that, despite there is some variation at the agency level, most of the action takes place between bureaus within agency. This is particularly the case for the time performance measure.

Figure 3: Scatter plots of the main variables
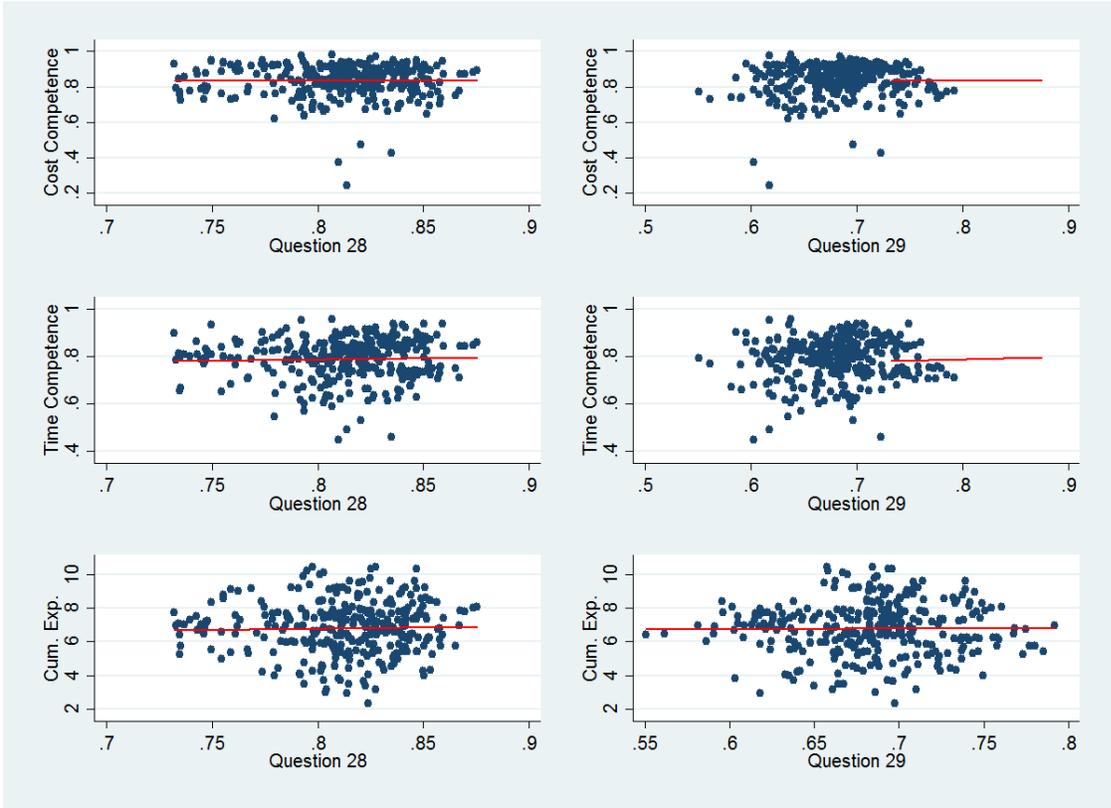
IV   Empirical Analysis

This section presents our empirical analysis and is subdivided in three parts. First, we analyze the persistency of performance within bureaus. Second, we look at the bureaus characteristics relating them to procurement outcomes. Finally, we look at how these outcomes are associated with procurers' choices in terms of suppliers' selection and types of awarding procedures used.

A. Persistency in buyers' performance

We begin our analysis from the basic, but fundamental question of whether contractual performance is a persistent feature of a bureau. To quantitatively address this issue, we estimate a linear regression model where the dependent variable, *performance*, is either CP or TP and the main independent variable is a bureau-specific *past performance* measure (either PCP or PTP). Specifically, for CP we estimate the following model:

Figure 4: Distribution of PCP and PTP across Bureaus and Agencies

(a) Cost Performance   (b) Time Performance
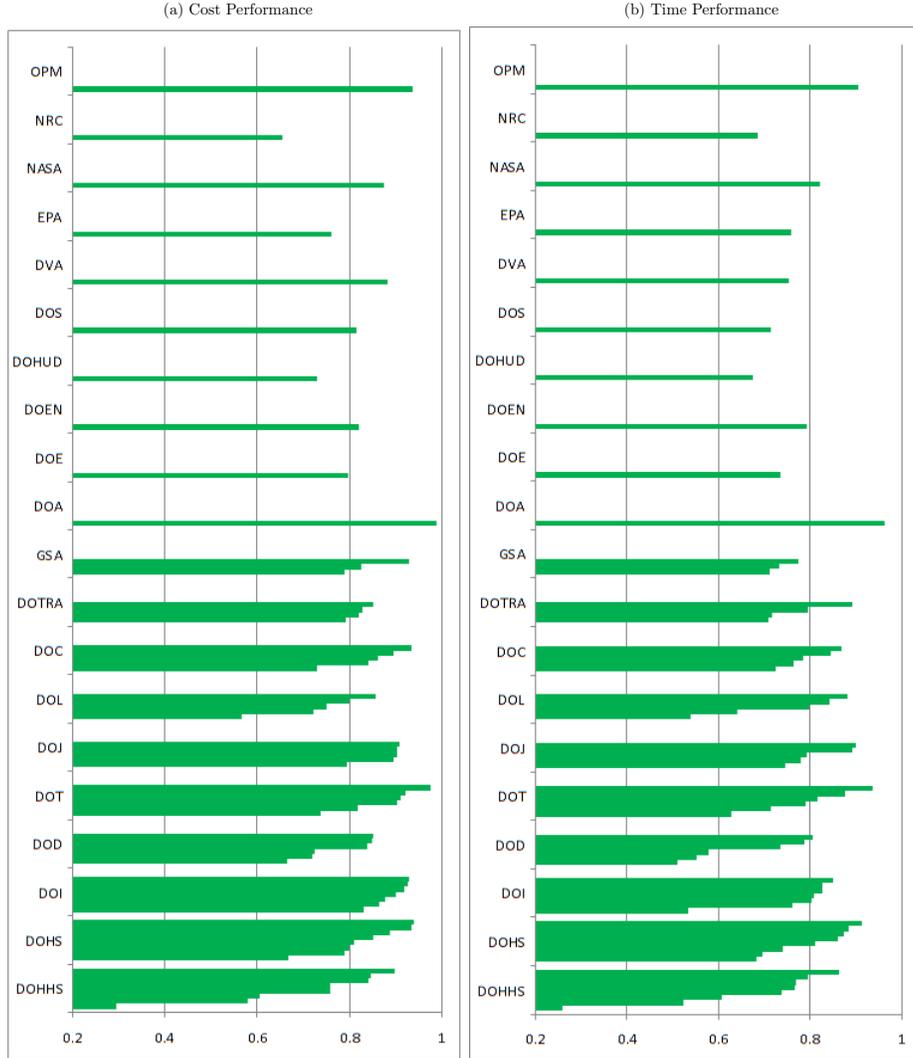
$$PC_{ijt} = \beta PCP_{jt} + \gamma Amount_i + \delta Experience_{jt} + \zeta CostPlus_i + \theta X_{ijt} + \epsilon_{ijt},$$

where X contains additional covariates. An analogous model is estimated for PTP. A positive coefficient ($\beta > 0$) would indicate persistency in performance. Regarding the other covariates, the complexity of works (proxies by *amount*, has an unambiguous meaning and we expect a negative effect of complexity on performances (i.e., negative γ). The greater experience of bureaus holding more contracts should have a positive impact on performance. Thus, we expect a positive sign for the effect of buyer's experience on performances (i.e., positive δ). Finally, when buyers prefer cost-plus contracts, they are losing effectiveness to

18

incentivize effort of sellers carrying out a contract and, hence, we expect a negative sign for the cost plus coefficient ($\zeta < 0$).

Table (2) and Table (3) present results for CP and TP respectively. Across all regressions we add controls in terms of fixed effects for year, bureau, service type and US state. Standard errors are estimated using the Huber-White sandwich estimators and clustered at (bureau * year) level. Model (1) includes PCP only, while the following models gradually expand the specification to include controls for cost plus contracts, the project value and the bureau's experience. Across all specification, the estimated effect of PCP remains positive and highly statistically significant. An analogous result is found for PTP. This indicates that bureaus with a good performance in the past are more successful in having contracts completed within what contractually agreed upon. Regarding the other covariates, cost plus is negative (as expected) and significant and so is the project value. Experience is, instead, positive but its statistical significance is only at the 10 percent level for CP, while it is at the 5 percent level for TP. Finally, the fact that despite the rich set of FE we observe a low explained variance indicates that it is hard to account for what drives contractual performance. In the next section we enter more in depth into this issue.

B. Buyers' characteristics and procurement outcomes

Having established that performance is a persistent trait across bureaus, we now seek to highlight what could explain this in terms of both bureaus features and actions. Since all of our measures of bureau's features are, however, obtained from the FEVS data, a preliminary step consists in showing that this survey's responses are indeed connected with our measures of performance in federal procurement. We have already seen, when discussing Figure 2, that also the overall measures of bureau's quality (Q28) tend to be stable over time. However, we have also discussed there that the PCP and PTP measures do not induce ranking of procurers's performance that are identical to those implied by Q28. Therefore, we begin the analysis by assessing the relationship between CP and TP with Q28 and then move on to the more specific bureau's characteristics measured by the FEVS data.

B.1. Bureau's overall quality

Table 2: Persistency of Cost Performance

	(1)	(2)	(3)	(4)
Past Cost Performance	0.95***	0.89***	0.84***	0.84***
	(0.030)	(0.030)	(0.030)	(0.030)
Cost Plus		-0.16***	-0.13***	-0.13***
		(0.011)	(0.011)	(0.011)
Log-Project Value			-0.023***	-0.023***
			(0.0014)	(0.0014)
Cum. Category Experience				0.32*
				(0.17)
N	334757	334757	334757	334757
R-squared	0.12	0.14	0.16	0.16
Year FEs	Yes	Yes	Yes	Yes
Bureau FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Product FEs	Yes	Yes	Yes	Yes

To assess the relationship between CP and TP with Q28, we estimate the same linear models described above, but with Q28 replacing PCP and PTP. Table 4 and Table 5 report the results which show that Q28 is indeed positively and significantly associated with both CP and TP. This indicates that the self-perceived measure of overall work quality that Q28 tries to elicit is indeed associated with contract performance, as measured in this study. We take this as an indication that we can rely on the survey data to try to related bureaus' performance with their characteristics.

Table 3: PERSISTENCY OF TIME PERFORMANCE

	(1)	(2)	(3)	(4)
Past Time Performance	0.99***	0.97***	0.90***	0.90***
	(0.053)	(0.052)	(0.047)	(0.046)
Cost Plus		-0.11***	-0.063***	-0.065***
		(0.010)	(0.0099)	(0.0099)
Log-Project Value			-0.036***	-0.036***
			(0.0020)	(0.0020)
Cum. Category Experience				0.92***
				(0.26)
N	334757	334757	334757	334757
R-squared	0.09	0.09	0.12	0.12
Year FEs	Yes	Yes	Yes	Yes
Bureau FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Product FEs	Yes	Yes	Yes	Yes

Table 4: Cost Performance and Self Perceived Quality (Q28)

	(1)	(2)	(3)	(4)
Perceived Quality	0.88**	0.85**	0.71*	0.71*
	(0.42)	(0.41)	(0.37)	(0.36)
Cost Plus		-0.18***	-0.15***	-0.15***
		(0.012)	(0.011)	(0.011)
Log-Project Value			-0.025***	-0.025***
			(0.0014)	(0.0014)
Cum. Category Experience				0.50**
				(0.25)
N	334757	334757	334757	334757
R-squared	0.09	0.11	0.13	0.13
Year FEs	Yes	Yes	Yes	Yes
Bureau FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Product FEs	Yes	Yes	Yes	Yes

Table 5: TIME PERFORMANCE AND SELF PERCEIVED QUALITY (Q28)

	(1)	(2)	(3)	(4)
Perceived Quality	1.77**	1.74**	1.52**	1.53**
	(0.80)	(0.79)	(0.71)	(0.71)
Cost Plus		-0.12***	-0.075***	-0.076***
		(0.012)	(0.011)	(0.011)
Log-Project Value			-0.039***	-0.039***
			(0.0020)	(0.0020)
Cum. Category Experience				0.96***
				(0.27)
N	334757	334757	334757	334757
R-squared	0.06	0.06	0.10	0.10
Year FEs	Yes	Yes	Yes	Yes
Bureau FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Product FEs	Yes	Yes	Yes	Yes

B.2. Bureaus' characteristics

The bureaus' characteristics that the FEVS data offers proxies for are bureaus' skills, incentive structure and level of internal cooperation. To assess their linkage with procurement performance, we use them in the linear models described earlier to replace past performance, with the idea of verifying whether any of these characteristics is able to account for why performance appeared to be a persistent trait of the buyers. The results are reported in Table 6 for CP and Table 7 for TP. For each ones of the three proxies for cooperation (Q20), skills (Q21) and incentives (Q23), we first introduce them individually in the model, and then consider a model with all the three together. It is interesting to not that, at least in terms of explained variance, all models in Table 6 and 7 have no less explanatory power than the ones including PCP and PTP. Nevertheless, the overall explained variance remains low with an R^2 of 13 percent in Table 6 and 10 percent in Table 7. Interestingly, this value remains constant within both tables as we gradually change the proxy used. This is in line with the idea that the wording of all the questions does not make completely obvious how to use them to isolate different, specific features. Nevertheless, some differences between the three proxies exists and, for both cost and time performance, it is cooperation in the workplace (Q20) that systematically has a positive and statistically significant role, regardless of whether it enters the regression alone, model (1), or along with the other two proxy variables, model (4). The negative effect of skills shall be interpreted carefully for at least three reasons. First, as mentioned above, the wording of Q21 is such that this variable is incapable of measuring whether the types of skills present in the bureau are at all adequate relative to the procurement process. Second, regressions (not reported here) with a different question asking whether the skill level of the bureau had improved recently indicate a positive association with CP and TP. Thus, a positive role of skill improvements on procurement outcomes seems present. Finally, if we substitute the measure of skill from the FEVS data with a different measure of bureaus' skill level developed by Clinton, Lewis and Richardson (2016) using the *2014 Survey on the Future of Government Service*, we obtain different results with skills being not statistically significantly associated with CP, but positively associated in a statistically significant way with TP. Although this latter survey

24

covers a single year, the skill measure it contains is likely more precise for our purposes.

Table 6: Cost Performance on Sources of Buyer's Heterogeneity

	(1)	(2)	(3)	(4)
Cost Plus	-0.15***	-0.15***	-0.15***	-0.15***
	(0.011)	(0.011)	(0.011)	(0.011)
Log-Project Value	-0.025***	-0.025***	-0.025***	-0.025***
	(0.0014)	(0.0013)	(0.0014)	(0.0014)
Cum. Category Experience	0.49*	0.50**	0.51**	0.49*
	(0.25)	(0.25)	(0.25)	(0.25)
Cooperation	0.88***			1.18***
	(0.23)			(0.33)
Skill		-0.25**		-0.47***
		(0.13)		(0.090)
Incentive			0.39***	-0.039
			(0.13)	(0.17)
N	334757	334757	334757	334757
R-squared	0.13	0.13	0.13	0.13
Year FEs	Yes	Yes	Yes	Yes
Bureau FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Product FEs	Yes	Yes	Yes	Yes

Table 7: Cost Performance on Sources of Buyer's Heterogeneity

	(1)	(2)	(3)	(4)
Cost Plus	-0.076***	-0.077***	-0.076***	-0.076***
	(0.011)	(0.011)	(0.011)	(0.011)
Log-Project Value	-0.039***	-0.039***	-0.039***	-0.038***
	(0.0020)	(0.0021)	(0.0020)	(0.0020)
Cum. Category Experience	0.93***	0.95***	0.97***	0.94***
	(0.27)	(0.27)	(0.27)	(0.26)
Cooperation	1.73***			1.81***
	(0.47)			(0.53)
Skill		-0.59**		-1.06***
		(0.26)		(0.18)
Incentive			0.81***	0.25
			(0.27)	(0.32)
N	334757	334757	334757	334757
R-squared	0.10	0.10	0.10	0.10
Year FEs	Yes	Yes	Yes	Yes
Bureau FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Product FEs	Yes	Yes	Yes	Yes

C. Buyers' choices and procurement outcomes

Not only given buyers' characteristics, but also some of their specific actions might contribute to explain their heterogeneous performance. In this section we focus on two of these actions that are observable through our dataset: the use of negotiations vs auctions and the selection of more effective sellers.

C.1. Choice of the procurement format

As discussed in section 2, a major channel identified by the literature in terms of what drives adequate procurement practices is the need to use negotiations for certain types of contracts and auctions for the others. In this part of our analysis, we present estimates of the following probit model for the use of negotiations vs auctions:

$$Negotiation_{ijt} = \gamma Amount_i + \delta Experience_{jt} + \zeta CostPlus_i + \theta X_{ijt} + \epsilon_{ijt}$$

Where i stands for the contract, j for the bureau, and t represents time. The results are presented in Table 8 where the first three columns report the results obtained for the subsample of low-quality bureaus and the last three columns use the subsample of high-quality bureaus. The sample split is performed by separating bureaus with a value of Q28 above or below the sample median Q28 (across bureaus). Since we know from earlier studies that the use of negotiations should be positively associated with cost plus, the project value and bureaus' experience, then we can use this sample split to verify whether, and to what extent, bureaus with high and low self perceived work quality differ. We find that the estimates for the two samples are quite similar, with the exception of the cost plus estimate in the low quality bureaus being not statistically significant. This provides some support for the idea that bureaus of different quality might obtain different procurement outcomes due to heterogeneity in how they select the procurement method to use.

Table 8: NEGOTIATIONS AND HIGH/LOW SELF PERCEIVED QUALITY

	(1)	(2)	(3)	(4)	(5)	(6)
Cost Plus	0.14	0.12	0.12	0.11***	0.10***	0.099***
	(0.090)	(0.085)	(0.084)	(0.022)	(0.022)	(0.021)
Log-Project Value		0.015**	0.015**		0.012***	0.012***
		(0.0074)	(0.0073)		(0.0044)	(0.0043)
Cum. Category Experience			2.46			1.22
			(1.89)			(0.93)
N	9845	9845	9845	7398	7398	7398
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Bureau FEs	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Product FEs	Yes	Yes	Yes	Yes	Yes	Yes

C.2. Choice of suppliers

Finally, the last channel that we explore regards whether the bureaus' heterogeneity is associated with how they deal with different sellers. Good contractual performance might emerge from bureaus' being able to select more competent sellers or by inducing every type of seller to behave well. To try to isolate what a high ability seller might be, we clearly have the complication that, for the same amount of past cost (and time) renegotiations, two sellers that are identical in terms of ability might appear from the data as very different in their performance due to either the different complexity of the contracts they faced or the different quality of the buyers they interacted with. To try to isolate these latter two forces and obtain a measure of seller's ability as close as possible to be a proxy for the seller's "type", we exploit the richness of the data and proceed as follows. First, we assign each contract in the dataset to a bin where bins are constructed on the basis of three dimensions: i) the nature of the service (to capture the different complexity of the different types of services), ii) the quartile of the contract value distribution it belongs to and iii) whether the buyer's Q28 value is above or below the median (to capture the buyers' quality). For each bin, we calculate the average CP and TP. Then, separately for each supplier, we calculate for each contract that he won whether his CP and TP was above the average of that contract's bin. Finally, we measure the *seller's cost ability* as the share of contract in which the supplier outperformed the (bin-specific) average CP, relative to the total number of contracts he won. The *seller's time ability* is defined in the same way using TP.

Albeit our proxy is an admittedly imperfect measure of seller's type, its usage within the linear regression models for CP and TP reveals a few interesting facts. Table 9 reports these results. The first two columns regards models where CP is the dependent variable, while TP is the dependent variable in the last two columns. Models (1) and (3) include the seller's ability as defined above, respectively for cost and time. Models (2) and (4) use, instead, include sellers fixed effects. To perform this analysis we must restrict attention to the subset of sellers winning at least two contracts. As Figure 5 illustrates, the majority of sellers in our sample wins at least two contracts and a non-negligible share of them wins more than 1,000 contracts. Regarding the regression results in Table 9, a first result of notice is the

29

marked increase in the models' explained variance once sellers' features are included. In this respect, our proxy for seller's ability is preferable to the fixed effect in terms of an even higher R^2 for both CP and TP. Furthermore, as expected both CP and TP are positively and significantly associated with performance. They also have a large magnitude. All these features are indicative that sellers with a greater ability are indeed better contractors in terms of what performance they will deliver. This is an important indication that selecting a good seller is important in determining the heterogeneity across buyers. Nevertheless, the fact that the explained variance remains fairly low in absolute terms indicates there remains scope also for buyers' heterogeneity to be driven by how different buyers guide sellers with the same ability.

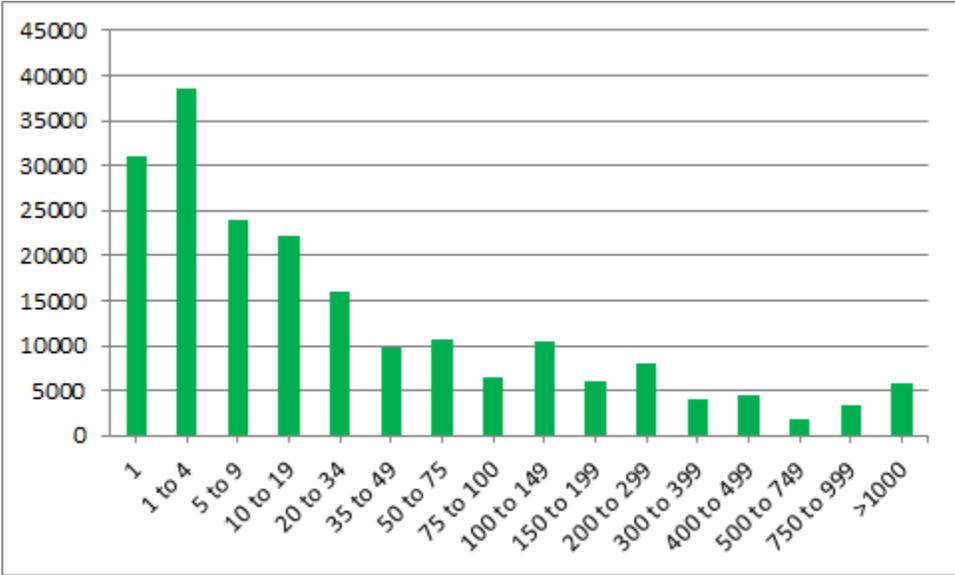Figure 5: Number of Contracts Won by Each Seller

Table 9: PERFORMANCE AND SELLER'S ABILITY

	(1)	(2)	(3)	(4)
Cost Plus	-0.062***	-0.100***	-0.0098	-0.055***
	(0.0080)	(0.014)	(0.0085)	(0.013)
Log-Project Value	-0.0033***	-0.015***	-0.0050***	-0.033***
	(0.00066)	(0.0016)	(0.0011)	(0.0022)
Cum. Category Experience	0.57***	1.05***	0.42***	1.03***
	(0.15)	(0.30)	(0.15)	(0.34)
Seller's Cost Ability	0.49***			
	(0.0069)			
Seller's Time Ability			0.65***	
			(0.0060)	
N	202790	202790	202790	202790
R-squared	0.55	0.48	0.58	0.44
Year FEs	Yes	Yes	Yes	Yes
Bureau FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Product FEs	Yes	Yes	Yes	Yes
Firm FEs	No	Yes	No	Yes

V   Conclusions

We have constructed new measures of agency and bureau procurement quality based on their past ability to effectively obtain the procured goods and service within the initially contracted costs and time limits. Coupled with survey data on agencies' perceived quality by employees, these performance-based quality measures were found to be highly heterogeneous across different agencies and time persistent. They were shown to be significantly related to better procurement outcomes.

Buyer's quality was found to affect performance via different channels, including a correct choice of the award procedure (e.g. auction vs negotiation) and the ability to select good quality contractors. The qualitative aspects that contribute the most on predicting procurement performance were found to be the perceived degree of cooperation within the unit, followed by the provision of appropriate incentives. This evidence indicates the importance of improving the quality of a purchasing unit by selecting highly skilled procurement managers, able to adopt management practices ensuring cooperation among different specialities (law, economics, engineering/management) and the motivation in the unit.

VI   References

- Bandiera Oriana, Prat Andrea and Tommaso Valletti (2009), Active and Passive Waste in Government Spending: Evidence from a Policy Experiment. American Economic Review, 99(4): 1278-1308.

- Bajari Patrick L., Robert S. McMillan, and Steve Tadelis, (2009), Auctions versus negotiations in procurement: An empirical analysis. Journal of Law, Economics and Organization, 25(2): 372_399.

- Bloom Nicholas, Lemos Renata, Sadun Raffaella and John Van Reenen (2014), Does Management Matter in Schools? NBER Working Paper No. 20667.

- Bloom Nicholas, Propper Carol, Seiler Stephan and John Van Reenen (2015), The Im-

pact of Competition on Management Quality: Evidence from Public Hospitals, Review of Economic Studies, 82 (2): 457-489.

- Bloom Nicholas and John Van Reenen (2007), Measuring and Explaining Management Practices Across Firms and Countries, Quarterly Journal of Economics, 122 (4): 1351-1408.

- Boyatzis, Richard (1982), The Competent Manager: A model for effective performance. John Wiley & Sons.

- Clinton, J.D., D.E. Lewis, M. Richardson (2016). "Characterizing the Ideology of Federal Agencies: An Approach to Measuring Difficult-to-Observe Organizational Characteristics," mimeo.

- Coviello, Decio, and Mario Mariniello (2014). Publicity Requirements in Public Procurement: Evidence from a regression discontinuity design. Journal of Public Economics 109, 76-100.

- Coviello, Decio, Guglielmo Andrea and Giancarlo Spagnolo (2016), The Effect of Discretion on Procurement Performance, CEP Discussion Paper, No 1427. London School of Economics and Political Science.

- Decarolis, Francesco (2014), Awarding Price, Contract Performance, and Bids Screening: Evidence from Procurement Auctions, American Economic Journal: Applied Economics, Vol. 6, Iss. 1, January 2014, pp. 108-132.

- Decarolis, Francesco, Pacini, Riccardo and Giancalo Spagnolo (2016), Contractors' Past Performance and Procurement Outcomes: A Firm-level Experiment," SIEPR Discussion Paper 16-036.

- Duflo, Esther and Abhijit, Banerjee (2000) Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry Quarterly Journal of Economics, 115 (3): 989-1017.

- Goldman, Eitan, Jörg Rocholl, and Jongil So (2013), Politically Cconnected Boards of Directors and the Allocation of Procurement Contracts." Review of Finance 17 (5): 1617-1648.

- Guccio, Calogero, Giacomo Pignataro, and Ilde Rizzo (2014). Evaluating the efficiency of public procurement contracts for cultural heritage conservation works in Italy. Journal of Cultural Economics, 38(1):43_70, 2014.

- Lewis, Gregory, and Patrick Bajari (2011) Procurement with Time Incentives: Theory and Evidence. Quarterly Journal of Economics, 126 (3): 1173-1211.

- Lewis, Gregory, and Patrick Bajari (2014), Moral hazard, Incentive Contracts, and Risk: Evidence from Procurement." The Review of Economic Studies, 81(3): 1201-1228.

- Liebman, Jeffrey B., and Neale MAHONEY (2013), Do Expiring Budgets Lead to Wasteful Year-end Spending." Evidence from Federal Procurement. NBER Working Paper No. 19481

- McClelland, D. C. (1973). Testing for competence rather than for intelligence. American Psychologist, 28, 1-14.

- Saussier Stephane and Jean Tirole (2015) Strengthening the Efficiency of Public Procurement http://cadmus.eui.eu/handle/1814/35680.

- Tassabehji, Rana, and Andrew Moorhouse (2008), The changing role of procurement: Developing professional effectiveness, Journal of Purchasing and Supply Management 14.1 (2008): 55-68.