Past Performance and Entry in Procurement: an Experimental Investigation*

Jeffrey V. Butler
University of California, Merced

Enrica Carbone
University of Campania "Luigi Vanvitelli"

Pierluigi Conzo
University of Turin
Collegio Carlo Alberto

Giancarlo Spagnolo†
Stockholm School of Economics-SITE
University of "Tor Vergata", EIEF & CEPR

This version: June 1, 2018

Abstract

There is widespread concern that incentive mechanisms based on past performance may hinder entry in procurement markets. We report results from a laboratory experiment assessing this concern. Within a simple dynamic procurement game where suppliers compete on price and quality we study how an incentive mechanism based on past performance affects outcomes and entry rates. Results indicate that some past performance based mechanisms indeed hinder entry, but when appropriately designed may significantly increase both entry and quality provision without increasing costs to the procurer.

JEL Codes: H57, L14, L15
Keywords: Bid subsidies, Entry, Past performance, Procurement, Quality, Supplier selection, Vendor rating.

*We are grateful to Gary Charness, Martin Dufwenberg, Dan Levine, Marco Pagano and Marco Pagnozzi for detailed comments and very useful discussions. We gratefully acknowledge research funding from the Swedish Competition Authority (Konkurrensverket).

†Address correspondence to: *Giancarlo Spagnolo, Stockholm School of Economics-SITE, Sveavägen 65, Box 6501, SE-11383, Stockholm; Fax: +46-8-31 81 86; E-mail: giancarlo.spagnolo(at)hhs.se.*

1 Introduction

Do selection criteria based on past performance deter entry in procurement markets? If buyers are allowed to formally consider suppliers' track records, does this necessarily hinder the ability of new firms, those with little or no performance history, to win contracts?

Understanding how procurement markets should be designed is a first order concern for firms and governments. Public procurement alone accounts for 15 to 20 percent of GDP in developed countries and is on the rise as continuing budget shortages cause governments to rely more and more on private providers. Private procurement is even larger and each well run firm has a system to evaluate and reward supplier past performance, typically with a preference in the allocation of new supply contracts. Given the economic importance of procurement markets, the lack of consensus among policy and law makers around the world on fundamental questions like the consequences and appropriateness of allowing current contracts to depend on past performance is surprising.

In the US, for example, the Federal Acquisition Regulation (FAR) *requires* government agencies to record contractors' past performance in a common database and to take it into account when awarding new contracts.[1] This policy was introduced by the Federal Procurement Streamlining Act (1994) with the intention of making public procurement less bureaucratic and more effective, closer to private procurement practices, but has recently drawn criticism. Several prominent US senators voiced their concern that the extensive use of past performance scores at the contractor selection stage could hinder the ability of new or small businesses to enter into and win competitions for public contracts. The debate led the Government Accountability Office to study dozens of procurement decisions across multiple government agencies in 2011. The resulting report, while inconclusive, contains some intriguing support for the senators' concern.[2] Despite extensive and costly policies aimed at fostering small business' access to US procurement markets,[3] the report " ... identified only one procurement in which offerors ... lacked relevant past performance."

On the other side of the pond, European regulators appear to have always been convinced that allowing the use of past performances indicators as criteria for selecting among contractors leads to manipulations in favor of local incumbents, hindering entry, cross-border procurement and common market integration—the main objective of the EU. For this reason EU Procurement Directives *prohibited* taking suppliers' track records into account when comparing their bids, with minor exceptions, until very recently. New directives open some scope for past-performance based selection.

[1] See `http://www.whitehouse.gov/omb/procurement_index_contract_perf`

[2] See Government Accountability Office (2011). The inquiry had a qualitative nature and did not reach clear conclusions in our reading.

[3] See e.g. Athey et al. (2013), and references therein.

The main reason US regulators and European public buyers push for the use of past-performance indicators in selecting among contractors is that they consider them essential to obtain good value for taxpayers' money. Direct and indirect litigation costs are high and court-enforced contracts are typically not sufficient to achieve satisfactory governance of all qualitative aspects of exchanges. Since procurement is rarely occasional, mechanisms that reward past performance, which are widely used in private procurement, can complement and improve substantially on what formal contracting can achieve (e.g. Kelman, 1990).[4]

Private buyers use past performance as a supplier selection criterion because they are concerned about the price and quality of the intermediate goods and services they buy, a crucial determinant of their competitiveness. Governments and regulators in charge of public procurement may also be interested in objectives other than the price/quality ratio of publicly purchased goods. For example, they are also usually concerned that the public procurement process is transparent and open in order to stimulate cross-border procurement and market integration—as in the case of the EU—as well as for obvious integrity/accountability reasons.[5] They may want to ensure that small businesses are not excluded from public procurement, a concern that in the US has led to large programs like the one started by the Small Business Act.[6] If the use of past-performance indicators in the selection of suppliers deters entry, however, then there may be a trade-off between the improvement in price/quality ratios buyers can secure using past performance indicators and the decrease in the likelihood that new, possibly more efficient or innovative suppliers will enter the market—as US senators fear for the US federal procurement market.

To shed light on this controversial issue, we implement in the lab a simple repeated procurement game with incentives implemented through a pre-announced and transparent bid subsidy linked to past quality provision. We then study the effects of these incentives on the entry decision of a more efficient supplier. Theoretically, it is well known, at least since Klein and Leffler (1981), that in markets where asymmetric information makes past performance important, new entrants must produce below cost to build up an advantageous track record. With imperfect financial markets, this may lead to entry deterrence of smaller, more financially constrained firms. However, quantitative studies on this possible effect are lacking, while policy makers

[4]This is why Steven Kelman introduced them in the big reform of US federal procurement he led when serving as Administrator of the Office of Federal Procurement Policy in the Office of Management and Budget during the Clinton administration, with the reform leading, among other things, to the Federal Procurement Streamlining Act (1994).

[5]This may not always be the case in practice. Djankov et al. (2002), in a cross section comparison of 85 counties, show that stricter regulation of entry is associated with higher levels of corruption and greater size of unofficial economic activity, suggesting entry is often regulated in the interest of regulators and not of the consumers. See also Bandiera et al. (2009) as well as Coviello and Gagliarducci (2016).

[6]This includes rules limiting bundling, the establishment of the Small Business Agency and "set aside" programs trying to stimulate small business entry in many procurement markets. Athey et al. (2013) provide empirical evidence that such "set-asides" typically foster entry by smaller firms at the cost of lower efficiency and revenues. They go on to show that a policy similar to the one we study, i.e., subsidizing small bidders, may foster entry without reducing efficiency or revenues. See also Marion (2007).

are increasingly wary of policy recommendations coming from complex theoretical models that disregard crucial institutional aspects and are sensitive to changes in simplifying assumptions needed to solve them. Some type of evidence is needed to convince policy makers. In this paper we provide more structured evidence than the qualitative assessments in the GAO report mentioned above. An ideal dataset to study these questions would contain observations of behavior in procurement markets that differ only in the role played by the supplier's past performance. For lack of such an exogenous variation in real-world procurement systems we conduct a tightly controlled laboratory experiment based on the simplest procurement game we could envisage that reflects the relevant environment and the issues at stake.[7]

We augment a three-period competitive procurement game among two suppliers with a decision by each period's winner on whether to provide costly quality in that period, and with the possibility of entry by a third, more efficient competitor in the final period, which represents the future. We then allow past performance to matter by adding a mechanism that generates a pre-announced advantage at the bidding stage for a supplier that provided high quality in the past. We implement this past performance advantage in the form of a transparent bid subsidy, a mechanism which Athey et al. (2013) single out as being particularly promising in real-world procurement auctions. Across treatments, we vary both the existence of an incentive mechanism and, when a mechanism is present, the relative size of the bid subsidy that a potential entrant—with no track record yet—may be given.

We use this framework, first and foremost, to ask whether past performance-based procurement can or must necessarily deter entry, our main research question. We then dig deeper to investigate precisely how the relative size of the entrant's advantage (bid subsidy) affects both the quality level delivered by sellers and the total costs paid by the buyer. This leads us to also evaluate the effects of these mechanisms on welfare functions with different weights for quality, price and entry, respectively.

One main novelty of our past performance mechanism is the provision of a bid subsidy to the potential entrant. This aspect of our design is meant to address a common misconception. It is often taken for granted in the policy debate that past performance mechanisms must be designed so that entrant firms with no history of production would start "without a bonus "—i.e., on an equal footing with an incumbent firm having the worst possible track record—which would obviously provide incumbent firms with an advantage that might deter new entrants.[8] However, in order to be effective in providing incentives, mechanisms must typically be based

[7]In this sense, our study belongs to the kind of experiments that Al Roth (1986) categorizes as "Whispering in the ears of princes." See e.g. Feldman and Ruffle (2015) and Abeler and Jäger (2015) for recent laboratory experiments that take a similar path to shed light on other important policy design questions. A field experiment would of course be ideal as an external validity test, but it is not easy to change procurement rules for parts of a country or region. Experiments with firms, on the other hand, can be arranged to clarify this question. See Decarolis et al. (2016) for a step in this direction.

[8]This assumes that a bonus of zero is the minimum possible score. The same argument applies after appropriate rescaling if the bonus scale includes negative values.

on clear rules that suppliers know and trust, which give commitment power to the buyer and allow these systems to be designed in a variety of ways. The mechanism may therefore award a positive rating to new entrants—e.g., the maximum possible rating, or the average rating in the market, putting entrants at less of a disadvantage—and ensure that this is taken into account by the scoring rule that selects the contractor, even if the contractor has never interacted with the buyer before.

As a preview of our results, our data suggest that there need not be a trade-off between past performance and entry in procurement. We do find that standard mechanisms—i.e., those that assign zero score to potential entrants—increase quality but reduce entry. In this sense, the concerns of the US senators are in principle justified. However, when the mechanism is designed so that it awards a positive score also to potential entrants, we find that both entry and quality significantly increase relative to the baseline treatment without an incentive mechanism. More surprisingly, we find that these increases in quality and entry do not raise the final price paid by the buyer—even incorporating all applicable bid subsidies. That is to say, these mechanisms are eliciting higher quality and more frequent entry at no additional cost to the buyer, a kind of "nirvana" result.[9]

Providing a modicum of evidence on the external validity of this phenomenon, an analogous result appears to be observed in a recent field study on the effects of the introduction of a past performance system for the suppliers of a large corporation (Decarolis et al., 2016). One potential explanation for this puzzling result is that the kind of multidimensional and intertemporal competition induced by a mechanism like this simply makes tacit collusion more difficult. The past performances mechanism makes players asymmetric after the first round, and tacit collusion being hindered by asymmetry is a recurrent finding in experimental oligopoly games in general (Mason et al., 1992; Fonseca and Norman, 2008), and in homogeneous Bertrand games in particular (Boone et al., 2012; Dugar and Mitra, 2016). In this sense, our results could be seen as consistent with these previous experimental and empirical findings. In our experiment, however, the asymmetry is endogenously generated by multidimensional intertemporal competition, not exogenously imposed. Also, Dufwenberg and Gneezy (2000, 2002) find that only by increasing the number of competitors to three or more or concealing losing bids in past games is the Bertrand equilibrium eventually reached in one-shot interactions. Therefore, prices consistently above the Bertrand solution in the first two rounds of our baseline treatment were somewhat expected. The increase in competition we observe when a past performance mechanism is introduced could then be due to the "increased complexity" of the competitive

[9]The result is puzzling because in a dynamic Bertrand-like environment, the prospects of a future advantage after winning the first contract and obtaining a bid preference should induce tougher competition with an incentive system in place, but only in the first stage (see e.g. Cabral and Villas Boas 2005). After that, the incumbent should recoup the investment, and overall (average) prices should in the end reflect the higher costs of higher quality provision.

environment, rather than to the asymmetry undermining tacit collusion. This is in line with the theoretical hypothesis of Gale and Sabourian (2005) that more complexity induces faster convergence to the competitive solution. Other more "behavioral" explanations for our "nirvana" results could also be envisaged. Unfortunately, our experiment was not designed to discriminate between these alternative explanations (we did not anticipate such a result), so clarifying the robustness and the origin of this puzzling effect can be seen as an interesting line for future work.

Our study is confined to a simple procurement game tested in a stylized laboratory setting. However, if confirmed by further empirical and experimental evidence, our results imply that the dual goals of providing incentives for quality provision and for increasing entry are not mutually exclusive—they are both achievable through an appropriately designed past performance mechanism. Moreover, since the reaction of prices to the presence of bid subsidies that we observe is weak, it seems that the increase in quality and entry may come at very little cost to buyers.

The remainder of the paper proceeds as follows. In the next section, we discuss our experimental design. In Section 3, we present the results from our experiment. Section 4 discusses how our results relate to the existing literature, while in Section 5 we provide concluding remarks.

2 Experimental design

In this section we first describe the dynamic procurement game common to all treatments, then the past performance mechanism which we vary across treatments. Next, we derive theoretical predictions and describe how the experiment was implemented.

2.1 The procurement game

To reproduce in the simplest setting possible the essential features of procurement we are interested in, we implement in the lab a procurement game consisting of a sequence of three stages of homogeneous-good price competition involving two Incumbent firms and one Entrant firm.[10]

Participants play between twelve and fifteen rounds of the three-stage procurement game.[11] Before each round, participants are randomly and anonymously divided into groups of three and then randomly assigned one of two roles: two participants in each three-person group are assigned the role of "Incumbent firm," while the third person in each group is assigned the role of

[10]Notice that each static stage is similar to the game studied in Dufwenberg and Gneezy (2000). We adopt their assumption that firms are fully informed because it simplifies the environment, allowing us and the subjects to better focus on the *per se* complex dynamic choices of price, quality provision and entry under different past performance regimes.

[11]The number of rounds varies due to time constraints. Our participants had little, if any, prior experience with experiments. Game play therefore proceeded relatively slowly. Each session featured as many rounds as we could feasibly implement given a two-hour pre-scheduled time constraint.

"Entrant firm."[12] We implement multiple rounds in order to allow participants the opportunity to learn how to play optimally. However, they are instructed that at the end of the experiment only one round will be randomly chosen to count towards their experimental earnings. This is a standard practice in experimental economics and serves to provide proper incentives in each round by, for example, ameliorating across-round hedging motives.

Within each round and each group of three, play proceeds as follows. In each of the three stages comprising the sequence, Incumbent firms submit a price bid. Bids are discrete within the set $\{0.00, \ldots, 4.50\}$ and submitted simultaneously, with the lowest bid winning.[13] Ties are broken by selecting a winner uniformly randomly from among the firms submitting the lowest bid. The winning firm must produce the good, and after the bidding it can choose whether to produce a high quality good at a relatively high cost c_H, or to produce a low quality good at a lower cost $c_L < c_H$.

Losing firms produce nothing and consequently earn nothing. At the end of each stage, each firm learns the bids of the other firms in their own groups and the quality production decision of the winning firm in their group.[14] Before the third stage begins, Incumbent firms are informed of the entry decision of the Entrant firm in their group.

While Incumbent firms participate in all three stages of a round, Entrant firms may only participate in the third stage, observing all bids and outcomes of the other two firms in their group. For each of the first two stages, Entrant firms earn a fixed outside wage of w.[15] Before the third stage of the round begins, Entrant firms decide whether to participate in this last stage—and forgo their outside wage, w, for this stage—or to stay out of even this last stage and earn w.[16] Incumbent firms never earn an outside wage in any period. The timing of the game is summarized graphically in Figure 1.

The other way Entrant firms differ from Incumbent firms is in their cost advantage: we give the Entrant firm a cost advantage in the production of low quality goods to create an efficiency justification for buyers to prefer entry, so $c_L^e < c_L$ for Entrant firms.[17]

We choose a three-stage procurement game to make the experiment as simple as possible

[12] We use the terms "Incumbent" and "Entrant" here for clarity of exposition. Neutral language was used in the experiment. Specifically, roles were referred to as "Firm A," "Firm B" and "Firm C," with the first two being incumbents and the latter the entrant.

[13] We capped allowable bids at 4.50 euros as a precaution against the unlikely possibility of firms colluding on very high bids. This maximum was set to be substantially higher than would be expected in any equilibrium of the game. The precaution turned out to be unnecessary, as even though setting an explicit upper bound on bids in all likelihood enhanced the opportunity for collusion by creating a focal point, successful collusion on bids of 4.50 euros was essentially non-existent in the data.

[14] Participants are not informed about choices in groups other than their own.

[15] We do this simply to provide the participant assigned the role of the Entrant with experimental earnings comparable to those we may expect Incumbents to earn. Since it is a fixed wage it should have no impact on the strategic environment.

[16] That is to say, either an Entrant firm's earnings will be $3w$ for the round if it stays out of the last auction, or $2w$ plus its earnings from the third stage, if any, if it enters.

[17] We limit the price advantage to low quality since an Entrant firm will never optimally choose to produce high quality—the round ends after the Entrant firm makes its production decision.

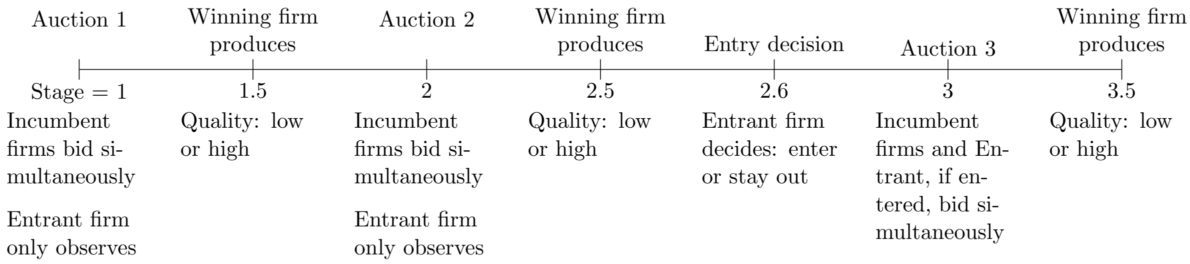

Figure 1: Timeline of Dynamic Procurement Auction Game

while still allowing for the features in which we are primarily interested. In particular, three is the minimum number of competitive stages allowing for investing in high quality (Stage 1) and potentially reaping the gains from such investment (Stage 2) before a new, more efficient firm has the option of entering the market (Stage 3).

2.2 Treatments

The experiment consists of four different treatments. The basic structure just outlined is common to all four treatments. What we vary across treatments is the presence of a past performance mechanism: three of the treatments involve a formal past performance mechanism while in the fourth treatment, "Baseline," no formal incentive mechanism is implemented. Across the three treatments involving a formal incentive mechanism we vary the incentive ascribed to Entrant firms.

2.2.1 Past performance incentive mechanism

While producing high quality is relatively costly, it may yield benefits. The essential characteristic of a past performance benefit is that, holding current stage bids constant, it puts a firm with good past performance at an advantage relative to competitors with a poor past performance. We implement the incentive in a simple and transparent way: by instituting a direct bid multiplier analogous to the one studied in Athey et al. (2013) that applies for only one subsequent stage. Specifically, for $t \in \{2, 3\}$, if the winning Incumbent firm in stage $t-1$ delivered the high quality good and also wins in stage t, this firm is paid a multiple B of its stage t bid by the procurer. An Incumbent firm with good past performance that wins the current stage with bid b is then paid Bb. Because this bid subsidy allows a firm with good past performance to earn a positive profit with a lower bid than an identical firm with poor past performance, it confers an advantage in terms of the ability of profitably winning the current stage.[18]

To analyze the market implications of assigning a positive score to a potential entrant, we

[18]Note that the bid advantage lasts only for one stage, i.e. the length of the buyer's memory of seller's past performance is just one stage.

also assign a bid multiplier, $\beta \geq 1$ to the Entrant firm. We vary the magnitude of β across treatments, considering three main cases: i) a "High Bonus" (**HB**) treatment where the bid multiplier for the Entrant is equal to the maximum possible for an Incumbent firm, i.e. $\beta = B$; ii) a "Low Bonus" (**LB**) treatment where the Entrant firm's bid multiplier is the minimum possible for the Incumbent, i.e. $\beta = 1$, no multiplier; and finally iii) a "Middle Bonus" (**MB**) treatment in which the Entrant firm's bid multiplier is between the maximum and the minimum possible for Incumbent firms, i.e., $\beta = \frac{(B+1)}{2}$. Treatment **HB** is analogous to supplier qualification and quality assurance systems common in the private sector where all qualified suppliers start with a fixed maximum number of points, lose points for bad performance, and may regain them through good performance but only up to the initial, maximum level. Points-based driver's license incentive systems are also designed this way in many countries. The **LB** treatment corresponds to a more standard past performance system where new firms without track records start out with minimal score. The remaining treatment, **MB**, represents a compromise between these two extremes, i.e. $\beta = \frac{(B+1)}{2}$, and corresponds to the case where Entrant firms enjoy the average market past performance score in a market where one of the two Incumbent firms has a good past performance. These rules are common knowledge among all players.

To choose specific parameter values, we balance two concerns: expository simplicity to facilitate experimental participants' understanding of the game against our desire for variation in equilibrium predictions across treatments. This balance of concerns suggested setting $B = 2$ and maintaining the same simple cost structure across all treatments: $c_L = 1.5$, $c_H = 2$ and $c_L^e = c_L - k$, where the Entrant firm's cost advantage $k = 1.375$. In all treatments the Entrant firm's outside wage, w, is 1. Our Baseline case without an incentive mechanism can be seen as implementing the parameters $B = \beta = 1$.

2.3 Equilibrium predictions

For equilibrium predictions we use the solution concept of subgame perfect Nash equilibrium (SPNE) with the additional restriction that firms never submit a sequence of bids so low as to guarantee themselves a negative profit from the remaining stages even if they win all remaining auctions. We view this as a reasonable restriction that rules out the type of non-credible threats that SPNE would rule out in a purely sequential-moves setting. It rules out, e.g., an Incumbent firm without a bonus bidding below c_L in the Stage 3 auction. Notice that this strategy would also be ruled out by a variety of formal equilibrium refinements such as sequential equilibrium or trembling hand perfection.[19] In the analysis that follows, denote Incumbent i's bid in auction k by b_i^k and the Entrant firm's bid in auction 3 by b_E. Note that low quality is always produced in Stage 3.5, since there is no monetary benefit to producing high quality yet there is a strict

[19] To see this notice that against a purely mixed strategy in which another firm places all possible bids with strictly positive probability, bidding below one's own cost in the Stage 3 auction can never be a best response.

increase in costs associated with doing so.

2.3.1 Baseline

In the Baseline treatment where no incentive mechanism is present it is easy to see that since high quality yields no advantage to any firm at any stage it cannot be produced in any equilibrium. This implies that Incumbent firms never bid below their cost of $c_L = 1.50$ (by assumption) in the Stage 3 auction, so that an Entrant firm who enters can always win with a bid of $b_E = 1.49$ against bids of $b_1 = b_2 = 1.50$ and obtain a profit of $1.49 - c_L^e = 1.49 - 0.125 = 1.374$.[20] Since the profit from staying out is $w = 1 < 1.374$, the Entrant firm always enters in Stage 2.6 and wins with $b_E \in \{1.49, 1.50\}$ in Stage 3.

At Stage 2.5, we have already established that low quality is produced so that, working backwards, there are two possible equilibria of the Stage 2 auction: i) both Incumbent firms bid at cost, i.e., $b_1 = b_2 = c_L = 1.50$; or ii) both firms bid at cost plus 0.01, i.e., $b_1 = b_2 = 1.51$. Because low quality is always produced, equilibrium expected profits from Stage 2 onward are either 0, in the former case, or 0.005 in the latter case. Working backwards once more, since low quality is always produced in Stage 1.5, bids in the Stage 1 auction are again either $b_1 = b_2 = c_L = 1.50$ or $b_1 = b_2 = c_L = 1.51$

Summing up, in the Baseline treatment equilibrium predictions are that: i) low quality is always produced; ii) entry always occurs; and iii) bids are always essentially $c_L = 1.50$.

2.3.2 HB treatment

In treatments with a formal incentive mechanism, equilibrium predictions are more complicated. Start with the HB treatment ($B = 2, \beta = 2$). In the Stage 3 auction, the lowest any Incumbent firm can bid in any subgame is 0.75. A bid below this would guarantee a negative profit even if the Incumbent has a past performance bonus. Since $\beta = 2$, an Entrant firm bidding $b_E = 0.74$ would earn a sure profit of $2 \times 0.74 - c_L^e = 1.355 > 1 = w$. Consequently, at Stage 2.6 Entry always occurs in equilibrium. Because there is no advantage to producing high quality at stage 2.5 but there *is* a strictly positive additional cost, low quality is always produced at Stage 2.5.

If neither incumbent enters Stage 2 with a bonus, equilibrium Stage 2 auction bids are the same as in the Baseline treatment: $b_1 = b_2 = 1.50$ or $b_1 = b_2 = 1.51$ with an expected total subsequent profit of either 0 or 0.005. If, however, one Incumbent firm has a past performance bonus – without loss of generality assume this is Incumbent firm 1 – this firm wins the Stage 2 auction in equilibrium with a bid of $b_1 \in \{1.49, 1.50\}$.[21] Total subsequent equilbrium profits for Incumbent firm 1 are either $2 \times 1.49 - c_L = 2.98 - 1.50 = 1.48$ or $2 \times 1.50 - c_L = 3.00 - 1.50 = 1.50$.

[20] A bid of $b_E = 1.50$ against bids of $b_1 = b_2 = 1.51$ could also occur in equilibrium, yielding the Entrant firm a profit of 1.375.

[21] The other Incumbent firm cannot bid strictly below 1.50 by our restriction on equilibrium strategies since this would result in a profit of at most $1.49 - 1.50 + 0.005 = -0.01 + 0.005 < 0$ from the remaining stages of the game.

Working backwards, since the increment to expected profits from producing high quality at Stage 1.5 is greater than the associated 0.50 increase in costs, high quality is always produced at Stage 1.5. At Stage 1, total expected subsequent profits from winning are either $\text{bid}_i - c_H + 1.48 = \text{bid} - 0.52$ or $\text{bid}_i - c_H + 1.50 = \text{bid} - 0.50$. Since firms are symmetric at this stage, in equilibrium firms bid away expected future profits. This results in four possible (essentially equivalent) sets of equilibrium bids: $b_1 = b2 = 0.52$, $b_1 = b2 = 0.53$, $b_1 = b2 = 0.50$ or $b_1 = b2 = 0.51$.

Consequently, our predicted behavior in HB treatment is: i) extremely low bids ($b_i < c_L, i = 1, 2$) in Stage 1 followed by high quality production in Stage 1.5; ii) bids at cost (c_L) in Stage 2 followed by low quality production in Stage 2.5; iii) entry in Stage 2.6 followed by bids of around cost (c_L) in Stage 3. That is to say, only one type of equilibrium exists in the HB treatment which we call "**entry accommodation**:" the Incumbent firm that wins the Stage 1 auction produces high quality in Stage 1.5, exploits this advantage to win the Stage 2 auction as well, but then cashes in on this advantage and accommodates entry by producing low quality in Stage 2.5.

Before moving on, it should be clear that owing to the discrete price grid we implement there will often be sets of equilibrium bids which are essentially identical except for being within a penny of one another. In the remainder of this section, for expositional simplicity, we will ignore this type of equilibrium multiplicity. For example, if $b_1 = b_2 = b$ and $b_1 = b_2 = b + 0.01$ both constitute equilibrium bids, we will only discuss the former bids.

2.3.3 LB treatment

In the LB treatment the Entrant firm is on equal footing with an Incumbent firm that has no past performance bonus ($B = 2, \beta = 1$). At Stage 3, there are four subgames to consider: Incumbent firm 1 has or does not have a bonus and entry did or did not occur at Stage 2.6. If Incumbent firm 1 has a past performance bonus and entry occurs, on the equilibrium path the Entrant firm wins the auction with a bid of $b_E = 0.74$ against other firm bids of $b_1 = 0.75, b_2 = 1.50$. The Entrant firm's profit in this subgame is $0.75 - c_L^e = 0.625$. If, on the other hand, Incumbent firm 1 has a bonus and the Entrant firm stays out, the Entrant firm's profit is $w = 1$ and Incumbent firm 1 wins the auction with a bid of $b_1 = 1.50$ against $b_2 = 1.51$ yielding Incumbent firm 1 a profit of $2 \times 1.50 - 1.50 = 1.50$. If neither Incumbent firm has a bonus and entry occurs, the the Entrant firm wins the Stage 3 auction with a bid of $b_E = 1.50$ against bids of $b_1 = b_2 = 1.51$ and earns a profit of $1.50 - c_L^e = 1.375$, while deciding to stay out again earns the Entrant firm a profit of 1.

Working backwards, at Stage 2.6 the Entrant decides to enter whenever neither Incumbent has a bonus ($1.375 > 1$). On the other hand, whenever an Incumbent firm has a bonus, the Entrant firm stays out at Stage 2.6 ($0.625 < 1$). Consequently, in Stage 2.5 an Incumbent firm can deter entry on the equilibrium path by producing high quality. The Incumbent's subsequent total profits following high quality production are 1.50, while producing low quality always results in zero subsequent profits. Because the increment to an Incumbent firm's profit associated with producing high quality is greater than the 0.50 cent increase in costs, the winning incumbent firm always produces high quality in Stage 2.5.

In Stage 2, there are again two subgames: either i) Incumbent firm 1 has a bonus; or ii) neither Incumbent firm has a bonus. In either case, our restriction on equilibrium bids implies that a firm with no bonus satisfies bid $- c_H \geq -$Stage 3 profit = bid $- c_H \geq -1.50$. In case i), Incumbent firm 1 wins the auction with a bid of $b_1 = 0.50$ (against a bid of $b_2 = 0.51$). Subsequent total profits are $0.50 = 2 \times 0.50 - c_H + 1.50$. In case ii), where neither Incumbent firm has a bonus, symmetric Bertrand competition for Stage 3 profits of 1.50 yields identical equilibrium bids of $b_1 = b_2 = 0.50$. These bids ensure zero total subsequent profit: $0.50 - c_H + 1.50 = -1.50 + 1.50 = 0$.

When choosing high or low quality in Stage 1.5, the winning Incumbent firm is choosing between subsequent total profits of either 0.50 by producing high quality or 0 by producing low quality. Since the 0.50 increase in future profit from producing high quality only exactly offsets the 0.50 cost difference between producing high or low quality, the winning Incumbent firm is indifferent between his two actions in Stage 1.5. Therefore, either high quality production or low quality production is possible in equilibrium at Stage 1.5.

Starting from Stage 1 all equilibria therefore yield the same subsequent total profits for Incumbent firms – zero profit. Because Incumbent firms are symmetric in Stage 1, equilibrium bid pairs consist of either $b_1 = b_2 = c_H = 2.00$ (in the equilibrium where high quality is produced in Stage 1.5) or $b_1 = b_2 = c_L = 1.50$ (when low quality is produced in Stage 1.5).

Summarizing, in the LB treatment there are two possible equilibria. In the first equilibrium: Stage 1 bids are high (c_H) and high quality is produced in Stage 1.5; Stage 2 bids are low ($< c_L$) and high quality is produced in Stage 2.5; in Stage 2.6 entry never occurs; in Stage 3, bids are moderate (c_L). In the second equilibrium: Stage 1 bids are moderate (c_L) and low quality is produced in Stage 1.5; Stage 2 bids are low ($< c_L$) and high quality is produced in Stage 2.5; in Stage 2.6 entry never occurs; in Stage 3, bids are moderate (c_L).

In either equilibrium, the high quality produced in Stage 2.5 always deters entry. Because of this latter feature, we refer to both of these equibuilibria as examples of "**Entry Deterrence**" equilibria.

2.3.4 MB treatment

Finally, consider the MB treatment ($B = 2, \beta = \frac{2+1}{2} = 1.5$). In Stage 3, our restriction on equilibrium bids implies that the lowest an Incumbent firm can bid in the Stage 3 auction is 0.75, if it has a bonus, or 1.50 if it does not. There are four subgames leading into Stage 3 to consider: Incumbent firm 1 has a bonus or neither firm has a bonus and, for each of these cases, entry did or did not occur.

In all subgames where entry did not occur, the Entrant firm's profit is 1. In the case where neither Incumbent firm has a bonus and entry did occur, the Entrant firm wins the auction with a bid of $b_E = 1.50$ against bids of $b_1 = b_2 = 1.51$, yielding a profit for the Entrant firm of $1.5 \times 1.5 - c_L^E = 2.25 - 0.125 = 2.125$. In the case where Incumbent firm 1 has a bonus and entry occurred, the Entrant firm wins the auction with a bid of $b_E = 0.75$ against bids of $b_1 = 0.76$ and $b_2 = 1.51$, which yields a profit of $0.75 \times 1.5 - c_L^E = 1.125 - 0.125 = 1$ for the Entrant firm.

Consequently, when choosing between entering or staying out in Stage 2.6, if neither Incumbent firm has a bonus the Entrant firm always chooses to enter ($2.125 > 1$). On the other hand, when Incumbent firm 1 has a bonus the Entrant is indifferent between his two options: both entry and staying out are possible on the equilibrium path.

Working backward, when choosing which quality level to produce in Stage 2.5, the winning Incumbent firm's subsequent total profits depend on the equilibrium being played. Any equilibrium where the Entrant firm enters in Stage 2.6 leaves zero profit for an Incumbent firm irrespective of its bonus. For these subgames, we can immediately see what strategies must be played in equilibria for all previous stages: they coincide with the "**Entry Accommodation**" equilibrium described in conjunction with the HB treatment above. In contrast, in the equilibrium without entry the Incumbent firm must produce high quality in Stage 2.5. In this case, strategies in all previous stages resemble the "**Entry Deterrence**" equilibria described in connection with the LB treatment above.

Summing up, in the MB treatment we can expect to observe two types of equilibria. On the one hand, any of the entry deterrence equilibria described in conjunction the LB treatment may occur. On the other hand, all of the entry accommodation equilibria described in conjunction with the HB treatment are also possible. As a result, we predict a mix of the features of these two types of equilibria in our data: overall, Stage 1 and 2 bids should be moderate and high quality should sometimes be produced in Stages 1.5 and 2.5. Entry should occur sometimes, and be particularly likely when low quality is produced in Stage 2.5.

2.4 Implementation

All sessions of the experiment were conducted in the laboratory facilities at the Einaudi Institute for Economics and Finance in Rome, Italy, using the software *z-Tree* (Fischbacher, 2007). Twelve sessions were conducted involving a total of 243 participants. Average earnings in the experiment were approximately 12 euros, including a show-up fee and payment for a risk elicitation task conducted after all rounds of the game were completed but *before* participants knew which round would be chosen to determine their earnings.[22] Because participants did not

[22]We included the risk elicitation task in order to ensure participants of a reasonable amount of money, as our equilibrium predictions suggested they would make little money from the auctions—a common dilemma when implementing Bertrand competition games in the lab. This concern turned out to be warranted, as competi-

know about the risk elicitation task when playing the auction game, it should not have affected decisions there. Each session lasted about two hours. Information on all four treatments is summarized in Table 1.

Insert Table 1 about here.

3 Results

Our experimental outcomes of interest are the proportion of winning firms producing high quality, the cost to the buyer—which we call the "buyer's (total) transfer" to avoid confusion with the sellers' costs of producing—as well as the proportion of Entrant firms choosing to enter. We first consider each of these outcomes in isolation and then consider buyers' welfare, which may incorporate some or all of these outcomes simultaneously.

3.1 Quality provision

Let us first examine quality provision, since encouraging high-quality goods provision is a primary reason buyers might prefer to implement some form of incentive mechanism. In Table 2, we report the average proportion of winning firms providing high quality. As expected, we observe a remarkable increase in high- quality provision in the first two stages in all treatments which involve an incentive mechanism relative to the baseline treatment, which lacks such a mechanism. For example, in Stage 1 about 80% of winning firms provide high quality whenever there is an incentive mechanism, whereas in the baseline treatment only 18% of winning firms provide high quality—a 340 % increase in the likelihood of high quality provision.[23] Averaging across all three stages (Table 2, last column), high quality provision is consistently about four times more likely with an incentive mechanism than without one.[24]

Insert Table 2 about here.

tion indeed drove profits from the auctions alone down to around €1.30, on average. While it may seem low, participants' average earnings of €12 for two hours of their time is commensurate with their opportunity costs. For example, work-study positions pay €5.50 per hour at a private college near the EIEF in Rome, whose pay structure we are familiar with. The risk elicitation task involved a sequence of choices between a sure payment of €5 and a lottery involving a 50% chance of a low payoff (€2.50) and a 50% chance of a high payoff, which increased over the sequence from €7.50 to €17 in steps of €0.50. More risk averse individuals should switch from preferring the sure payment to the lottery later in the sequence, so we take this switch point as an index of each participant's risk aversion. If there were multiple switch points, we follow much of the literature using related mechanisms and only consider the first switch point. One choice in the sequence was randomly chosen to count, with uncertainty being resolved, if necessary, by flipping a coin.

[23] It is somewhat surprising that even in the Baseline treatment with no incentives, 18% of winning firms choose to provide costly high quality. A likely explanation is a framing or labeling effect of high quality producing "good reputation" even when there are no concomitant financial incentives. This could be an interesting point worthy of further investigation in future research. However, for now we only note that this framing effect is constant across our treatments so that it should not contaminate across-treatment comparisons.

[24] In the Appendix (Table A1), we report a battery of pairwise non-parametric Mann-Whitney tests confirming the statistical significance of many of the large differences observed in the raw numbers: in Stages 1 and 2, Mann-Whitney tests reveal that quality provision in the baseline treatment is significantly different from all other treatments; differences among the non-baseline treatments themselves are generally not significant.

Result 1: *The introduction of an incentive mechanism significantly increases supplied quality.*

More formally, in Table 3 we estimate probit models of the binary decision to provide high quality in each of the stages separately (columns 1-3). In column 4, we pool observations from all three stages and estimate a Tobit model, using as the dependent variable the proportion of the three stages in which the winning firm provided high quality. In each of these estimates we control for dynamic effects, such as learning, by including the round of the observation as a control.[25] In these and all subsequent model estimates, unless otherwise noted, we cluster standard errors by session to allow for arbitrary within-session correlation of behavior. Confirming appearances in the raw data, we find that high-quality provision is significantly higher in all of our incentive mechanism treatments relative to the baseline treatment (the excluded category).

Finally, notice that in all treatments except the baseline, quality provision declines precipitously from the second stage to the third. This suggests that participants generally understood the strategic incentives inherent in each three-stage sequence, as there is no incentive to produce high quality in Stage 3. At the same time, even in Stage 3 quality provision is significantly lower in the baseline treatment than in any other treatment. One plausible explanation that would provide a further unintended benefit of implementing an incentive mechanism is that participants acquired a "habit" of high-quality provision in the first two stages which carried over to the third stage. Other possible explanations include "framing" or symbolic effects generated by the incentive mechanism. In any event, the effect is relatively small in magnitude, so we do not focus on it here.

Insert Table 3 about here.

3.2 Entry

Having confirmed that introducing an incentive mechanism can substantially increase costly quality provision, we are now in a position to address the central question of our inquiry: is there necessarily a trade-off between past performance and entry?

In Table 4, we report the proportion of Entrant firms choosing to enter Stage 3. These raw data suggest that an incentive mechanism which assigns no bid subsidy to the Entrant firm (LB) may indeed hinder entry, as feared by US senators and EU regulators. At the same time, however, our data suggest that a *properly calibrated* incentive mechanism need not hinder entry. Indeed, in both treatments where the Entrant firm is not assigned the worst possible score—MB and HB—our incentive mechanism tends to *increase* entry, without reducing quality provision.[26]

[25] In the Appendix (Table B1), we allow for more flexible dynamic effects by introducing a full set of round dummies into our model estimates. Nothing changes either qualitatively or in terms of statistical significance.

[26] As before, a battery of pairwise non-parametric tests by treatment is reported in the Appendix (Table A2),

Insert Table 4 about here.

To get a more formal sense of the significance of the effect of our incentive mechanism on entry, in Table 5 we report marginal effects estimates from a probit model using, as the dependent variable, an indicator taking the value one if the Entrant firm decided to enter the Stage 3 auction. On the right-hand side, we include a set of treatment dummy variables with the baseline treatment as the excluded category. To account for dynamic patterns in a simple way, we control for the round of the observation.[27] We find that entry is significantly higher relative to the baseline treatment, both economically and statistically, whenever the Entrant firm is not assigned the worst possible score. In treatments MB and HB, the estimated marginal effect of an incentive mechanism is to increase entry by 8 to 10 percentage points. On the other hand, we also find that the decline in entry observed in the raw data when Entrant firms are assigned poor score (LB) is not statistically significant.

Insert Table 5 about here.

Result 2a: *The introduction of an incentive mechanism that assigns no score to an entrant reduces the frequency of entry, although the effect is not statistically significant.*

Result 2b: *The introduction of an appropriately designed incentive mechanism that assigns a positive score to an entrant significantly increases the frequency of entry relative to the benchmark treatment without incentives, and does not reduce quality provision.*

Result 2a shows that the concerns raised in the policy debate about the possibility that rewarding past performance may hinder entry of new suppliers are justified and are captured by our experimental set up. Result 2b is our first main result. It shows that the answer to our main research question is negative: there is *not* necessarily a trade-off between past performance/quality and entry. A well designed and calibrated incentive mechanism that rewards past performance with a higher chance of winning (e.g. with a bid preference) can achieve both higher quality *and* higher entry.

3.3 Buyer's transfer

Because our results suggest that the effect of past performance on entry depends on the relative level of the Entrant firm's bid subsidy, and quality provision is costly for the supplier, a natural question to ask at this point is whether the most desirable outcome of high quality coupled with high entry comes at a significant increase in costs to the buyer. To avoid confusion with firms'

supporting the notion that the introduction of an incentive mechanism can either significantly increase or decrease entry, depending on the relative score assigned to the Entrant firm.

[27] A more flexible specification for dynamic patterns, incorporating a full set of round dummy variables, can be found in the Appendix, Table B2. This more flexible specification does not yield substantially different estimates.

costs of production, in the discussion that follows we refer to the total amount the buyer pays to the winning seller, accounting for any relevant bid subsidy, as the "buyer's transfer."

In Table 6 we report average buyers' transfers by treatment and stage, as well as the average buyer's transfer across all three stages. Somewhat surprisingly (at least for us) our data suggest there is only a very mild effect of even large bid subsidies on buyers' transfers, even though supplied quality increases.

Buyers' transfers are generally lower in the first stage when there is a an incentive mechanism than when there is not, reflecting competition for the bid advantage that high quality production entails in the subsequent stage. Considering the average buyer's transfer across all three stages, there is a surprisingly mild effect of our incentive mechanism on buyers' transfers, even though costly quality provision increases dramatically. This suggests that sellers incorporate future advantages when constructing their Stage 1 bids, and that competition on bids increases with the introduction of the incentive mechanism, guaranteeing that prices for the buyer do not increase together with quality provision.

Insert Table 6 about here.

To confirm this appearance, in Table 7 we present OLS estimates of buyers' transfers across treatments and stages. As usual, we control here for dynamic effects in a simple manner and report estimates allowing for more flexible dynamic patterns in the Appendix (Table B3). In both specifications, we find that introducing an incentive mechanism significantly lowers buyers' transfers in Stage 1. In subsequent stages, buyers' transfers may be higher or lower depending on the size of the Entrant firm's bonus, but the effects are generally mild and statistically non-significant. Considering buyers' transfers averaged over all three stages (Column 4), introducing an incentive mechanism never has a significant effect. This latter finding strengthens the interpretation that profit opportunities provided by bid subsidies are fully competed away and that the increase in costly quality provision comes at no cost to the buyer.

Insert Table 7 about here.

Result 3: *The introduction of an incentive mechanism that increases quality and entry does not increase the transfer paid by the buyer.*

This is our second main result, and perhaps the most surprising one. The increase in costly quality provision and in the frequency of entry generated by the incentive mechanism come at no additional cost to the buyer. An increase in supplier competition, linked to either the (endogenous) asymmetry or to the increased complexity introduced by the incentives mechanism, appears to us a likely explanation.[28]

[28] As will be discussed in the literature review, several previous experiments showed that asymmetry may have

3.4 Buyer's preferences: theoretical and empirical welfare functions

As a final exercise before concluding, in this section we construct a welfare function for buyers and examine how buyer's welfare varies, both theoretically and empirically, over our treatments.[29] In particular, we suppose that the buyer derives utility from three additively separable components: buyer's transfer (negatively), quality and entry. We model this in a flexible manner by assuming buyer's welfare is a simple weighted average of these three components. We then compare the welfare generated by each of our treatments—both theoretically, using equilibrium predictions, and empirically, using the experimental data—in two cases: i) buyers place equal weight on entry, quality and buyer's transfer; and ii) buyers do not care directly about entry, but rather divide all weight equally between the remaining two components. One can think of the first case as representing the situation in the EU, where increasing cross-border entry *per se* is a main political objective; the latter case may be closer to the US, where entry is valued only insofar as it increases efficiency and value-for-money for the taxpayer.

The welfare function we consider is $W = \alpha D + \delta Q + \gamma Pr(E)$, where $\alpha + \delta + \gamma \leq 1$; $D = \frac{4.5 - \frac{\sum_{t=1}^{3} T_t}{3}}{4.5}$; and T_t is the transfer from buyer to seller (i.e., [winning bid]*[relevant bid multiplier]) in Stage t. Notice that since 4.5 is the maximum allowable bid in the experiment, D is a measure of the "discount" below the maximum possible price buyers could pay excluding bid subsidies. This serves as a convenient normalization of the buyer's transfer component of welfare on a $0-1$ scale. The other two components of the welfare function are straightforward: Q is the proportion of the three stages in which high quality is produced; and $Pr(E)$ is the probability that—or the proportion of observations in which—entry occurs in the third stage. Weights are also normalized so that $\delta = (1 - \alpha - \gamma)$.

Using the parameters chosen for the experiment, we calculate the buyer's theoretical welfare by computing the equilibrium values of D, Q and $Pr(E)$ for each treatment and then evaluating buyer's welfare in each treatment for the welfare function weights implied by the two cases mentioned above: case i) $\alpha = \gamma = \frac{1}{3}$; and case ii) $\alpha = \gamma = \frac{1}{2}$. We report the buyer's theoretical welfare levels in these two cases in Table 8.

Insert Table 8 about here.

In case i) where buyers care about entry, quality and transfers equally, we find the highest buyer welfare in the HB treatment (when $B = \beta = 2$, $W = 0.71$), where the theoretical equilibrium frequency of entry is largest. On the other hand, in case ii) where buyers do

pro-competitive effects in oligopolistic environments. The argument that increased complexity may foster competition is theoretical instead, and goes back to Gale and Sabourian (2005). A recent empirical study by Decarolis et al. (2016) appears to find an analogous "nirvana" result for the buyer – introducing an incentive mechanism for suppliers increases supplied quality but not the price paid by the buyer – providing some reassurance about the external validity of Result 3.

[29] We are grateful to Gary Charness for suggesting this last exercise.

not care about entry directly, but only about quality and transfers instead, buyer's welfare is maximized in the LB treatment (when $B = 2$ and $\beta = 1$, $W = 0.73$), where the possibility of entry constrains bids and increases quality. Importantly, in both cases we consider that having an incentive mechanism in place increases buyer's welfare. To determine whether a similar result holds in our data, we next consider the empirical analogue of our theoretical buyer's welfare function. We measure quality, Q, by the average proportion of winning firms providing high quality across all three stages. We measure entry probability, $Pr(E)$, as the average proportion of Entrant firms entering in Stage 3. As our measure of buyer's transfers, we calculate D according to the formula described above. Table 9 reports our empirical estimates of buyer's welfare.

Insert Table 9 about here.

As with theoretical welfare, for both sets of weights considered buyers can always achieve higher welfare with an incentive mechanism than without. In contrast to our theoretical analysis, however, buyer's welfare is always maximized in the MB treatment, where Entrant firms are given neither the highest nor lowest possible score. This difference is likely due to Entrant firms basing their entry decisions on the bid subsidy to a lesser extent than our simple theoretical framework predicts. For example, entering Stage 3 with probability of less than one when the Entrant firm's bid-subsidy is relatively high (HB), as we observe in the data, reduces the empirical welfare advantage of HB over MB when buyers care about entry directly.

Result 4: *Introducing an incentive mechanism increases buyer's welfare, whether or not the buyer cares directly about the likelihood of entry.*

3.5 Predictions and behaviour

In this subsection we compare the experimental results with our theoretical prediction. The most evident inconsistencies between predictions and actual behavior appear in bidding behavior and quality provision. First, strong price competition does not emerge as subjects overbid especially in Stage 1 (Stage 2) in the Baseline and HB (LB) treatments (see Table C1 in the Appendix). Second, quality provision is higher than expected in Stage 2.5 of the HB and MB treatments. In the LB treatment subjects seem to behave according to the first equilibrium discussed in Section 2.3.3 above, with large winning bids and high quality provision in Stages 1 and 1.5 respectively, though in stage 2.5 the provision of high quality is lower than expected (see also Tables C1 and C2 in the Appendix). The behavior observed in the MB treatment is consistent with the theoretical predictions for the entry accommodation equilibrium as described in Section 2.3.2, with participants bidding low prices in Stage 1 and exploiting it in Stage 2 – they accommodate entry by producing low quality in Stage 2.5. Even if in this latter Stage high quality was less frequent than in the equivalent Stage of the HB treatment (where entry accommodation is the predicted strategy), high quality production in the MB is still higher than expected. In summary, subjects on average bid less aggressively and are more likely to produce high quality than theory predicts. Possible explanations for these inconsistencies are a limited ability to apply backward induction – perhaps due to the complexity of the game – as well as risk aversion or self-image motives.

The first explanation can only partially account for the reported inconsistencies. Even if a small share of subjects produce high quality when this is not a profit-maximizing strategy (e.g. in Stage 3.5), most of them seem to understand the game incentives. For instance, entry is lower whenever entry-deterrence is the most commonly played strategy (e.g. in the LB treatment) and the winning bids get close to the low-quality production cost in the last stage of all the treatments.

Asymmetries in risk aversion may explain non-aggressive bidding in our context as more risk averse bidders may lower their profits to improve the chance of winning (Campo 2012). In order to test this hypothesis, we first regress the winning bids at each bidding stage on our measure of participants' risk aversion (this measure is explained in footnote 22), controlling for a set of treatment dummies and the round of the observation. Results are reported in Table C2 in the Appendix and show that risk aversion has the expected sign but is not significant, leading us to exclude risk aversion as a primary explanation for the discrepancies between theory and behavior.

Another potential explanation for prudent bidding may be framing. We framed our experiment as a procurement auction, which – as Seifert and Strecker (2003) argue – may lead participants to bid more defensively than in alternative frames such as sales auctions. They show that the "sales" vs. "procurement" auction frame affects deviations from the dominant strategy, with participants bidding aggressively in a sales context and defensively in a procurement context. In addition, as overbidding (i.e. bidding defensively —above the private cost) especially occurs in our experiment in Stages 1 and 2, subjects may perceive higher winning chances if they bid less aggressively.

We also check whether the probability of choosing high quality is affected by participants' risk aversion. Table C3 in the Appendix shows that risk aversion negatively predicts high-quality production in Stages 1.5 and 2.5, highlighting that less risk-averse subjects tend to invest more often in bidding low prices and to avoid cheating the buyer. Considering within-treatment patterns, less risk-averse participants seem to provide high-quality more often in the stages and treatments where they are less expected to do so, i.e. Stages 1.5 and 2.5 of the Baseline treatment and Stage 2.5 of the HB treatment (results available upon request). This may suggest that when the cost of deviating from equilibrium increases, the risk-averse subjects are those more likely to play Nash.

The effects of risk aversion are robust to the inclusion among the controls of the bid previously submitted by the winner (Table C4 in the Appendix). The previous bid has negative effects on the quality produced in Stages 2.5 and 3.5, highlighting that the subjects winning the auction with large bids are more likely to provide low-quality in order to maximize profits. This evidence, jointly with the negligible fraction of subjects producing high-quality in the last Stage, would also reduce the role of self-image motives in explaining why high-quality is produced when it is not profit-maximizing to do so.

4 Related literature

Our work focuses on past-performance based incentive mechanisms and transparent selection rules to which the buyer can (or must) commit, as is often the case for large corporations and government procurement.

Closer to our paper are the laboratory experiments by Brosig and Reiss (2007) and Brosig-Koch and Heinrich (2014). The former analyzes capacity-constrained suppliers' decisions to enter and bid in the various stages of a sequential procurement game. It finds that entry and bidding in the sequential procurement auctions are indeed affected by the opportunity cost of early bidding generated by the capacity constraints. It also finds that entry decisions and average bids systematically deviate from equilibrium predictions, and that giving subject additional information on winners and prices tends to reduce the extent of this deviation. The latter asks what would happen if transparency rules that impose open price competition were removed and more discretion was left to public buyers, so that information on past behavior could matter for supplier selection. It finds that when buyers have discretion to choose among sellers, the latter invest in providing high quality. In contrast, when buyers lack the discretion needed to reward past behavior, sellers provide instead low quality to reduce short-term costs. Consequently, in the absence of a strucctured mechanisms rewarding past performance (such as the one discussed in this paper) buyer discretion increases market efficiency, with the benefits accruing entirely to buyers. Our study complements these previous studies as it also deals with entry decisions and past performances in a sequential laboratory procurement game. Our setup, however, is rather different (no capacity constraints, new entrant at the last stage, and pre-announced bid subsidies for past quality provision), as are the main research questions we focus on.

Related to our work are also experimental studies of one-shot and finitely repeated homogeneous good Bertrand price competition, as they share several features with our procurement game. Dufwenberg and Gneezy (2000, 2002) study finite-price one-shot homogeneous Bertrand games, repeated to allow for learning but with random and anonymous re-matching between the repetitions, that can be alternatively interpreted as homogeneous Bertrand competition with discrete prices or as first-price sealed bid procurement auctions. They find that only by increasing the number of competitors to three or more is the Bertrand equilibrium reached after learning, and that disclosing the losing bids in previous rounds leads to higher prices than only disclosing the past winning bids. Even though our environment is more complex than a one-shot Bertrand game (three repetitions in each round, with possible entry in the third, and quality choices), the fact that we observe prices consistently above the Bertrand solution in our Baseline treatment—the most closely related to a one-shot Bertrand game—can be seen as consistent with their results, as the first stages of each round are similar to a Bertrand duopoly and information on losing bids is disclosed after each bidding stage. This suggests that with more competitors, or without disclosing past losing bids, in our Baseline treatment bids would possibly be lower. On the other hand, our finding that competition seemingly increases with the introduction of transparent past-performance based mechanisms, leading to higher quality but not higher prices, suggests that these mechanisms, by introducing complexity in terms of multidimensional and intertemporal competition, may have effects similar to an increase in the number of competitors or the non-disclosure of losing bids. Our result is also consistent with several previous experimental studies, starting with Mason et al. (1992) and including Fonseca and Norman (2008), Boone et al. (2012), Dugar and Mitra (2016), addressing the effect of asymmetry in static and finitely repeated oligopolies and finding that it typically increases competition. In particular, the fact that we find more entry in HB than in Baseline appears in line with Dugar and Mitra's (2016) finding that subjects play the asymmetric Bertrand solution more often the larger the cost asymmetry among players. In contrast to these studies, however, in our treatments with a past-performance based mechanism the asymmetry between incumbents is not exogenous, but endogenously generated by the decision of the supplier to invest, provide high quality and earn the bid subsidy for the following period.

Closely related to our study is the recent strand of literature that focuses on past performance in procurement. Examples here include Calzolari and Spagnolo (2009), Board (2011), and Albano et al. (2017). Overall, these papers suggest that when important dimensions of the exchange are not contractible and there are many competing suppliers, a dynamic incentive mechanism based on past performance must complement standard competitive auctions to obtain decent value for money. In contrast to our study, however, none of these papers directly focuses on the consequences of dynamic incentive mechanisms for the entry of new firms.

Finally, on the empirical side outside of a laboratory environment, we are aware of only two studies that shed light on issues closely related to our paper: Koning and Van de Meerendonk (2014) and Decarolis et al. (2016). The former study documents an improvement in service quality following an increase in the scoring-weight given to supplier, in the scoring rule of public procurements of work-to-welfare programs. The latter study documents a significant increase in quality following the announcement of the introduction of a past-performance based vendor rating system in a large utility company, not followed by a corresponding increase in price. Both of these empirical studies appear broadly consistent with our experimental results.

5 Concluding remarks

In this paper, we ask whether the use of indicators based on past performance always entails a trade-off between increased quality provision and reduced entry of new suppliers in procurement markets. This question is important to private and government procurement design, and is at the center of a transatlantic policy debate as current regulations in the US and Europe reflect differing answers. In the US, where past-performance based mechanisms are currently required in Federal procurement, the Senate recently expressed concerns that such past-performance based selection criteria could hinder small businesses' ability to enter and successfully compete for contracts. On the other hand, in Europe, where regulators explicitly prohibit the use of past-performance indicators as criteria for selecting contractors on the grounds that they discriminate against cross-border entrants, public buyers and their national representatives have had recent limited success in overturning the prohibition.

We have investigated this question experimentally, augmenting a simple repeated procurement game with quality choices and potential entry by a more efficient supplier. Our treatments differed in the presence and design of a past-performance based mechanism that rewards high quality provision in a transparent way, i.e. with a pre-announced rule that assigns a bid subsidy for the subsequent procurement auction to suppliers that provide high quality in the current period.

Our results indicate that poorly calibrated mechanisms may indeed hinder entry, but that a trade-off between quality and entry is not necessary. To the contrary, we find that a well calibrated mechanism, in which new entrants with no history of past performance are awarded a moderate or high score – as is sometimes the case in private sector vendor rating systems or, for example, with point systems for drivers' licenses – actually *fosters* entry and, at the same time, delivers a substantial increase in quality.

Perhaps more surprisingly, we find that the increase in both costly quality provision and in entry made possible by properly calibrated incentive mechanisms may come at very little cost to the buyer. In our data, the total cost to buyers does not increase when the incentive mechanism is introduced, even in cases where costly quality provision and entry increase substantially. The introduction of bid subsidies for good past performance appears therefore to benefit the buyer by driving winning bids down enough to fully offset the potential increase in procurement costs due to bid subsidies and the costly quality provision they generate. This "nirvana" result for the buyer has already found some confirmation in the field, has obvious implication for public and private buyers, but is rather puzzling from a theoretical perspective. Additional experimental and field work appears therefore warranted to further test its robustness and to identify the driving forces behind it.

Summing up, our results imply that there need not be a trade-off between the use of appropriately designed past-performance based mechanisms and entry by new firms into a procurement market. A well calibrated mechanism based on past-performance may instead increase entry and quality provision simultaneously, without increasing the cost for the procurer. If confirmed in further studies, our findings suggest that the emphasis placed on past performance by the revised Federal Acquisition Regulation and by private procurer is fully justified, and that European regulators may have been imposing large deadweight losses on their citizens by forbidding the use of past-performance indicators as selection criteria in public procurement. Policy makers should probably stop quarrelling about whether a generic past-performance based mechanism should be introduced, and focus instead on *how* such a mechanism should be designed in different contexts.

References

[1] Abeler, Johannes, and Simon Jäger. 2015. "Complex tax incentives". *American Economic Journal: Economic Policy*, 7.3: 1-28.

[2] Albano, Gian Luigi, Bernardino Cesi, and Alberto Iozzi. 2017. "Public procurement with unverifiable quality: The case for discriminatory competitive procedures." *Journal of Public Economics*, 145: 14-26.

[3] Athey Susan, Coey Dominique, and Jonathan Levin. 2013. "Set-Asides and Subsidies in Auctions." *American Economic Journal: Microeconomics*, 5(1): 1-27.

[4] Bandiera, Oriana and Andrea Prat and Tommaso Valletti. 2009. "Active and passive waste in government spending: evidence from a policy experiment." *American Economic Review*, 99(4): 1278-1308.

[5] Bar-Isaac, Heski and Steve Tadelis. 2008. "Seller Reputation." *Foundations and Trends in Microeconomics*, 4: 273-351.

[6] Board, Simon. 2011. "Relational Contracts and the Value of Loyality," *American Economic Review*, 101(7): 3349–67.

[7] Boone, Jan, Marï¿œa Jose Larraï¿œn Aylwin, Wieland Mï¿œller, and Amrita Ray Chaudhuri. 2012. "Bertrand competition with asymmetric costs: experimental evidence." *Economics Letters*, 117: 134-137.

[8] Brosig, Jeannette and J. Philipp Reiss. 2007. "Entry decisions and bidding behavior in sequential first-price procurement auctions: An experimental study." *Games and Economic Behavior* 58(1): 50-74.

[9] Brosig-Koch, Jeanette and Timo Heinrich. 2014. "Reputation and Mechanism Choice in Procurement Auctions - An Experiment." *Production and Operations Management.*, 23(2): 210-220.

[10] Cabral, Luis and Miguel Villas-Boas. 2005. "Bertrand Supertraps." *Management Science*, 51(4): 599-613.

[11] Calzolari, Giacomo and Giancarlo Spagnolo. 2009. "Relational Contracts and Competitive Screening." *CEPR Discussion paper* No. 7434.

[12] Campo, Sandra. 2012. "Risk aversion and asymmetry in procurement auctions: Identification, estimation and application to construction procurements", *Journal of Econometrics*, 168 [1]: 96-107.

[13] Coviello, Decio and Stefano Gagliarducci. 2016. "Tenure in Office and Public Procurement." *American Economic Journal: Economic Policy*, forth.

[14] Decarolis, Francesco, Riccardo Pacini and Giancarlo Spagnolo. 2016. "Past Performance and Procurement Outcomes", (No. w22814). National Bureau of Economic Research.

[15] Dellarocas, Chrysanthos, Federico Dini and Giancarlo Spagnolo. 2006. "Designing Reputation (Feedback) Mechanisms" in *Handbook of Procurement*, Nicola Dimitri, Gustavo Piga and Giancarlo Spagnolo, eds. Cambridge University Press.

[16] Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2002. "The Regulation of Entry". *The Quarterly Journal of Economics*, 117(1): 1-37.

[17] Dufwenberg, Martin and Uri Gneezy. 2000. "Price Competition & Market Concentration: An Experimental Study", *International Journal of Industrial Organization*, 18(1): 7-22

[18] Dufwenberg, Martin and Uri Gneezy. 2002. "Information Disclosure in Auctions: an Experiment", *Journal of Economic Behavior and Organization*, 48: 431-444.

[19] Dugar, Subhasish, and Arnab Mitra. 2016. "Bertrand competition with asymmetric marginal costs" *Economic Inquiry*, 54(3): 1631-1647.

[20] European Commission. 2011a. Green Paper on the modernization of EU public procurement policy Towards a more efficient European Procurement Market. Available at `http://ec.europa.eu/internal_market/consultations/index_en.htm`

[21] European Commission. 2011b. Green Paper on the modernization of EU public procurement policy Towards a more efficient European Procurement Market. Synthesis of Replies. Available at `http://ec.europa.eu/internal_market/consultations/docs/2011/public_procurement/synthesis_document_en.pdf`

[22] Feldman, Naomi E., and Bradley J. Ruffle. 2015. "The impact of including, adding, and subtracting a tax on demand." *American Economic Journal: Economic Policy*, 7.1: 95-118.

[23] Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10(2): 171–178.

[24] Fonseca, Miguel A., and Hans-Theo Normann. 2008. "Mergers, asymmetries and collusion: experimental evidence." *The Economic Journal*, 118: 387-400.

[25] Government Accountability Office. 2011. (GAO-12-102R, Oct. 18, 2011). Prior Experience and Past Performance as Evaluation Criteria in the Award of Federal Construction Contracts. Available at `http://www.gao.gov/products/GAO-12-102R`

[26] Gale, Douglas and Hamid Sabourian. 2005. "Complexity and Competition." *Econometrica*, 73(s): 739-769.

[27] Heinrich, Timo. 2011. "Communication and reputation in procurement auctions – Some empirical evidence." *Economics Letters*, 114(2): 164–167.

[28] Kelman, Steven. 1990. Procurement and Public Management: The Fear of Discretion and the Quality of Government Performance: *American Enterprise Institute Press.*

[29] Koning, Pierre, and Arthur Van De Meerendonk. 2014. "The Impact of Scoring Weights on Price and Quality Outcomes: An Application to the Procurement of Welfare-to-Work Contracts," *European Economic Review*, 71: 1-14.

[30] Mailath, George and Larry Samuelson. 2006. Repeated Games and Reputation: Long-Term Relationships. *Oxford: Oxford University Press*

[31] Marion, Justin. 2007. "Are bid preferences benign? The effect of small business subsidies in highway procurement auctions," *Journal of Public Economics*, 91: 1591-1624.

[32] Mason, Charles F., Owen R. Phillips, and Clifford Nowell. 1992. "Duopoly behavior in asymmetric markets: an experimental evaluation." *The Review of Economics and Statistics*, 74: 662-670.

[33] Milgrom, Paul and John Roberts. 1982. "Predation, reputation and entry deterrence," *Journal of Economic Theory*, 27: 280-312.

[34] Roth, Alvin. 1986. "Laboratory Experimentation in Economics," *Economics and Philosophy*, 2: 245-273.

[35] Seifert, Stefan, and Stefan Strecker. 2003. "A Note on Over- and Underbidding in Vickrey Auctions - Do We Need a New Theory?", Available at `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.304.5116&rep=rep1&type=pdf`.