Procurement with specialized firms

Jan Boone and Christoph Schottmüller*

December 3, 2015

Abstract

We analyze optimal procurement mechanisms when firms are specialized. The procurement agency has incomplete information concerning the firms' cost functions and values high quality as well as low price. Lower type firms are cheaper (more expensive) than higher type firms when providing low (high) quality.

With specialized firms, distortion is limited and a mass of types earns zero profits. The optimal mechanism can be inefficient: types providing lower second best welfare win against types providing higher second best welfare. As standard scoring rule auctions cannot always implement the optimal mechanism, we introduce a new auction format implementing the optimal mechanism.

JEL: D82, H75, L51

keywords: procurement, specialized firms, deregulation, countervailing incentives

*Boone: Department of Economics, University of Tilburg Tilec, CentER and CEPR. Schottmüller: Department of Economics, University of Copenhagen; Tilec. We gratefully acknowledge financial support from the Dutch National Science Foundation (VICI 453.07.003). We thank audiences at the University of Copenhagen and the University of Edinburgh for comments. We are grateful to Cedric Argenton, Humberto Moreira, Eric van Damme, the editor David Martimort and three anonymous referees whose helpful comments and suggestions improved the paper considerably.

1. Introduction

Consider a government that needs to procure electric power. On the one hand, it is interested in a low price per kWh. On the other hand, the government is concerned about the effects of power generation on the environment; say, carbon emissions.[1] For the moment, fix the quantity that needs to be procured and let q ("quality") denote an inverse measure of carbon emissions, that is high q denotes low emissions. Then for a given technology, say a coal power plant, marginal abatement costs (MC) are increasing. But for a technology like gas which is a more expensive fuel, marginal abatement costs are lower than for coal at each q (see, for instance, Johnson et al., 2013). If high emissions (low q) are allowed, coal is cheaper than gas. However, if high environmental standards are enforced, gas is cheaper than coal. As standards are raised further, technologies like wind and solar energy come to the fore. Depending on the required level of q different technologies lead to lower costs.

Figure 1 is well known from micro economic textbooks (see, for instance, McAfee and Lewis, 2007) and has a similar flavour. Here, q on the horizontal axis refers to quantity produced by a firm. The figure shows the cost concepts average (total) costs (AC) and marginal costs (MC) for different firms and has two important characteristics. First, a firm has a (finite) scale that minimizes average costs –characterized by the MC curve intersecting the AC curve. The textbook argues that in the real world firms tend to face economies of scale at low output levels but dis-economies of scale at high output levels. Second, one firm is not unambiguously "better" than the other. Depending on the scale required, either firm can have lower costs per unit of output. Once we move beyond the micro textbook however, say to mechanism design, these cost functions tend

[1] "Green Public Procurement" has been on the political agenda for already quite some years; see, for example, Council of the European Union (2006).

to be replaced by simpler ones where firms can be unambiguously ranked in terms of efficiency.[2] The highest type is then more efficient than all other types irrespective of q, either as abatement or as scale.

This article analyzes a procurement setting with cost functions that have this property that the identity of the firm with lowest cost varies with q. We follow the literature on procurement and incentive regulation and assume that firms have private information with regard to their cost functions (see, for example Laffont and Tirole, 1987, 1993; Che, 1993). This private information is represented by a "type" which is assumed to be a scalar. Whereas this literature assumes that higher types have lower costs for all q, we allow firms to be specialized: one firm has lower cost at scale q_1 whereas another is more efficient at scale q_2. In particular, each firm has a scale where it is "best": no other firm can mimic this firm at this scale and earn the same or higher profits. In terms of the two types in figure 1, no type wants to mimick the other if each produces at q with minimal AC while earning zero rents.

[Figure 1 about here.]

In line with the carbon abatement example above, we use the terminology quality when referring to q in this article. Other examples of quality in a procurement setting are the following. Think of delivery time as (the sole) component of quality and suppose firms differ in production costs. In order to achieve a fast delivery time, firms might have to break already signed contracts with other buyers and this would require them to pay compensation to these other buyers. Consider the case where producers with lower production costs were able to get more/bigger contracts with other buyers in the past. If the penalties for breach of contract are sufficiently high, this implies that firms

[2] The literature on countervailing incentives, which is dicussed more thoroughly below, is the exception to this rule.

with high production costs might have the lowest total cost –consisting of production cost and penalty payments for breach of contract– for short delivery times, i.e. a high quality project. For lower qualities (longer delivery time), however, the firms with low production costs are cheaper as the long delivery time means that they do not have to breach other contracts.

While the "specialization" in the previous example resulted from differences in outside options, such specialization can also be observed as a result of prior investments in certain machines, technologies or employees. To illustrate, OMA is a leading international partnership in architecture.[3] Its projects include Parc des Expositions in Toulouse, Il Fondaco dei Tedeschi in Venice and the Commonwealth Institute in London. But this does not imply that OMA is a serious contender for designing your new house or a small school.

Often the distinction between firms specialized in high quality and firms specialized in low quality/low costs makes headlines in newly liberalized sectors. Examples of sectors that have been liberalized over the past years are postal services, air transport and railway. Some players compete with low prices and lower quality in, for instance, the following sense: only make deliveries twice a week (instead of 6 days a week), operate planes with reduced seat pitch and limited on board service as well as offering less connections and use slow (i.e. not high speed) trains. Incumbents in these sectors tend to be specialized in high quality at a high price because of pre-liberalization investments and their organizational form. If high quality incumbents compete against entrants specialized in low quality, which firm should win the procurement? How can a planner best play off one specialized firm against another?

Put differently, what is the optimal way to procure when firms are specialized?

[3] See `http://www.oma.eu/oma`.

Without specialization, the procurement literature cited above derives the following well known principles: only the highest possible type produces his first best quality ("no distortion at the top"), all other types' quality is distorted in the same direction (usually downward), all active types (except one) earn strictly positive rents and these rents, welfare and the probability of winning increase in type.

Taking specialization into account changes these principles in a way analogous to the countervailing incentives literature (see section 2).[4] In particular, there can be more than one firm with undistorted quality. We analyze a case where there are three such types. We have no distortion at the top (for two types) and no distortion at the bottom. In this case, the worst type is actually in the middle of the type space. Starting from this type, we branch out to better types by moving both to lower and higher types. Quality is distorted either upwards or downwards depending on whether the quality a type is specialized in is below or above his first best quality. There is an interval of types that produce their specialized quality: these types earn zero rents themselves and do not generate information rents for other types. In fact, it is possible to have all firms active –producing their specialized quality– without paying any (information) rent; although this is not optimal in general. Finally, rents are U-shaped.

We show that in the optimal mechanism the (second best) efficient firm does not necessarily win the procurement. This leads to new commitment problems for the procurer. Further, the optimal mechanism cannot be implemented by standard scoring rule auctions because such a scoring rule cannot discriminate between firms producing their specialized quality level earning zero rents. We propose a dual score auction that can implement the optimal mechanism in dominant strategies. With the optimal scoring

[4]The delivery time example above is typical for this literature which assumes that better firms, i.e. those with lower production costs, have also better outside options.

rule, initially the winning probability falls with the quality bid and then increases with q.

The set up of the article is as follows. We first give a review of the literature. In section 3, we present the model. Section 4 analyzes the case where first best welfare is monotonically increasing in type whereas section 5 deals with U-shaped first best welfare. In the latter case, we find a discrimination result, i.e. some types with lower second best welfare are preferred to types with higher second best welfare. Section 6 shows that it is not possible to implement the optimal mechanism with a scoring rule auction when specialization matters. We then propose an alternative way of implementation. Section 7 concludes. Proofs are relegated to the appendix. In the supplementary material to this article, we show how the optimal mechanism can be derived if some of the assumptions of section 3 do not hold.

2. Review of the literature

Our article is related to the literature on procurement, especially to those papers in which more than price matters, e.g. Laffont and Tirole (1987), Che (1993), Branco (1997) or Asker and Cantillon (2008). This literature shows how quality (or quantity) is distorted away from first best for rent extraction purposes. It also analyzes how simple auctions can implement the optimal mechanism. These papers assume that firms are not specialized, i.e. higher types have lower costs for all quality levels. This assumption seems to be too strong in many settings, e.g. the examples mentioned in the introduction.

Asker and Cantillon (2010) are an exception in the procurement literature. They analyze a model where firms differ in both marginal and fixed costs; each can be either high or low. Hence, there are four types. Which of the two mixed types has lower overall costs depends on the quality level, i.e. firms are partially specialized although this is not the main focus of their article. Our article shares some results with Asker and Cantillon (2010), e.g. quality can be upward and downward distorted. In contrast, we (i) analyze a situation of pure specialization, (ii) use more general cost functions, (iii) have a continuum of one-dimensional types and (iv) propose an auction that implements the optimal mechanism. This leads also to qualitatively new results, e.g. that the optimal mechanism is second best inefficient.

Our paper is also related to Ganuza and Pechlivanos (2000) who analyze a procurement model with horizontally differentiated firms. In their setting, it is costly for firms to explore their own costs for a given quality ("design"). These exploration costs are so high that the principal finds it optimal to set the desired quality level in a first stage. Then firms explore their costs for this quality level and bid in a discriminatory price auction in a second stage. They find that the optimal quality choice promotes heterogeneity between firms, i.e. the quality choice favors the more preferred firm even more but then the price auction discriminates against this firm. In our setting, firms know their cost functions from the start and the quality level to be provided is therefore not fixed ex ante. This leads to more conventional quality distortions that strengthen weak types.

Our article connects the literature on competitive procurement with the literature on countervailing incentives, see Lewis and Sappington (1989) and Maggi and Rodriguez-Clare (1995) for the seminal contributions and Jullien (2000) for the most general treatment. These papers analyze principal agent problems where higher types have lower marginal costs (of quality) but higher fixed costs. By assuming that firms are specialized, our model uses the same cost functions as the countervailing incentives literature. We contribute by allowing for several agents bidding for the contract whereas the countervailing incentive literature focuses on settings with one principal and one agent. That is, we put the countervailing incentives model in an auction context in the same sense in which Laffont and Tirole (1987) do this with the standard monopolistic screening model. Our result that the participation constraint is binding for a mass of types is typical for the countervailing incentives literature. Also the distortions of quality turn out to be the same as in the countervailing incentives model. Some of our results cannot occur in a principal agent model and do not occur when the standard monopolistic screening model is put into an auction context as in Laffont and Tirole (1987), e.g. second best inefficiency of the optimal provider choice and our results on scoring rule auctions. We also face new technical issues. For example, the standard proof that local incentive compatibility implies global incentive compatibility does not go through in our framework.

3. Model

We consider the case where a social planner procures a service of quality $q \in \mathbb{R}_+$. The gross value of this service is denoted by $S(q)$ with $S_q(q) > 0, S_{qq}(q) \leq 0$. The cost of production is denoted by the three times continuously differentiable cost function $c(q, \theta)$ where a firm's type θ is private information of the firm. There are n firms and each firm's type is drawn independently from a distribution F on $[\underline{\theta}, \bar{\theta}]$ which has a strictly positive and differentiable density f.

To explain how our set up differs from standard models, define $\theta^s(q)$ as the type producing quality q at minimal costs:

$$\theta^s(q) = \arg\min_{\theta\in[\underline{\theta},\bar{\theta}]} c(q,\theta) \tag{1}$$

If type $\theta^s(q)$ is asked to produce quality q, the planner can pay this type $c(q,\theta^s(q))$; this gives $\theta^s(q)$ zero rents and no other type can mimic $\theta^s(q)$ and earn rents. We say that $\theta^s(q)$ is specialized in quality q. In a standard model, we get $\theta^s(q) = \bar{\theta}$ for each $q > 0$ whereas $\theta^s(0) = [\underline{\theta},\bar{\theta}]$. The highest type has lowest cost to produce any quality $q > 0$. For any type $\theta < \bar{\theta}$, the only way to reduce all rents to zero is to exclude it from production (produce quality 0). We are interested in the case where there are interior solutions for $\theta^s(q)$. We define $k(\theta)$ as the quality level in which type θ is specialized:[5]

$$\theta = \arg\min_{\theta'\in[\underline{\theta},\bar{\theta}]} c(k(\theta),\theta') \tag{2}$$

In words, if type θ produces quality $k(\theta)$, the planner can give θ zero rents and no other type can mimic θ.

A tractable cost function that gives us this idea of specialization is given by

$$c(q,\theta) = h_1(q) + h_2(\theta) - \alpha q\theta \tag{3}$$

where $\alpha > 0$, h_1 is strictly increasing and convex in q and h_2 is convex in θ. Assumption 1 explains the features of this cost function and why we need them. We first go over some examples to illustrate the range of models captured by this cost function.

Example 1. *Assume that* $c(q,\theta) = \bar{\nu}(\theta) + (\bar{\theta} - \theta)q$, *where* $\theta \in [0,\bar{\theta}]$ *and the function* $\bar{\nu} \geq 0$ *is increasing in* θ. *This is the countervailing incentives set-up of Lewis and*

[5]Loosely speaking, we can think of k as the inverse of θ^s ($\theta^s(k(\theta)) \equiv \theta$); but because θ^s can be set-valued, we prefer not to pursue this formally.

Sappington (1989) and Maggi and Rodriguez-Clare (1995), where in our model higher θ implies lower marginal costs. We find $k(\theta) = \bar{\nu}'(\theta)$.

Maggi and Rodriguez-Clare (1995) interpret $\bar{\nu}(\theta)$ as the value of a foregone outside option and associate types with lower marginal costs with better outside options. This setup can model the delivery time example in the introduction: $\bar{\nu}(\theta)$ is then the cost of production whereas $(\bar{\theta} - \theta)q$ are the compensation payments to other buyers that have to be made in order to be able to deliver a high quality, i.e. a short delivery time.

With this cost function we find that $c(q, \theta)/q$ is decreasing in q. That is, there is no finite efficient scale where c/q is minimized as in figure 1. Similarly, returning to the example of carbon dioxide abatement in power plants, marginal abatement costs are increasing. Interpreting higher q as cleaner power, we need $c_{qq} > 0$ to model this. Different values of θ then refer to power companies with differing fractions of coal and gas power stations and solar thermal electric plants etc. (Johnson et al., 2013; Steen, 2015; DNV, 2014).[6] Example 1 does not allow for either of these cases, but equation (3) does.

Example 2. *Assume $c(q, \theta) = (q - \theta)^2 + q(1 - \theta/2)$ where θ is distributed uniformly on $[0, 1]$. Then we find that $k(\theta) = 4\theta/5$ and $\theta^s(q) = q/(1 - q/4)$.*

Figure 2 shows this cost function both as a function of q and as a function of θ. Figure 2a draws c as a function of q for three values of $\theta = 0, 0.5, 1.0$ and shows the quality level $k(\theta)$ at which each θ is most efficient (marked by red dots). These quality levels are given by $k(\theta) = 0, 0.4, 0.8$ resp. Figure 2b draws c as a function of θ for these three quality levels $q = 0, 0.4, 0.8$. Varying θ for these three values q, of course, shows that c is minimized for $\theta^s(q) = 0, 0.5$ and 1.0 resp.

[6] Note that the relevant information is the mix of the *free capacity* of the company. As contracts between the company and other customers are usually confidential, our assumption that the principal cannot observe the type seems reasonable in this example.

[Figure 2 about here.]

Finally, the blue and red cost curves in figure 1 are based on this cost function with $\theta = 0.5, 1.0$ *resp. Note that the value of* q *at which average cost* c/q *is minimized (which equals* θ *for the function here) does not necessarily coincide with* $k(\theta)$ *which is the value of* q *where type* θ *has lowest cost of all other types.*

The cost function in this example can be seen as a combination of horizontal and vertical differentiation. The former is captured by $(q-\theta)^2$. To see this, consider the case where firms are distributed on a Hotelling line, where their "address" $\theta \in [0,1]$ gives the quality level that they can produce without any "transportation"/adjustment costs. Producing $q \neq \theta$ involves quadratic transportation costs. The part $q(1-\theta/2)$ captures vertical differentiation: higher θ firms are better at producing each quality level q. These two parts together model that firms are specialized.

Example 3. *Assume* $c(q,\theta) = \frac{1}{2}q^2 - \theta q + \theta k$*. Thus,* $k(\theta) = k$*.*

This example, like example 1, reflects the idea that a firm with high fixed costs (θk) has lower marginal costs ($c_q = q - \theta$) of producing quality. For example, a firm that produces with a more capital intensive technology might have lower marginal costs for quality but higher fixed costs. The difference with the first example is that here the average cost curve is U shaped.

The following assumption specifies the properties of c in equation (3) and the properties of S and distribution function F that we use below.

Assumption 1. *We assume that*

- *the function* $c(q,\theta)$ *satisfies* $c_{qq} > 0, c_{q\theta} < 0, c_{\theta\theta} \geq 0, c_{\theta\theta q} = 0,\ c_{qq\theta} = 0,$

- *for $q \in I\!R_+$ it is the case that $S(q)$ is high enough compared to $c(q,\theta)$ so that the planner always wishes to procure (regardless of the type realization) and*
- *the function F satisfies the monotone hazard rate properties $\frac{d((1-F(\theta))/f(\theta))}{d\theta} < 0$ and $\frac{d(F(\theta)/f(\theta))}{d\theta} > 0$.*

These assumptions are standard in the literature. The first part says that c is convex in q, higher θ implies lower marginal costs c_q (the Spence-Mirrlees condition) and c is convex in θ. Finally, $c_{\theta\theta q} = 0$ and $c_{qq\theta} = 0$ (i.e. $c_{q\theta}$ constant) enable us to exclude stochastic contracts (e.g. a contract where player i's quality depends on other players' types) and will ensure that the standard monotonicity condition is satisfied. These assumptions imply that the cost function can be written as equation (3).

To ease the exposition, we assume that it is always socially desirable for the service to be supplied. A simple sufficient condition for this is $S(k(\theta)) - c(k(\theta),\theta) \geq 0$ for each $\theta \in [\underline{\theta},\bar{\theta}]$.[7] The third part is the monotone hazard rate (MHR) assumption. This assumption will allow us to use a first order approach by ensuring monotonicity of the resulting solution. Usually this assumption is only made "in one direction". However, in the literature on countervailing incentives it is standard to have MHR "in both directions", see for example Lewis and Sappington (1989), Maggi and Rodriguez-Clare (1995) or Jullien (2000).[8]

The assumption that $c_{\theta\theta} = h_2'' \geq 0$ implies that we can characterize $k(\theta)$ in (2) with the first order condition as

$$c_\theta(k(\theta),\theta) = 0. \tag{4}$$

[7]If we do not make this assumption, the virtual surplus (see below) can turn negative and for some realizations of θ the planner decides not to procure at all. Although straightforward to incorporate, we want to stress that with specialization production can be guaranteed for any realization of θ with zero rents.

[8]The normal, uniform and exponential distribution satisfy MHR. See Bagnoli and Bergstrom (2005) for a more complete overview.

Hence, we find that $k(\theta) = h_2'(\theta)/\alpha$.

The definition of k and assumption 1 lead to the following properties of k.

Lemma 1. *The function k is increasing and differentiable in θ. Furthermore,*

$$c_\theta(q,\theta) = h_2'(\theta) - \alpha q \begin{cases} > 0 & \textit{if } q < k(\theta) \\ < 0 & \textit{if } q > k(\theta). \end{cases}$$

For high values of q, a higher type θ produces q more cheaply. This is the usual assumption. We allow for the possibility where low values of q are actually more cheaply produced by lower types. To illustrate, high type firms may have invested in (human) capital that makes it actually relatively expensive to produce low quality. If the quality of the product is mainly determined by the qualification of the staff, these firms might have more expensive but also more qualified workers. Think of hiring OMA to build a small school.

If $k(\theta)$ is close to zero for all types, our model reduces to a standard model as analyzed in the earlier literature. In this sense, our model encompasses earlier procurement models. It is therefore not surprising that the solution of these earlier models shows up as a special case of our solution (see case 1 in proposition 1).

As c_θ can be both positive and negative, it is not clear how first best welfare varies with θ. First best quality is defined as

$$q^{fb}(\theta) = \arg\max_q S(q) - c(q,\theta) \tag{5}$$

which is uniquely defined as $S_{qq} \leq 0$ and $c_{qq} > 0$ by assumption 1. First best welfare is denoted by

$$W^{fb}(\theta) = S(q^{fb}(\theta)) - c(q^{fb}(\theta),\theta). \tag{6}$$

Hence, we find that

$$\frac{dW^{fb}(\theta)}{d\theta} = -c_\theta(q^{fb}(\theta), \theta). \tag{7}$$

In a standard model $dW^{fb}/d\theta > 0$ (as $c_\theta < 0$) and the highest type is best from a first best point of view. With c_θ changing sign in our set up, a lot of shapes are possible for W^{fb}. For concreteness, we choose one of them here. We assume that W^{fb} is quasiconvex in θ. That is, we focus on the case where best types are either firms that specialize in producing low quality cheaply or firms specializing in high quality.[9] This is a tension often seen in liberalized industries, as discussed in the introduction.

To get to quasiconvexity, we want $-c_\theta(q^{fb}(\theta), \theta) < 0$ and hence $q^{fb}(\theta) < k(\theta)$ for low θ and $-c_\theta(q^{fb}(\theta), \theta) > 0$ and hence $q^{fb}(\theta) > k(\theta)$ for high θ. A necessary and sufficient condition for this is that if q^{fb} intersects k, it intersects from below ($q^{fb}_\theta(\theta) > k_\theta(\theta) = h_2''(\theta)/\alpha$). This can be written as follows.

Assumption 2. *Assume that* $\frac{\alpha}{-S_{qq}(k(\theta))+h_1''(k(\theta))} > \frac{h_2''(\theta)}{\alpha}$.

With this assumption, we get the following result.

Lemma 2. *First best welfare $W^{fb}(\theta)$ is quasiconvex in θ. There is at most one type, denoted θ_w, at which $q^{fb}(\theta_w) = k(\theta_w)$. Furthermore, $q^{fb}_\theta(\theta_w) > k_\theta(\theta_w)$.*[10]

By the quasiconvexity of $W^{fb}(\theta)$ and $dW^{fb}(\theta_w)/d\theta = -c_\theta(k(\theta_w), \theta_w) = 0$, it follows that W^{fb} is minimized at θ_w. From a first best point of view, $\theta_w \in [\underline{\theta}, \bar{\theta}]$ (if it exists) is the worst type. Roughly speaking, types $\theta < \theta_w$ are better because they are cheaper and types $\theta > \theta_w$ are better as they produce higher quality.

[9] Combining example 1 with lemma 1, we focuses on the case where $k'(\theta) = \bar{v}''(\theta) \geq 0$. This is the cost function analyzed in section 4.1.1 of Maggi and Rodriguez-Clare (1995). The supplementary material analyzes another case where types in the middle can be "best" which is related to 4.1.2 in Maggi and Rodriguez-Clare (1995).

[10] With $k_\theta \geq 0$ (lemma 1) and $q^{fb}_\theta(\theta_w) > k_\theta(\theta_w)$ we rule out cost functions related to figures 2 and 4 in Maggi and Rodriguez-Clare (1995).

As first best welfare is quasiconvex in θ, we only need to consider two cases. Either first best welfare is monotone in θ or it is first decreasing and then increasing in θ. To ease the exposition, we will think of the highest type $\bar{\theta}$ as the best type, i.e. the type with the highest first best welfare. It should, however, be noted that analysis and results would not change if the lowest type was best (and by lemma 2 there are no other cases). The two cases that we focus on in this article are therefore:

Definition 1. *We consider the two cases:*

(WM) *first best welfare is monotone in* θ*:* $\frac{dW^{fb}(\theta)}{d\theta} > 0$ *for all* $\theta \in [\underline{\theta}, \bar{\theta}]$ *and*

(WNM) *a* θ_w *exists such that* $\frac{dW^{fb}(\theta)}{d\theta} < 0$ *for* $\theta \in [\underline{\theta}, \theta_w)$ *and* $\frac{dW^{fb}(\theta)}{d\theta} > 0$ *for* $\theta \in (\theta_w, \bar{\theta}]$*; further* $W^{fb}(\bar{\theta}) > W^{fb}(\underline{\theta})$*.*

WM is the case where $q^{fb} > k$ while q^{fb} intersects k in WNM. Following lemma 2, the only cases we neglect are therefore (i) $k > q^{fb}$ and W^{fb} is decreasing and (ii) q^{fb} intersects k but $W^{fb}(\underline{\theta}) \geq W^{fb}(\bar{\theta})$. These cases are, however, quite similar to WM and WNM respectively and could be solved with the same methods.

Returning to examples 2 and 3 above, with the assumption that $S(q) = q$, we can illustrate WM and WNM as follows. In example 2, we find that $q^{fb}(\theta) = 5\theta/4$. First best welfare is $W^{fb}(\theta) = \frac{9}{16}\theta^2$ which is increasing in $\theta \in [0, 1]$. In example 3, with $k \in (1 + \underline{\theta}, 1 + \bar{\theta})$, we find that $q^{fb}(\theta) = 1 + \theta$ and $dW^{fb}(\theta)/d\theta = 1 + \theta - k$. Hence, with $(k - 1) \in (\underline{\theta}, \bar{\theta})$ first best welfare increases for $\theta > k - 1$ and decreases for $\theta < k - 1$.

Now we are able to set up the mechanism design problem. The planner only needs one firm to supply the desired service or product. As $n \geq 2$ firms are able to supply, the planner needs to determine (i) which firm wins the procurement, (ii) what quality level this firm supplies and (iii) how much money is transferred to firms in return.

The planner offers a menu of choices for firms and each firm chooses the option that maximizes its expected profits. The profits of a type θ firm if it has to provide quality q with probability x and receives transfer t is $t - xc(q, \theta)$. The planner's objective is to maximize the expected value of $S(q)$ minus the expected transfer payments to all firms.

Following Myerson (1981), we use a direct revelation mechanism. That is, we design a menu of choices $(q^i(\Theta), x^i(\Theta), t^i(\Theta))_{i=1\ldots n}$ meaning that firm i receives transfer $t^i(\Theta)$ and has to provide the quality level $q^i(\Theta)$ with probability $x^i(\Theta)$ if the vector of types is $\Theta = (\theta^1, \ldots, \theta^n)$. The menu has to be designed such that it is incentive compatible (IC). That is, it is optimal for each firm i to truthfully reveal its type θ^i given that all other firms truthfully reveal their types.

If type θ misrepresented as $\hat{\theta}$, expected profits would be

$$\pi^i(\hat{\theta}, \theta) = \mathop{\mathbb{E}}_{\theta^{-i}} \left[t^i(\hat{\theta}, \theta^{-i}) - x^i(\hat{\theta}, \theta^{-i}) c(q^i(\hat{\theta}, \theta^{-i}), \theta) \right]. \tag{8}$$

With a slight abuse of notation we define the rent function $\pi^i(\theta)$ as

$$\pi^i(\theta) = \max_{\hat{\theta}} \pi^i(\hat{\theta}, \theta).$$

Using an envelope argument, incentive compatibility requires

$$\pi^i_\theta(\theta) = \mathop{\mathbb{E}}_{\theta^{-i}} \left[-x^i(\Theta) c_\theta(q^i(\Theta), \theta) \right] \tag{9}$$

for almost all types. This equation makes sure that the first order condition for truthful revelation is satisfied for type θ, i.e. it ensures that the derivative of (8) with respect to $\hat{\theta}$ is zero when evaluated at $\hat{\theta} = \theta$.

Example 1 (continued). *IC takes the form*

$$\pi_\theta(\theta) = \mathop{\mathbb{E}}_{\theta^{-i}}[-x^i(\Theta)(q(\Theta) - \bar{\nu}'(\theta))] \tag{10}$$

which corresponds to the IC constraint in Maggi and Rodriguez-Clare (1995) with two exceptions. First, as we have competition between firms, there is the probability that i wins the procurement ($x^i(\Theta)$) and potentially i's quality depends on its own type θ^i as well as the others' types θ^{-i}. Lemma 3 shows that q^i does, in fact, not depend on θ^{-i}. This allows us to follow the countervailing incentives literature in intuition and proofs.

If (9) holds, we say that *local IC* is satisfied. It is well known in the procurement literature (Laffont and Tirole, 1987) that local IC implies global IC under the usual regularity conditions. The standard proof does, however, not apply in our framework where firms are specialized: the constant sign condition $\partial^2\pi^i/\partial x^i\partial\theta > 0$ on which this proof relies is not satisfied in our framework as $c_\theta(q,\theta)$ can change sign. We will nevertheless first neglect non-local incentive constraints and use a first order approach; we refer to this as the *relaxed program.* After deriving the solution to this program, we verify that the non-local IC constraints do not bind under our assumptions. For the remainder of this section, we refer with "optimal mechanism" to the optimal mechanism of the relaxed program.

Finally, as firms can decide not to participate, a firm must have expected profits at least as good as its outside option. Because c_θ can switch sign, it is not clear for which type(s) this constraint is binding. Hence, we need to explicitly track the individual rationality constraint

$$\pi^i(\theta) \geq 0 \tag{11}$$

where we normalize firms' outside option to zero. Using the definition of π^i to substitute

π^i for t^i, the principal's relaxed program is therefore

$$\max_{(q^i,x^i,\pi^i)_{i=1\dots n}} \mathop{\mathbb{E}}_{\Theta} \left\{ \sum_{i=1}^{n} \left[x^i(\Theta) \left(S(q^i(\Theta)) - c(q^i(\Theta), \theta^i) \right) - \pi^i(\theta^i) \right] \right\} \tag{12}$$

subject to (9), (11) and the feasibility constraints $x^i(\Theta) \in [0,1]$ and $\sum_{i=1}^{n} x^i(\Theta) \leq 1$.

We proceed in three steps to solve this problem. First, we show that given any feasible allocation rule $(x^i(\Theta))_{i=1\dots n}$, the optimal q^i depends only on i's type θ^i. Second, we show that for any given allocation rule $(x^i(\Theta))_{i=1\dots n}$, the optimal $(q^i(\theta^i))_{i=1\dots n}$ is independent of the allocation rule. Third, we derive the optimal allocation rule and rents given the optimal $(q^i(\theta^i))_{i=1\dots n}$ derived in the second step. This three step procedure is the same as used in Laffont and Tirole (1987) where the second step consists of a problem that is similar to a countervailing incentives problem.

The following lemma establishes the first result: q^i depends only on θ^i and not on other firms' types. Hence, we can write $q^i(\theta^i)$ from here on. The intuition for this result is the following. If q^i depends on the types of the other firms, firm i is essentially facing a stochastic contract. The mechanism designer can gain if she assigns the expected quality (conditional on being contracted) to type θ^i instead of this stochastic scheme because the objective in (12) is concave in q^i. By the assumption $c_{qq\theta} = 0$, (9) is linear in q^i. Hence, assigning this expected quality will not affect (the slope of) the rent function.

Lemma 3. *The optimal quality schedule q^i does not depend on θ^{-i}.*

Using lemma 3, we can rewrite the objective in (12) as

$$\mathop{\mathbb{E}}_{\Theta}\left\{\sum_{i=1}^{n}\left[x^i(\Theta)\left(S(q^i(\theta^i))-c(q^i(\theta^i),\theta^i)\right)-\pi^i(\theta^i)\right]\right\}$$

$$=\sum_{i=1}^{n}\int_{\underline{\theta}}^{\bar{\theta}}\dots\int_{\underline{\theta}}^{\bar{\theta}}\left[x^i(\Theta)\left(S(q^i(\theta^i))-c(q^i(\theta^i),\theta^i)\right)-\pi^i(\theta^i)\right]$$

$$f(\theta^1)\dots f(\theta^{i-1})f(\theta^{i+1})\dots f(\theta^n)\,d\theta^1\dots d\theta^{i-1}\,d\theta^{i+1}\dots d\theta^n\,f(\theta^i)\,d\theta^i$$

$$=\sum_{i=1}^{n}\mathop{\mathbb{E}}_{\theta^i}\left[X^i(\theta^i)\left(S(q^i(\theta^i))-c(q^i(\theta^i),\theta^i)\right)-\pi^i(\theta^i)\right]$$

where we use Fubini's theorem and the fact that θ^i and θ^{-i} are independent for the first equality. For the second equality, we use the notation $X^i(\theta^i)=\mathbb{E}_{\theta^{-i}}\,x^i(\Theta)$, that is, $X^i(\theta^i)$ is the probability with which a type θ^i of firm i expects to be contracted. Note that in the last expression the term in square brackets depends only on firm i and not on types, qualities, selection probabilities or rents of other firms.

Now consider the problem of the second step where we take an arbitrary allocation rule $(x^i(\Theta))_{i=1\dots n}$ – and therefore also all $X^i(\theta^i)$ – as given. This second step maximization problem (over $(q^i(\theta^i),\pi^i(\theta^i))_{i=1\dots n}$) is then separable across firms as the objective is a sum in which the ith summand depends only on firm i. That is, the maximization problem over q^i,π^i in the second step for one particular firm i is

$$\max_{q^i,\pi^i}\int_{\underline{\theta}}^{\bar{\theta}}f(\theta^i)[X^i(\theta^i)(S(q^i(\theta^i))-c(q^i(\theta^i),\theta^i))-\pi^i(\theta^i)]\,d\theta^i \tag{13}$$

subject to (9) and (11). This problem is similar to a problem of monopoly regulation. In contrast to Laffont and Tirole (1987), (13) turns out to be a problem of countervailing incentives because of our specialization assumption. We will show in the following section (proposition 1) that the solution for q^i in (13) does not depend on the choice rule $(x^i)_{i=1,\dots,n}$ that we treated as given, i.e. the optimal q^i is the same for any x^i. This allows us to plug this optimal q^i into the principal's objective and maximize – in the third step – over the optimal choice rule $(x^i)_{i=1,\dots,n}$ (and rents $(\pi^i)_{i=1\dots n}$). As objective and constraints are linear in x^i, it is not surprising that the optimal choice rule is based on a suitably defined "virtual valuation", i.e. the firm with the highest virtual valuation is contracted.

For the following, it is useful to note that the optimal control function q^i in problem (13) has to maximize the Hamiltonian function $H = f[X^i(S - c) - \pi^i] + \lambda^i X^i c_\theta$ where $\lambda^i(\theta^i)$ denotes the costate associated with (9). The optimal q^i will – for types where $X_i(\theta^i) \neq 0$ – then satisfy the first order condition

$$f(\theta^i)(S_q(q^i(\theta^i)) - c_q(q^i(\theta^i), \theta^i)) + \lambda^i(\theta^i) c_{q\theta}(q^i(\theta^i), \theta^i) = 0. \tag{14}$$

The condition for optimal q^i includes a first best welfare term $(S_q - c_q)$ and a rent extraction term, i.e. increasing $q^i(\theta^i)$ will increase the slope of the rent function by (9) and $c_{q\theta} < 0$. If IC is binding downwards (upwards), this increases the rent for types above (below) θ^i. As firms are specialized, it is not clear whether higher or lower types are "better" and therefore upwards as well as downwards binding IC is possible. As in Maggi and Rodriguez-Clare (1995), the following notation proves useful. Let $q^h(\theta)$ denote the solution to

$$S_q(q(\theta)) - c_q(q(\theta), \theta) + \frac{1 - F(\theta)}{f(\theta)} c_{q\theta}(q(\theta), \theta) = 0 \tag{15}$$

and $q^l(\theta)$ the solution to

$$S_q(q(\theta)) - c_q(q(\theta), \theta) - \frac{F(\theta)}{f(\theta)} c_{q\theta}(q(\theta), \theta) = 0. \tag{16}$$

Put differently, q^h is the solution to the first order condition (14) when $\lambda^i(\theta^i) = 1 - F(\theta^i)$ –i.e. the IC constraint is binding downwards and increasing $q^i(\theta^i)$ creates an information rent for all types above θ^i– and q^l is the solution to (14) if $\lambda^i(\theta^i) = -F(\theta^i)$ –i.e. the IC constraint is binding upward and increasing $q^i(\theta^i)$ creates an information rent for all types below θ^i. Note that $q^l \geq q^{fb} \geq q^h$ with strict inequality for all but the boundary types as $c_{q\theta} < 0$.

The intuition follows the logic of the countervailing incentives literature. Firms θ producing $q(\theta) > k(\theta)$ have an incentive to report $\hat{\theta} < \theta$ to pretend to have higher costs ($c_\theta < 0$) and raise the transfer they receive. Firms θ producing $q(\theta) < k(\theta)$ report $\hat{\theta} > \theta$ to raise their transfer ($c_\theta > 0$). The planner wants to prevent mimicking while keeping information rents low. Hence, $c_{q\theta} < 0$ implies that q is distorted downwards in the former case and upwards in the latter.

4. First best welfare monotone

We will now characterize the optimal mechanism for the WM-case. It turns out that all firms are treated symmetrically which means that we can write $q(\theta)$ instead of $q^i(\theta)$, $\pi(\theta)$ instead of $\pi^i(\theta)$ etc. There are two cases to consider. In the first case, the solution (given by equation (15)) is similar to a setting where firms are not specialized. Put differently, optimal qualities are so high above $k(\theta)$ that higher types have lower costs in the relevant quality range. Consequently, the solution in this case is essentially the solution of a standard problem known in the literature. In the second case, low types up to a type $\theta_b \geq \underline{\theta}$ have zero profits (but with different quality levels) and from θ_b onwards, $q(\theta)$ follows q^h. In this case, the assumption that firms are specialized is relevant: all types below θ_b are assigned the quality $k(\theta)$ they are specialized in. In both cases, IC binds only downwards, i.e. high types would like to mimic low types (not the other way around).

Proposition 1. *The optimal mechanism treats all firms symmetrically, i.e.* $q^i = q$, $x^i = x$ *and* $\pi^i = \pi$ *for all* $i = 1, \dots, n$. *We define type* θ*'s virtual valuation as follows:*

$$VV(\theta) = S(q(\theta)) - c(q(\theta), \theta) + \frac{1 - F(\theta)}{f(\theta)} c_\theta(q(\theta), \theta). \tag{17}$$

There are two cases:

1. *If* $c_\theta(q^h(\underline{\theta}), \underline{\theta}) < 0$*, then* $q^h(\theta)$ *in equation (15) gives the optimal quality for all* $\theta \in [\underline{\theta}, \bar{\theta}]$*. Firm* i *with highest* $VV(\theta^i)$ *wins the procurement, i.e.* $x^i(\theta^i, \theta^{-i}) = 1$ *if* $VV(\theta^i) > VV(\theta^j)$ *for all* $j \neq i$*. We have* $\pi_\theta(\theta), q_\theta(\theta), X_\theta(\theta) > 0$ *for each* $\theta \in [\underline{\theta}, \bar{\theta}]$.

2. *If* $c_\theta(q^h(\underline{\theta}), \underline{\theta}) \geq 0$*, then there exists a largest* $\theta_b \geq \underline{\theta}$ *such that*

$$q(\theta) = k(\theta) \text{ for all } \theta \in [\underline{\theta}, \theta_b]$$

and θ_b *is determined by the unique solution to*

$$S_q(k(\theta_b)) - c_q(k(\theta_b), \theta_b) + \frac{1 - F(\theta_b)}{f(\theta_b)} c_{q\theta}(k(\theta_b), \theta_b) = 0. \tag{18}$$

For all $\theta > \theta_b$*, quality* $q(\theta) = q^h(\theta)$*. Firm* i *with highest* $VV(\theta^i)$ *wins the procurement, i.e.* $x^i(\theta^i, \theta^{-i}) = 1$ *if* $VV(\theta^i) > VV(\theta^j)$ *for all* $j \neq i$*. We have*

$$\pi(\theta) = 0 \text{ for all } \theta \in [\underline{\theta}, \theta_b],$$

$$\pi_\theta(\theta) > 0 \text{ for all } \theta \in (\theta_b, \bar{\theta}], \text{ and}$$

$$X_\theta(\theta), q_\theta(\theta) \geq 0 \text{ for all } \theta \in [\underline{\theta}, \bar{\theta}].$$

The relaxed solution is globally incentive compatible.

The virtual valuation includes next to first best welfare $(S - c)$ a rent extraction term. Roughly speaking, contracting a type θ^i with a higher probability, i.e. increasing $x^i(\theta^i, \theta^{-i})$, changes the slope of the rent function $\pi(\theta)$; see equation (9). If $q(\theta^i) > k(\theta^i)$, the rent function is increasing more steeply when $x^i(\theta^i, \theta^{-i})$ is increased. Hence, types above θ^i will get a higher rent. $1 - F(\theta^i)$ is the weight of the types that benefit from this higher rent. But if firm i produces its specialized quality $k(\theta^i)$, no other firm can mimic it profitably and hence the information rent disappears (as $c_\theta(k(\theta), \theta) = 0$). The firm with the highest VV is contracted.

In the WM case, VV is increasing in type. Apart from MHR and WM, the fact that firms are specialized is another reason for this: to say that firms are specialized we used $c_{\theta\theta} \geq 0$. This implies that the effect of marginally increasing the probability of being contracted on the rents—i.e. on π_θ in (9) bearing in mind that $c_\theta(q(\theta), \theta) \leq 0$ in the optimal mechanism—is smaller for higher types. This effect lets the principal prefer higher types and hence the highest θ is contracted. Thus, the planner chooses the type that generates the highest surplus; in this sense, the provider choice x is not distorted.

Quality q, however, is distorted. Although $S - c$ would be maximized by q^{fb}, this is too expensive to implement. The IC constraint causes a downward distortion of quality to reduce rents. There are two ways in which this can be done. First, the standard way: q^h as defined by equation (15). By reducing $q(\theta)$ below $q^{fb}(\theta)$, $S - c$ is reduced, but it allows for a reduction of rents as it becomes less attractive for types $\theta' > \theta$ to mimic θ. With specialized firms, there is a second way in which rents can be reduced: let θ produce quality $k(\theta)$ with zero profits. No type can profitably mimic θ in this case. This second option is clearly preferable to the first option if $q^h(\theta) < k(\theta) < q^{fb}(\theta)$ because

k leads then to less distortion than q^h without generating rents for higher types. In this sense, *distortion is limited* (by k) if firms are specialized. This is the case for types below θ_b. Given that the rent effect, i.e. c_θ in IC, continuously approaches 0 as $q \to k$, it is not surprising that the first option is optimal whenever $k(\theta) < q^h(\theta) < q^{fb}(\theta)$ which is the case for types above θ_b. Note that the quality distortion does not depend on the number of firms. Hence, we obtain the same distortion as in the principal agent model with countervailing incentives.[11]

In principle, the principal could guarantee production without paying rents (by setting $q(\theta) = k(\theta)$ for all θ). This is in contrast to standard models where rents can only be reduced to zero by excluding types from production. This, of course, leads to a risk that the service is not procured at all, depending on the draws of θ. Put differently, guaranteeing the service leads to strictly positive rents in standard models but not in our model. In this sense, the principal can *extract more rents* when firms are specialized and still guarantee the service.

Another way to compare our results with the standard procurement model is to compare case 2 of proposition 1 with the optimal menu in the hypothetical case where the cost function is $c(q,\theta) = h_1(q) + \mathbb{E}_\theta[h_2(\theta)] - \alpha q\theta$ instead of $c(q,\theta) = h_1(q) + h_2(\theta) - \alpha q\theta$; i.e. all firms have the same fixed costs and there is no tradeoff between marginal and fixed costs. This is a standard procurement model where the firm with the highest type is contracted and q^h is the optimal quality schedule. Furthermore, rents are strictly increasing and only type $\underline{\theta}$ has zero rents. It follows that also in this sense having specialized firms leads to lower rents (for all types apart from $\underline{\theta}$) and less quality distortion (for types below θ_b).

[11] To illustrate, if $k'(\theta) = 0$, some types are also bunched on the same quality: see figure 1 in Maggi and Rodriguez-Clare (1995) with $k'(\theta) = \bar{\nu}''(\theta) = 0$.

Example 2 (continued). *Let* $S(q) = q$ *such that* $q^{fb}(\theta) = 5\theta/4 > 4\theta/5 = k(\theta)$. *Equation (15) becomes*

$$1 - 2(q - \theta) - \left(1 - \frac{\theta}{2}\right) + (1 - \theta)\left(-\frac{5}{2}\right) = 0.$$

Hence, $q^h(\theta) = max\{0, 5\theta/2 - 5/4\}$. *Case 2 of proposition 1 applies because* $c_\theta(q^h(0), 0) = 0 \geq 0$. *The* q^h *function intersects* k *at* $\theta_b = 25/34$. *Hence, the optimal quality schedule is*

$$q(\theta) = \begin{cases} \frac{5}{2}\theta - \frac{5}{4} & \text{if } \theta \geq \frac{25}{34} \\ \frac{4}{5}\theta & \text{if } \theta < \frac{25}{34}. \end{cases}$$

The virtual valuation is

$$VV(\theta) = \begin{cases} \frac{25}{16} - \frac{17}{4}\theta + \frac{13}{4}\theta^2 & \text{if } \theta \geq \frac{25}{34} \\ \frac{9}{25}\theta^2 & \text{if } \theta < \frac{25}{34}. \end{cases}$$

As the virtual valuation is increasing in type, the firm with the highest type is contracted which implies that $X(\theta) = \theta^{n-1}$ *as types are uniformly distributed. Expected profits of types* $\theta \leq 25/34$ *are 0. Expected profits for types higher than* $25/34$ *are*

$$\pi(\theta) = \int_{\theta_b}^{\theta} -X(s)c_\theta(q(s), s)\, ds = \frac{17}{4(n+1)}\theta^{n+1} - \frac{25}{8n}\theta^n + \frac{25}{8n(n+1)}\left(\frac{25}{34}\right)^n.$$

5. First best welfare non-monotone

In this section, we analyze the case where first best welfare is first decreasing and then increasing in type. The lowest type $\underline{\theta}$ is no longer worst (in a first best sense) and therefore can have positive profits under the optimal mechanism. One can think of the WNM case as having two standard menus. One for lower θ in which lower types are better, the incentive constraint is upward binding, profits are decreasing in type and quality is distorted upwards. The other for higher θ with higher types being better, profits increasing in type, the incentive constraint downward binding and quality distorted downwards.

The way to connect these two standard menus is an interval of types with zero profits (but differing quality levels). Incentive compatibility within the zero profit interval is no problem here: Each zero profit type θ will produce the quality level $k(\theta)$ at which he has lower costs than any other type. The following proposition describes the optimal menu in the WNM case.

Proposition 2. *The optimal mechanism treats all firms in a symmetric way, i.e. $q^i = q$, $VV^i = VV$ and $\pi^i = \pi$ for all $i = 1, \ldots, n$.*
There exist θ_1 and θ_2, with $\theta_1 < \theta_2$, such that θ_1 and θ_2 are uniquely defined by $q^l(\theta_1) = k(\theta_1)$ and $q^h(\theta_2) = k(\theta_2)$. Virtual valuation is given by

$$VV(\theta) = \begin{cases} S(q(\theta)) - c(q(\theta), \theta) - \frac{F(\theta)}{f(\theta)} c_\theta(q(\theta), \theta) & \text{if } \theta < \theta_1 \\ S(q(\theta)) - c(q(\theta), \theta) & \text{if } \theta \in [\theta_1, \theta_2] \\ S(q(\theta)) - c(q(\theta), \theta) + \frac{1-F(\theta)}{f(\theta)} c_\theta(q(\theta), \theta) & \text{if } \theta > \theta_2. \end{cases} \tag{19}$$

Quality is determined by

$$q(\theta) = \begin{cases} q^h(\theta) & \text{for all } \theta > \theta_2 \\ k(\theta) & \text{for all } \theta \in [\theta_1, \theta_2] \\ q^l(\theta) & \text{for all } \theta < \theta_1. \end{cases} \tag{20}$$

The firm with the highest $VV(\theta)$ is contracted, i.e. $x^i(\theta^i, \theta^{-i}) = 1$ if $VV(\theta^i) > VV(\theta^j)$

for all $j \neq i$. *We have*

$$\begin{aligned} \pi(\theta) &= 0 \textit{ for all } \theta \in [\theta_1, \theta_2] \\ \pi_\theta(\theta) &< 0 \textit{ for all } \theta < \theta_1 \\ \pi_\theta(\theta) &> 0 \textit{ for all } \theta > \theta_2 \\ q_\theta(\theta) &\geq 0. \end{aligned}$$

Type θ_w*, who has the lowest first best welfare of all types, is in the zero profit interval and produces his first best quality. It holds that*

$$X_\theta(\theta) \leq 0 \textit{ for all } \theta < \theta_w$$

$$X_\theta(\theta) \geq 0 \textit{ for all } \theta > \theta_w.$$

The relaxed solution is globally incentive compatible.

[Figure 3 about here.]

Figure 3 illustrates proposition 2.[12] It follows from lemma 2 that θ_w is the worst type from a welfare point of view. Moreover, $q^{fb}(\theta_w) = k(\theta_w)$ implies that we can implement $q^{fb}(\theta)$ for θ_w without creating information rents ($\pi(\theta_w) = 0$ and no other type can profitably mimic θ_w). Hence, there is "no distortion at the bottom". For $\theta > \theta_w$, the quality schedule follows proposition 1: distort quality downwards to reduce rents by choosing either $k(\theta)$ or $q^h(\theta)$ depending on which yields higher welfare. The switch from one to the other happens at θ_2. At $\bar{\theta}$ we have "no distortion at the top": $q^h(\bar{\theta}) = q^{fb}(\bar{\theta})$.

[12]Similar figures appear in Maggi and Rodriguez-Clare (1995), see their figure 3 and 5. The reason is that our quality distortion does not depend on the number of firms and is therefore the same as in the setting with countervailing incentives.

Consider $\theta < \theta_w$. Moving to the left from θ_w we move to types that are better than θ_w. Types would like to mimic θ's above them and IC constraints bind upwards. To reduce rents, we need to distort quality upwards, i.e. $q(\theta) > q^{fb}(\theta)$: this follows from (9) combined with $c_\theta(q^{fb}(\theta), \theta) > 0$ for $\theta < \theta_w$ and $c_{q\theta} < 0$. There are two ways to do this: either $k(\theta)$ or $q^l(\theta)$ as defined in (16). Of these two, the quality level that is closest to $q^{fb}(\theta)$ is optimal; at θ_1 we switch from one to the other. Finally, $q^l(\underline{\theta}) = q^{fb}(\underline{\theta})$: also here we have "no distortion at the top".

The selection rule for the winner of the procurement is based on VV. As welfare is quasiconvex and $c_{\theta\theta} \geq 0$, VV is also quasiconvex. The worst type θ_w has the lowest probability of winning ($X(\theta_w) = 0$). Moving either to the left or to the right from θ_w increases the probability of winning as we move to better types. This non-monotonicity of VV and X causes two new issues. In standard models and in case 1 of proposition 1, VV, welfare and profits are strictly increasing in type. Although the shapes of these three functions differ, they all point in one direction: higher types lead to higher welfare, higher probability of winning and higher profits. Thus, first, the outcome is ex post efficient: the type generating the highest welfare wins. Second, the outcome is easy to implement: the type with the highest profit wins and therefore letting firms bid in a second price mechanism leads to a winner which generates the highest welfare. This is not optimal in our model; we come back to this in the next section.

Equation (19) shows that second best welfare $W^{sb}(\theta) = S(q(\theta)) - c(q(\theta), \theta)$ differs from $VV(\theta)$ by the information rent term. The planner, when assigning the contract, is willing to deviate from efficiency to reduce rents. In particular, by playing out high types against low types, the rents of high types can be reduced (and the other way around). A similar result is well known in auctions with asymmetric bidders. Myerson (1981) shows that it is optimal to discriminate between bidders drawing their valuations from different distributions. For example, if bidder A draws his valuation from a distribution putting more weight on high values and bidder B draws from a distribution with low values, the revenue maximizing auction will favor B. What is new in our case is that there is only one distribution from which types are drawn: types are ex ante symmetric. Discrimination in our model is due to different parts of the same distribution governing the distortion: For low θ, the left tail is relevant and for high types the right tail of the distribution matters for distortion.

The next result shows that it is not hard to find two types, where the winning type is not the one generating highest second best welfare. That is, the optimal mechanism is not second best efficient.

Corollary 1. *The optimal allocation is not second best efficient in the sense that there exist types θ', θ'' such that θ' wins against θ'' although $W^{sb}(\theta'') > W^{sb}(\theta')$.*

We discuss two implications of this result. First, as mentioned in the introduction, the specialization of firms often comes to the fore in industries that are being deregulated. In many of these industries, incumbent firms used to invest a lot in quality during regulation (for instance, because the regulation scheme in place stimulated this with subsidies). After deregulation, new firms come in which provide low quality at a low cost. The incumbents then tend to complain that procurement is biased toward low costs. In a standard set-up, this does not make sense: the high quality winner is more efficient at every quality level. In our set-up, it is possible that a high quality firm generates higher surplus in the optimal mechanism but loses against a low cost competitor.

In other words, although in second best a high type generates higher quality and higher welfare than some low type, it can happen that the low type wins the procurement contract. Incidentally, the opposite can happen as well: a high type wins from a low type although the latter generates higher (second best) welfare.

Second, this result reiterates that commitment is important. In standard models, commitment is important when it comes to quality. The mechanism induces firms to reveal their type. Once types are known, there is an incentive for planner and firm to renegotiate quality to move closer to first best. Commitment is needed to prevent this. In addition to this, we have a commitment problem with respect to selecting the winner. Once the planner knows firms' types, she may want to choose the one that yields highest (second best) welfare, however this is not optimal from an ex ante perspective.

6. Scoring rule auctions

A scoring rule auction is a procurement mechanism in which the principal designs a scoring rule and the firm bidding the highest score is contracted. A scoring rule is a function which assigns to each price/quality pair a real number that is called the "score". If price enters this function linearly, the scoring rule is said to be quasilinear. A second score auction is a straightforward extension of the famous Vickrey auction: The highest bidder is contracted and has to provide a quality/price combination resulting in the second highest score bid.

Scoring rule auctions are used in practice and have also received attention in the academic literature, see Asker and Cantillon (2008). Arguably, the procurement guidelines of the European Union favor scoring rules. If the procurement procedure is based on the concept of "best economic value", the procurement agency has to publish the relative weighting of the different criteria ex ante. Hence, the procurement mechanism will resemble a scoring rule auction.[13] Furthermore, Che (1993) shows that the optimal

[13]The guidelines allow for one alternative to the concept of best economic value: the criterion of

mechanism in a standard procurement model is implementable through a quasilinear second score auction. We will show that this result does not necessarily hold when firms are specialized even when we allow for general scoring rule auctions. We will then introduce an extended scoring rule auction which can implement the optimal mechanism with specialized firms.

In a second score auction, it is a dominant strategy to bid the highest score one can provide at non-negative profits. Denoting the scoring rule by $s(q,p)$, a firm of type θ will therefore have the bid

$$bid(\theta) = \max_{p,q} s(q,p) \qquad s.t.: p \geq c(q,\theta). \tag{21}$$

Naturally, the constraint will be binding and therefore we can write

$$bid(\theta) = \max_{q} s(q, c(q,\theta)).$$

Using the envelope theorem, bids change in type according to

$$bid_\theta(\theta) = s_p(q(\theta), c(q(\theta),\theta))c_\theta(q(\theta),\theta). \tag{22}$$

The last equation implies that $bid_\theta(\theta) = 0$ for all types with $q(\theta) = k(\theta)$. Recall that the optimal mechanism assigns $q(\theta) = k(\theta)$ to the types in the zero profit interval. Hence, all types with zero profits will have the same bid in a scoring rule auction implementing the optimal quality schedule. However, in the optimal mechanism as described in propositions 1 and 2, types in the zero profit interval typically have different virtual valuations and therefore different probabilities of being contracted.

In the appendix, we show that a similar reasoning also holds in first score auctions

"lowest price". Such a focus on price is clearly not optimal when quality matters.

which leads to the following result.

Proposition 3. *Generically, a scoring rule auction cannot implement the optimal mechanism in the WNM case. In the WM case, scoring rule auctions cannot implement the optimal mechanism in case 2 of proposition 1.*

In short, standard scoring rule auctions usually do not work when firms are specialized. Another way to implement the optimal mechanism in the non-specialized setup is to have a second price auction for the right to be regulated as a monopolist. That is, the procurement agency commits to the optimal regulation menu for the monopoly case and the winner of the second price auction can pick a quality/transfer pair from this menu. Again this does not work when firms are specialized. For instance, all types in the zero profit interval would bid zero in the second price auction. Hence, the procurement agency would have to treat those types in the same way although they have different virtual valuations in the optimal mechanism.

As indicated above, the main problem is that scoring rule auctions cannot discriminate between types with zero profits. This suggests that, at least, a tie breaking rule is needed to implement the optimal mechanism. In the following, we propose such an extended scoring rule auction. Our *dual-score auction with tie breaking* works in the following way. We use a simplified scoring rule auction in which firms bid a quality and a minimum price at which they are willing to provide this quality.[14] The principal sets up two scoring rules –A,B– and depending on whether the quality is below or above $q(\theta_w)$ a firm's bid is evaluated with scoring rule A or B. The firm with the highest score is contracted and has to provide the quality it bid. The price is determined as in a Vickrey auction: It is the price that yields the same score as the second highest score.

[14]In an earlier version of this paper, we presented a slightly more complicated auction that implemented the optimal mechanism in which firms bid a score without comitting to a quality.

As indicated above, types in the interval $[\theta_1, \theta_2]$ will have the same score in this auction. To break ties, the auctioneer uses the following rule: The bidder whose bid quality q maximizes $VV(q^{-1}(q))$ wins where VV and q^{-1} refer to the virtual valuation and the inverse of the quality function in the optimal mechanism derived in propostions 1 and 2.[15]

This means that the auctioneer has to announce the two scoring rules, the quality level $q(\theta_w)$ and the tie-breaking rule based on the virtual valuation. The scoring rules are designed in a way that will ensure that each type finds it profit maximizing to bid the optimal quality. We use the scoring rules $s_A(q,p) = G\left(S(q) - p + \Delta(q)\right)$ and $s_B(q,p) = S(q) - p + \Delta(q)$ where G is a strictly increasing function determined in the appendix and[16]

$$\Delta(q) = \begin{cases} \int_{q(\underline{\theta})}^{q} \frac{\lambda(q^{-1}(s))}{f(q^{-1}(s))} c_{q\theta}(s, q^{-1}(s))\, ds & \text{for } q \in [q(\underline{\theta}), q(\theta_w)] \\ \int_{q(\theta_w)}^{q} \frac{\lambda(q^{-1}(s))}{f(q^{-1}(s))} c_{q\theta}(s, q^{-1}(s))\, ds & \text{for } q \in (q(\theta_w), q(\bar{\theta})] \\ -\infty & \text{else.} \end{cases} \tag{23}$$

Because of the Vickrey design, it is dominant to bid $c(q,\theta)$ as minimum price where q is the quality the firm bids. In case of winning, a firm has profits of $S(q) + \Delta(q) - D - c(q,\theta)$ where D is a constant that depends only on the second highest score.[17] Note that the quality q maximizing these profits also maximizes the firm's score given that the bid price equals $c(q,\theta)$. As the same quality maximizes the probability of winning as well as the profits conditional on winning, it is a dominant strategy to bid this quality. The scoring rule is designed such that this quality is exactly the quality from the optimal

[15]Bids using qualities that are not in $[q(\underline{\theta}), q(\bar{\theta})]$ are loosing.

[16]Recall that $\lambda(\theta)$ is defined as $-F(\theta)$ if $\theta < \theta_1$, $1 - F(\theta)$ if $\theta > \theta_2$ and as the unique solution to $S_q(k(\theta)) - c_q(k(\theta), \theta) + \lambda c_{q\theta}(k(\theta), \theta) = 0$ for $\theta \in [\theta_1, \theta_2]$.

[17]D will be either $score^{(2)}$ or $G^{-1}(score^{(2)})$ depending on whether the winning firm bids a quality above or below $q(\theta_w)$.

mechanism. It turns out that the firms' rents in the auction are the same as in the optimal mechanism which leads to the following result.

Proposition 4. *The optimal mechanism can be implemented in dominant strategies by the dual-score auction with tie breaking described above.*

Clearly our dual score auction with tiebreaking is somewhat more complicated than a standard scoring rule auction. Interestingly, the added complication is mainly on the side of the auctioneer. Firms have a dominant strategy and compute their optimal bids in exactly the same way as in standard scoring rule auction. The added complication derives from the tie-breaking rule and the more complicated scoring rule. The exact tie-breaking rule is, however, quite irrelevant for firms: Ties only occur (with positive probability) for types that earn zero profits whenever they win. Consequently, these types are indifferent between winning and losing and do not have to care about the details of the tie-breaking rule. The scoring rule is not really more complicated in itself, i.e. it is still a function mapping qualities and prices into scores. The only complication is that it is a composite function piecing together two standard scoring rules. Again the complication is on the auctioneer's side while it makes little difference for the firms whether the auctioneer arrived at the rule in one or two steps.

The dual score auction also illustrates the additional commitment problem we discussed in section 5: Even if we abstract from the possibility to renegotiate qualities and prices, the auctioneer might want to accept a different bid than the one winning according to the auction rules. The reason is that two firms might be evaluated according to different scoring rules. Although types with a higher score will lead to a higher payoff for the auctioneer within scoring rule A and within scoring rule B, the opposite might be true when comparing types that are not both evaluated using the same rule. This is, again, an additional complication on the auctioneer's side.

Example 2 (continued)**.** *In the WM case, the dual-score auction with tiebreaking reduces to a* single-score *auction with tiebreaking or, put differently,* $\theta_w = \underline{\theta}$*. The optimal score is then* $s(q,p) = q - p + \Delta(q)$ *where (for* $q \in [q(\underline{\theta}), q(\bar{\theta})]$ *)*

$$\Delta(q) = \int_{q(\underline{\theta})}^{q} \frac{\lambda(q^{-1}(s))}{f(q^{-1}(s))} c_{q\theta}(s, q^{-1}(s))\, ds.$$

Plugging in the optimal mechanism derived earlier, gives[18]

$$\Delta_B(q) = \begin{cases} -\frac{9}{16}q^2 & \text{if } q \in [0, \frac{10}{17}) \\ \frac{25}{68} + \frac{1}{2}q^2 - \frac{5}{4}q & \text{if } q \in [\frac{10}{17}, \frac{5}{4}] \\ -\infty & \text{else.} \end{cases}$$

In the second score auction, it is a dominant strategy to bid the quality maximizing the score one can deliver at zero profits, i.e. $\max_q q - c(q,\theta) + \Delta(q)$*. Note that this maximization problem is strictly concave on* $[0, 5/4]$ *and the objective is continuously differentiable (even at* $10/17$*). If the* $\arg\max$ *to this maximization problem is in* $[0, 10/17)$*, the first order condition gives* $q(\theta) = 4\theta/5 = k(\theta)$ *(which is the optimal* q *for* $\theta < 25/34$*). If the* $\arg\max$ *is in* $[10/17, 5/4]$*, the first order condition gives* $q(\theta) = 5\theta/2 - 5/4$ *which is the optimal* q *for* $\theta \geq 25/34$*. In fact, the* $\arg\max$ *is below* $10/17$ *if and only if* $\theta < 25/34$*. For* $\theta < 25/34$*, we get a score of* $4\theta/5 - (\theta^2/25 + 4\theta/5 - 2\theta^2/5) - 9\theta^2/25 = 0$ *whereas the scores for higher types are positive and increasing in type. Hence, all types* $\theta < 25/34$ *have the same bid. In case all firms have types below* $25/34$*, the tie breaking rule chooses the highest type.*

As the dual-score auction can implement the optimal mechanism, let's return to

[18] λ can be derived from (14) using the optimal q. This gives $\lambda(\theta) = 1 - \theta$ for $\theta \geq 25/34$ and $\lambda(\theta) = 9\theta/25$ for $\theta < 25/34$.

our green energy example to see the consequences of technological progress. Innovation tends to make green energy like solar and wind energy cheaper. Hence, it becomes more attractive for firms to invest in such energy sources. The question is: how do the government's procurement rules react to these developments? At first sight, one may think that such green technological developments lead to "greener" procurement rules. In fact, they do not.

Think of such technological progress as increasing firms' θ over time. For concreteness, assume that $f(\theta) = a\theta + 1 - \frac{1}{2}a$ and $F(\theta) = \frac{1}{2}a\theta^2 + (1 - \frac{1}{2}a)\theta$ for $\theta \in [0, 1]$ and $a \in [0, 2]$. We interpret an increase in a as green innovations becoming available over time. It is routine to verify that $F(\theta)/f(\theta)$ falls with a whereas $(1 - F(\theta))/f(\theta)$ increases with a. It follows from proposition 2 that $q(\theta)$ falls for all $\theta < \theta_1$ and for all $\theta > \theta_2$ while leaving $q(\theta)$ unaffected for $\theta \in [\theta_1, \theta_2]$. Hence, the procurement rules reduce the quality demanded from a given type θ. Further, equation (19) shows that VV follows the same pattern.[19]

Summarizing, the government's rules bias against green technologies the more such technologies become available. A given type θ is required to provide lower quality and higher types are less likely to win. This may be hard to explain to environmental groups that argue that government rules should embrace green technologies, the more these become available.[20]

Further, from a dynamic perspective, if the government would like to stimulate firms to invest in green technologies it should deviate from the static optimal procurement rules derived above. Indeed, following these rules introduces a bias against high θ firms,

[19] As a increases, F/f increases and $(1 - F)/f$ falls while $c_\theta > 0$ for $\theta < \theta_1$ and $c_\theta < 0$ for $\theta > \theta_2$. There is also the indirect effect of a on $q(\theta)$ but by the definition of $q^{h,l}$ in equations (15, 16) this only has a second order effect on VV.

[20] As a technical point, without specialization the effect of $(1 - F)/f$ on quality q would be present as well, but a change in a would not affect the choice of the winner.

making it less attractive for firms to upgrade their θ. Again this requires commitment. The government should specify ex ante its energy procurement rules for each of the coming years in which it plans to procure electric power.

7. Conclusion

We analyzed a procurement setting in which the procurement agency cares not only about the price but also about the quality of the product. In the introduction we gave some examples where a firm is more efficient than another firm in producing some quality level but not necessarily in all quality levels.

Standard procurement models do not account for this possibility because "type" denotes efficiency and not how a firm is specialized. Put differently, a more efficient type produces cheaper at any quality level. We relax this assumption and allow each type to be specialized, i.e. to be the most efficient type for some quality level. This leads to a bunching of types on zero profits. The intuition is that distorting quality further than the quality level a type is specialized in (for rent extraction reasons) is not necessary: A type producing "his quality level" with expected profits of zero cannot be mimicked by any other type. Hence, the incentive constraint is slack and an interval of zero profit types is feasible. In short, distortion is limited and more rents are extracted if firms are specialized.

If we assume that first best welfare is U-shaped, e.g. there are gains from being specialized in low costs even from a welfare point of view, we get an interesting discrimination result. Types with lower second best welfare can be preferred to types with higher second best welfare. This is similar to auctions with asymmetric bidders where discriminatory mechanisms are well known. The commitment to favor some worse types allows the principal to reduce the rents of the best types. Put differently, competitive pressure can be exerted even by firms that are clearly worse. Further, “gold plating” can be optimal in the sense that some types produce quality levels above their first best levels. Finally, high quality firms can lose against low quality firms offering lower (second best) welfare. A complaint that is often heard in newly liberalized sectors.

Optimal mechanisms as the one derived in this article are usually not implemented one-to-one in reality. This makes our qualitative results on implementation especially relevant. First, the discrimination result mentioned above implies that the commitment problem of the principal is more severe when firms are specialized: As the optimal mechanism is second best inefficient, the principal might ex post want to contract another firm than the one the optimal mechanism demands. This commitment problem is in addition to the standard commitment problem; i.e. the principal would like to renegotiate quality and price after learning the firms’ types. Second, standard scoring rule auctions cannot implement the optimal mechanism when firms are specialized. The main problem is that all the types with zero expected rents will have the same bid in such a scoring rule auction. In standard mechanism design problems there is only one active type with zero rents. Here we have a range of such types. Therefore, well specified tie breaking rules become crucial when firms are specialized.

Appendix

Proof of lemma 2 From the first order condition for q^{fb}, we derive that

$$q_\theta^{fb} = \frac{\alpha}{-S_{qq}(q^{fb}(\theta)) + h_1''(q^{fb}(\theta))} > 0.$$

Hence, $q_\theta^{fb}(\theta) > k_\theta(\theta)$ at a type where $q^{fb}(\theta) = k(\theta)$ if and only if

$$\frac{\alpha}{-S_{qq}(q^{fb}(\theta)) + h_1''(q^{fb}(\theta))} > \frac{h_2''(\theta)}{\alpha}$$

which holds by assumption 2. Hence, q^{fb} can intersect k at at most one type and only from below. The type at which $q^{fb}(\theta) = k(\theta)$ is denoted by θ_w. As $W_\theta^{fb}(\theta) = -c_\theta(q^{fb}(\theta), \theta)$, this implies that W^{fb} has to be first de- and then increasing if q^{fb} intersects k and W^{fb} has to be monotone if q^{fb} does not intersect k; see lemma 1. This implies quasiconvexity. *Q.E.D.*

Proof of lemma 3 Take a direct mechanism consisting of rents $(\pi^i)_{i=1,\dots,n}$, choice rules $(x^i)_{i=1,\dots,n}$ and quality schedules $(q^i)_{i=1,\dots,n}$. Pick one particular i and suppose that q^i depends on θ^{-i}. We will now show that we can use $\hat{q}(\theta^i) = \frac{E_{\theta^{-i}}[x^i(\Theta)q^i(\Theta)]}{E_{\theta^{-i}}[x^i(\Theta)]}$ instead of q^i. Clearly $\hat{q}^i$ depends only on θ^i and we will show that (i) rents stay the same, (ii) incentive compatibility still holds, (iii) the principal's objective is weakly higher when using $\hat{q}^i$ instead of q^i.

First, we show that the slope of the rent function (i.e. first order incentive compatibility) stays the same. Cost function $c(q, \theta^i) = h_1(q) + h_2(\theta^i) - \alpha\theta^i q$ implies that (9) can be written as

$$\pi_\theta^i(\theta^i) = E_{\theta^{-i}}[-x^i(\Theta)h_2'(\theta^i) + x^i(\Theta)\alpha q^i(\Theta)].$$

Using the definition of $\hat{q}^i$ we get

$$\begin{aligned}\pi_\theta^i(\theta^i) &= E_{\theta^{-i}}[-x^i(\Theta)h_2'(\theta^i) + x^i(\Theta)\alpha q^i(\Theta)] \\ &= E_{\theta^{-i}}[-x^i(\Theta)h_2'(\theta^i)] + \alpha\hat{q}^i(\theta^i)E_{\theta^{-i}}[x^i(\Theta)] = E_{\theta^{-i}}[-x^i(\Theta)h_2'(\theta^i) + \alpha\hat{q}^i(\theta^i)x^i(\Theta)] \\ &= E_{\theta^{-i}}[x^i(\Theta)c_\theta(\hat{q}^i(\theta^i),\theta^i)].\end{aligned}$$

Put differently, (9) still holds if we use $\hat{q}^i$ instead of q^i (while keeping the same $(x^i)_{i=1,\dots,n}$). The mechanism $(\hat{q}^i, x^i, \pi^i)_{i=1,\dots,n}$ is therefore feasible in the relaxed program. Note that this mechanism satisfies the participation constraint because the original mechanism $(q^i, x^i, \pi^i)_{i=1,\dots,n}$ was assumed to do so.

Note that $\hat{q}^i$ is basically the expected quality provided by firm i in the original mechanism. Given that the principal's valuation is concave and costs are convex, it is therefore not surprising that the principal's objective $S - c - \pi$ is (weakly) higher when using $\hat{q}^i$ instead of q^i. More formally, write the principal's objective in the relaxed program as

$$\sum_{i=1}^n \int_{\theta^i} E_{\theta^{-i}}[x^i(\Theta)\left(S(q^i(\Theta)) - c(q^i(\Theta),\theta^i)\right)] - \pi^i(\theta^i)dF(\theta^i).$$

Now the integrand (for a given θ^i) is concave in q^i by assumption. This implies that substituting the expected q^i, i.e. $\hat{q}^i$, instead of q^i will increase the integrand. As this is true for any given θ^i, it is also true when we integrate over θ^i. In detail, define $K(q,\theta^i) = (S(q) - c(q,\theta^i))\,E_{\theta^{-i}}[x^i(\Theta)]$ and note that K is concave in q by our assumptions on S and c. Let H_{θ^i} denote the distribution over $[\underline{\theta},\bar{\theta}]^{n-1}$ that has density $h_{\theta^i}(\theta^{-i}) = f(\theta^1) * \cdots * f(\theta^{i-1}) * f(\theta^{i+1}) * \cdots * f(\theta^n) * \frac{x^i(\theta^i,\theta^{-i})}{E_{\tilde{\theta}^{-i}}[x^i(\theta^i,\tilde{\theta}^{-i})]}$. Taking expectations with respect to H_{θ^i} will be denoted by $E_{H_{\theta^i}}$. Note in particular that $\hat{q}^i(\theta^i) = E_{H_{\theta^i}} q^i(\theta^i,\theta^{-i})$

by the definition of $\hat{q}^i$. Then, the principal's objective can be written as

$$\sum_{i=1}^{n}\int_{\theta^i} E_{\theta^{-i}}\left[\frac{x^i(\theta^i,\theta^{-i})}{E_{\tilde{\theta}^{-i}}[x^i(\theta^i,\tilde{\theta}^{-i})]}K(q^i(\Theta),\theta^i)\right]-\pi^i(\theta^i)\,dF(\theta^i)$$

$$=\sum_{i=1}^{n}\int_{\theta^i}\int_{[\underline{\theta},\bar{\theta}]^{n-1}} K(q^i(\Theta),\theta^i)dH_{\theta^i}-\pi^i(\theta^i)\,dF(\theta^i)$$

$$=\sum_{i=1}^{n}\int_{\theta^i} E_{H_{\theta^i}}[K(q^i(\Theta),\theta^i)]-\pi^i(\theta^i)\,dF(\theta^i)\le\sum_{i=1}^{n}\int_{\theta^i} K(E_{H_{\theta^i}}[q^i(\Theta)],\theta^i)-\pi^i(\theta^i)\,dF(\theta^i)$$

$$=\sum_{i=1}^{n}\int_{\theta^i} K(\hat{q}^i(\theta^i),\theta^i)-\pi^i(\theta^i)\,dF(\theta^i)=\sum_{i=1}^{n}\int_{\theta^i} E_{\theta^{-i}}[x^i(\Theta)]\left(S(\hat{q}^i(\theta^i))-c(\hat{q}^i(\theta^i),\theta^i)\right)-\pi^i(\theta^i)dF(\theta^i)$$

$$=\sum_{i=1}^{n}\int_{\theta^i} E_{\theta^{-i}}[x^i(\Theta)\left(S(\hat{q}^i(\theta^i))-c(\hat{q}^i(\theta^i),\theta^i)\right)]-\pi^i(\theta^i)dF(\theta^i)$$

where the inequality follows from the fact that K is concave in q and the last step is true because no term in $S(\hat{q}^i(\theta^i))-c(\hat{q}^i(\theta^i),\theta^i)$ depends on θ^{-i}. The last expression is, of course, the principal's payoff when using the mechanism $(\hat{q}^i,x^i,\pi^i)_{i=1,\ldots,n}$. Consequently, the principal can achieve an at least as high payoff by using $\hat{q}^i$ which depends only on θ^i as she can by using a q^i that depends on θ^{-i} (in the relaxed program). As i was arbitrary, this concludes the proof. *Q.E.D.*

Proof of proposition 1 The structure of this proof is similar to Laffont and Tirole (1993, pp. 315). The only difference is that we solve a countervailing incentive problem instead of a standard principal agent problem in order to obtain the optimal q^i. In a first step, we will determine the optimal q^i for a given x^i. In a second step, we determine then the optimal x^i.

Following lemma 3, q^i depends on θ^i only. For a given $x^i(\cdot)$ (and therefore a given $X^i(\cdot)$), the principal's maximization with respect to q^i (12) can then be written as

$$\max_{(q^i,\pi^i)_{i=1\ldots n}}\sum_{i=1}^{n}\mathbb{E}_{\theta^i}\left\{\left[X^i(\theta^i)\left(S(q^i(\theta^i))-c(q^i(\theta^i),\theta^i)\right)-\pi^i(\theta^i)\right]\right\}$$

subject to (9) and (11). As X^i is given for the moment, the maximization problem is separable across firms, i.e. the maximization over q^i and π^i does not depend on q^j, π^j or θ^j for $j \neq i$. Hence, we can treat the maximization for firm i separately which then becomes

$$\max_{q^i,\pi^i} \int_{\underline{\theta}}^{\bar{\theta}} \left[X^i(\theta)\left(S(q^i(\theta)) - c(q^i(\theta),\theta)\right) - \pi^i(\theta)\right] f(\theta)\, d\theta \tag{24}$$

subject to

$$\begin{aligned} \pi^i_\theta(\theta) &= -X^i(\theta)c_\theta(q^i(\theta),\theta) \\ \pi^i(\theta) &\geq 0. \end{aligned}$$

This is an optimal control problem where q^i is the control and π^i is the state variable. To show that the quality schedule proposed in the proposition solves this problem, we use a sufficiency result for optimal control problems with pure state constraints (Seierstad and Sydsaeter, 1987, Thm. 1, ch. 5.2; adjusted to our notation) :

Theorem 1. *Let (q^*,π^*) be an admissible pair in problem (24). Let $\lambda^i : [\underline{\theta},\bar{\theta}] \to \mathbb{R}$ be a continuous and piecewise continuously differentiable function and $\eta^i : [\underline{\theta},\bar{\theta}] \to \mathbb{R}_+$ be a piecewise continuous function such that $\eta^i(\theta)\pi^*(\theta) = 0$ for all $\theta \in [\underline{\theta},\bar{\theta}]$. If the following properties are satisfied, (q^*,π^*) solves problem (24):*

- *q^* maximizes $H(\pi^*(\theta),q,\lambda^i(\theta),\theta) = f(\theta)\left[X^i(\theta)(S(q) - c(q,\theta)) - \pi(\theta)\right] + \lambda^i(\theta)X^i(\theta)c_\theta(q,\theta)$ for every $\theta \in [\underline{\theta},\bar{\theta}]$*
- *$\lambda^i_\theta(\theta) = -f(\theta) + \eta^i(\theta)$*
- *$\lambda^i(\bar{\theta})\pi^*(\bar{\theta}) = \lambda^i(\underline{\theta})\pi^*(\underline{\theta}) = 0, \lambda^i(\bar{\theta}) \geq 0$ and $\lambda^i(\underline{\theta}) \geq 0$*
- *$H(\pi,q^*(\theta),\lambda^i(\theta),\theta) = f(\theta)\left[X^i(\theta)(S(q^*(\theta)) - c(q^*(\theta),\theta)) - \pi\right] + \lambda^i(\theta)X^i(\theta)c_\theta(q,\theta)$ is concave in π for all θ.*

We immediately turn to case 2 of the proposition as the proof of case 1 resembles the proof of case 2 for types $\theta > \theta_b$. Before turning to the conditions of the theorem, we establish that θ_b is unambiguously defined. Note that the left hand side of (18) is increasing in θ_b as—by assumptions 1-2—its derivative can be written as

$$c_{q\theta}(k(\theta),\theta)\left(-1+\frac{(-S_{qq}(q(\theta))+c_{qq}(k(\theta),\theta))c_{\theta\theta}(k(\theta),\theta)}{c_{q\theta}^2(k(\theta),\theta)}+\frac{d\frac{1-F(\theta)}{f(\theta)}}{d\theta}\right)>0$$

Hence, θ_b is uniquely defined by (18). Note that this implies that q^h intersects k only at type θ^b and that $q^h(\theta) > (<)k(\theta)$ for all $\theta > (<)\theta_b$.

To check the conditions of the theorem, we propose $\lambda^i(\theta) = 1-F(\theta)$ and $q(\theta) = q^h(\theta)$ for $\theta \geq \theta_b$ and let $\lambda^i(\theta)$ be implicitly (and uniquely) defined by (14) and $q(\theta) = k(\theta)$ for $\theta < \theta_b$. Thus we propose that λ^i, q^i are independent from the chosen X^i (or x^i). We show that this is consistent with the optimality requirements of the theorem.

For π^i we propose $\pi^i(\theta) = 0$ for $\theta \leq \theta_b$ and $\pi^i(\theta) = \int_{\theta_b}^{\theta} -X^i(\tilde{\theta})c_\theta(q(\tilde{\theta}),\tilde{\theta})\,d\tilde{\theta}$ for $\theta > \theta_b$. Thus, π^i does depend on the choice of X^i and these will be determined jointly.

For the proposed q, $H_q(\pi^*(\theta), q(\theta), \lambda(\theta), \theta) = 0$. As $H_{qq} < 0$ by the assumptions $S_{qq} \leq 0$, $c_{qq} > 0$ and $c_{qq\theta} = 0$, the first requirement of the theorem is met.

Note that the proposed λ^i is continuous by the definition of θ_b and the continuity of c, S and f. Because we assumed that f is a C^1 function, S is a C^2 function and c is a C^3 function, λ^i is continuously differentiable at all types with the possible exception of θ_b. Furthermore, λ^i does not depend on X^i nor on the firm identifier i and we can therefore write λ without the firm identifier i.

To check the second condition of the theorem, define $\eta(\theta)$ (which also does not depend on X^i nor on the firm identifier i) as

$$\eta(\theta) = \begin{cases} 0 & \text{for } \theta \geq \theta^b \\ \lambda_\theta(\theta) + f(\theta) & \text{otherwise.} \end{cases}$$

This clearly satisfies the second condition and piecewise continuity but it remains to check that $\eta(\theta) \geq 0$ for all $\theta < \theta_b$. This is done in two steps:

First, in the WM case, $dW^{fb}/d\theta = -c_\theta(q^{fb}(\theta), \theta) > 0$ and therefore $q^{fb}(\theta) > k(\theta)$. This implies together with $q^h(\theta) < k(\theta)$ for all $\theta < \theta_b$ that $q^h(\theta) < k(\theta) = q(\theta) < q^{fb}$ for types $\theta < \theta_b$. Therefore, $\lambda(\theta) \in (0, 1 - F(\theta))$ for $\theta < \theta_b$.

Second, differentiating (14) with respect to θ evaluated at $q(\theta) = k(\theta)$, yields—after plugging in $k_\theta = -c_{\theta\theta}/c_{q\theta}$—for types $\theta < \theta_b$:

$$c_{q\theta}(k(\theta), \theta) \left(-1 + \frac{(-S_{qq}(q(\theta)) + c_{qq}(k(\theta), \theta))c_{\theta\theta}(k(\theta), \theta)}{c_{q\theta}^2(k(\theta), \theta)} + \frac{d\frac{\lambda(\theta)}{f(\theta)}}{d\theta} \right) = 0. \tag{25}$$

By assumption 2 –with $c_{q\theta} = -\alpha, c_{\theta\theta} = h_2'', c_{qq} = h_1''$– the sum of the first two terms in brackets is negative. Hence, the third term has to be positive. This implies $\lambda_\theta(\theta) \geq \lambda(\theta) f_\theta(\theta)/f(\theta)$. Therefore, we get

$$\eta(\theta) \geq \frac{1}{f(\theta)} \left(f^2(\theta) + \lambda(\theta) f_\theta(\theta) \right) \geq 0$$

where the second inequality follows from MHR and $\lambda(\theta) \in (0, 1 - F(\theta))$.

The last two conditions of the theorem are clearly met: $\lambda(\bar{\theta}) = \pi(\underline{\theta}) = 0$ and H is linear (and therefore concave) in π.

Hence, we showed that the quality schedule in the proposition is optimal and derived the optimal λ (for any given selection rule $(x^i)_{i=1\ldots n}$). Both q and λ do not depend on firm identifier i nor the given x^i.

Under the optimal q, $q(\theta) = q^h(\theta) > k(\theta)$ for $\theta > \theta_b$ and $q(\theta) = k(\theta)$ for $\theta < \theta_b$. (9) implies then $\pi^i_\theta(\theta) = 0$ for $\theta < \theta_b$ and $\pi^i_\theta(\theta) > 0$ for $\theta > \theta_b$. Hence, (11) is binding in the optimal mechanism for θ if and only if $\theta \leq \theta_b$. For a given x^i, the optimal rents $\pi^i(\theta)$ for $\theta \geq \theta_b$ are therefore given by $\int_{\theta_b}^{\theta} -X^i(\tilde{\theta})c_\theta(q(\tilde{\theta}), \tilde{\theta})\, d\tilde{\theta}$ as we proposed above. As $c_\theta(q(\theta), \theta) = 0$ for $\theta \leq \theta_b$ under the optimal q, we can equivalently write $\pi^i(\theta) = \int_{\underline{\theta}}^{\theta} -X^i(\tilde{\theta})c_\theta(q(\tilde{\theta}), \tilde{\theta})\, d\tilde{\theta}$.

When solving for the optimal decision rule x^i we can therefore plug $\pi^i(\theta^i) = \int_{\underline{\theta}}^{\theta^i} -X^i(\tilde{\theta})c_\theta(q(\tilde{\theta}), \tilde{\theta})\, d\tilde{\theta}$ into the principal's objective. Also plugging in the optimal q derived above, the principal's problem becomes

$$\max_{(x^i)_{i=1\ldots n}} \int_\Theta \left\{ \sum_{i=1}^n \left[x^i(\Theta) \left(S(q(\theta^i)) - c(q(\theta^i), \theta^i) \right) + \int_{\underline{\theta}}^{\theta^i} X^i(\tilde{\theta})c_\theta(q(\tilde{\theta}), \tilde{\theta})\, d\tilde{\theta} \right] \right\} f(\theta^1)\ldots f(\theta^n)\, d\Theta$$

subject to the feasibility constraints $x^i(\Theta) \geq 0$ and $\sum_i x^i(\Theta) \leq 1$ for all Θ. Using integration by parts and $X^i(\theta^i) = \mathbb{E}_{\Theta^{-i}}\, x^i(\Theta)$, we eliminate the inner integral and get

$$\max_{(x^i)_{i=1\ldots n}} \int_\Theta \left\{ \sum_{i=1}^n \left[x^i(\Theta) \left(S(q(\theta^i)) - c(q(\theta^i), \theta^i) + \frac{1 - F(\theta^i)}{f(\theta^i)} c_\theta(q(\theta^i), \theta^i) \right) \right] \right\} f(\theta^1)\ldots f(\theta^n)\, d\Theta.$$

As this expression is linear in each x^i, it is optimal to contract the firm i, i.e. set $x^i(\Theta) = 1$, for which the virtual valuation $VV(\theta^i) \equiv S(q(\theta^i)) - c(q(\theta^i), \theta^i) + \frac{1-F(\theta^i)}{f(\theta^i)} c_\theta(q(\theta^i), \theta^i)$ is highest.[21]

It remains to show the monotonicity results in proposition 1. Monotonicity of q follows from $k_\theta = -c_{\theta\theta}/c_{q\theta} \geq 0$ for $\theta \leq \theta_b$. For types $\theta > \theta_b$, $q_\theta(\theta) > 0$ holds as $q^h_\theta > 0$ under assumption 1 (by MHR and $c_{qq\theta} = c_{q\theta\theta} = 0$). The virtual valuation is increasing

[21] If the virtual valuation of all firms happens to be below 0, it is optimal not to procure. We ignore this case as, by assumption 1, S is high enough such that it is always optimal to procure.

in type as

$$\frac{dVV}{d\theta} = -c_\theta(q(\theta),\theta)\left(1 - \frac{d\frac{1-F(\theta)}{f(\theta)}}{d\theta}\right) + c_{\theta\theta}(q(\theta),\theta)\frac{1-F(\theta)}{f(\theta)} \geq 0 \tag{26}$$

where the inequality holds because the term in brackets is positive by MHR for $\theta > \theta_b$. The inequality holds strictly for $\theta > \theta_b$ and also for $\theta \leq \theta_b$ if $c_{\theta\theta} > 0$.

Global incentive compatibility of the solution in the relaxed program is shown in lemma 4 below. *Q.E.D.*

Lemma 4. *The relaxed solution in proposition 1 is globally incentive compatible.*

Proof of lemma 4 For global incentive compatibility we first show that no θ can profitably misrepresent as $\hat{\theta} > \theta$. This is true if

$$\pi(\theta) - \pi(\hat{\theta}) - X(\hat{\theta})[c(q(\hat{\theta}),\hat{\theta}) - c(q(\hat{\theta}),\theta)] \geq 0.$$

Using (9), this can be rewritten as

$$\int_\theta^{\hat{\theta}} X(t)c_\theta(q(t),t) - X(\hat{\theta})c_\theta(q(\hat{\theta}),t)\,dt \geq 0.$$

This last inequality can be rewritten as

$$\int_\theta^{\hat{\theta}}\int_t^{\hat{\theta}} X_\theta(s)c_\theta(q(s),t) + X(s)c_{q\theta}(q(s),t)q_\theta(s)\,ds\,dt \leq 0. \tag{27}$$

The second term of the integrand is negative by the monotonicity of $q(\theta)$ in proposition 1. Note that we saw in the proof of proposition 1 that $c_\theta(q(\theta),\theta) \leq 0$ for all types. As $t \leq s$ and $c_{\theta\theta} \geq 0$, clearly $c_t(q(s),t) \leq 0$ in the first term of the integrand. As $X_\theta \geq 0$ in proposition 1, inequality (27) has to hold.

To show that no θ gains by misrepresenting as $\hat{\theta} < \theta$ we use the following notation introduced in equation (8):

$$\pi(\hat{\theta}, \theta) = t(\hat{\theta}) - X(\hat{\theta})c(q(\hat{\theta}), \theta)$$

The idea is to define the following cost function

$$\tilde{c}(a, \theta) = \min\{c(q(a), a), c(q(a), \theta)\} \tag{28}$$

where $q(a)$ is the optimal quality schedule derived in proposition 1. Next define

$$\tilde{\pi}(a, \theta) = t(a) - X(a)\tilde{c}(a, \theta). \tag{29}$$

The following inequalities show that the solution derived above satisfies IC globally as well:

$$\begin{aligned}
& \pi(\hat{\theta}, \theta) - \pi(\theta, \theta) \\
\leq\ & \tilde{\pi}(\hat{\theta}, \theta) - \tilde{\pi}(\theta, \theta) \\
=\ & \int_{\theta}^{\hat{\theta}} \frac{\partial \tilde{\pi}(a, \theta)}{\partial a} da \\
=\ & \int_{\hat{\theta}}^{\theta} \left(\left. \frac{\partial \pi(a, \theta)}{\partial a} \right|_{\theta = a} - \frac{\partial \tilde{\pi}(a, \theta)}{\partial a} \right) da && (30) \\
=\ & \int_{\hat{\theta}}^{\theta} X_\theta(a)(\tilde{c}(a, \theta) - c(q(a), a)) + X(a)(\tilde{c}_a(a, \theta) - c_q(q(a), a)q_\theta(a))da && (31) \\
\leq\ & 0
\end{aligned}$$

where the first inequality follows from the definition of $\tilde{c}(\cdot)$ and the observation that $\tilde{\pi}(\theta, \theta) = \pi(\theta, \theta)$. Equation (30) follows because $\left. \frac{\partial \pi(a, \theta)}{\partial a} \right|_{\theta = a} = 0$ by the first order condition of truthful revelation. Equation (31) follows from the definitions of the derivatives of $\pi(a, \theta)$ and $\tilde{\pi}(a, \theta)$ w.r.t. a. The final inequality follows from the properties of the optimal mechanism $X_\theta(a), q_\theta(a) \geq 0$ and the following three observations. First, by definition of $\tilde{c}(\cdot)$ we have

$$\tilde{c}(a,\theta) - c(q(a),a) \leq 0$$

Second, for values of a where $\tilde{c}(a,\theta) = c(q(a),\theta)$ we have

$$\tilde{c}_a(a,\theta) - c_q(q(a),a)q_a(a) = (c_q(q(a),\theta) - c_q(q(a),a))q_a(a) \leq 0$$

because $c_{q\theta} \leq 0$ and $\theta \geq a$. Finally for values where $\tilde{c}(a,\theta) = c(q(a),a)$ we have

$$\tilde{c}_a(a,\theta) - c_q(q(a),a)q_a(a) = \left.\frac{\partial c(q(a),\theta)}{\partial \theta}\right|_{\theta=a} \leq 0$$

because in our solution $c_\theta(q(\theta),\theta) \leq 0$ for all θ. *Q.E.D.*

Proof of proposition 2 Type θ_w is determined by the intersection of q^{fb} and k which is unique by lemma 2. We have to show that $\theta_1 < \theta_w < \theta_2$. Assumptions 1 (in particular, MHR and the assumptions on third derivatives of c) and 2 imply that the left hand sides of (15) and (16) are both increasing in θ if $q(\theta) = k(\theta)$. Hence, θ_1 and θ_2 are unique. As

$$S_q(k(\theta)) - c_q(k(\theta),\theta) - \frac{F(\theta)}{f(\theta)}c_{q\theta}(k(\theta),\theta) > S_q(k(\theta)) - c_q(k(\theta),\theta) + \frac{1-F(\theta)}{f(\theta)}c_{q\theta}(k(\theta),\theta)$$

for all θ, it follows that $\theta_1 < \theta_2$. As $S_q(k(\theta_w)) - c_q(k(\theta_w),\theta_w) = 0$, $\theta_w \in (\theta_1,\theta_2)$.

The optimal contract of proposition 2 for types above θ_w is similar to the optimal contract in proposition 1. It is straightforward to check the optimality conditions of theorem 1 as in the proof of proposition 1. Define $\lambda(\theta) = -F(\theta)$ for types $\theta \leq \theta_1$ and let $\lambda(\theta)$ be defined by (14) with $q(\theta) = k(\theta)$ for types $\theta \in (\theta_1, \theta_w]$. Note that $q^{fb}(\theta) < q(\theta) = k(\theta) < q^l(\theta)$ holds for types in (θ_1,θ_w). Hence, $\lambda(\theta) \in (-F(\theta), 0)$ for

$\theta \in (\theta_1, \theta_w)$. From there, all steps of the proof of proposition 1 showing the optimality of q and λ go through. By (9), π^i is decreasing on $[\underline{\theta}, \theta_1)$, constant on (θ_1, θ_2) and increasing on $(\theta_2, \bar{\theta}]$ under the optimal q. Hence, (11) is binding for θ if and only if $\theta \in [\theta_1, \theta_2]$. Using (9) and $\theta_w \in [\theta_1, \theta_2]$, rents (for any given x^i) under the optimal q can be written as $\pi^i(\theta^i) = \int_{\theta_w}^{\theta^i} -X^i(\tilde{\theta}) c_\theta(q(\tilde{\theta}), \tilde{\theta})\, d\tilde{\theta}$.

To derive the optimal x^i, we plug the optimal q and $\pi^i(\theta^i) = \int_{\theta_w}^{\theta^i} -X^i(\tilde{\theta}) c_\theta(q(\tilde{\theta}), \tilde{\theta})\, d\tilde{\theta}$ into the principal's objective function. This leads to the maximization problem:

$$\max_{(x^i)_{i=1\dots n}} \sum_{i=1}^{n} \int_{\Theta} \left\{ \left[x^i(\Theta) \left(S(q^i(\theta^i)) - c(q^i(\theta^i), \theta^i) \right) + \int_{\theta_w}^{\theta^i} X^i(\tilde{\theta}) c_\theta(q(\tilde{\theta}), \tilde{\theta})\, d\tilde{\theta} \right] \right\} f(\theta^1) \dots f(\theta^n)\, d\Theta$$

$$= \max_{(x^i)_{i=1\dots n}} \sum_{i=1}^{n} \int_{\Theta^{-i}} \left\{ \int_{\underline{\theta}}^{\theta_w} \left[x^i(\Theta) \left(S(q^i(\theta^i)) - c(q^i(\theta^i), \theta^i) \right) - \int_{\theta^i}^{\theta_w} X^i(\tilde{\theta}) c_\theta(q(\tilde{\theta}), \tilde{\theta})\, d\tilde{\theta} \right] f(\theta^i)\, d\theta^i \right.$$

$$\left. + \int_{\theta_w}^{\bar{\theta}} \left[x^i(\Theta) \left(S(q^i(\theta^i)) - c(q^i(\theta^i), \theta^i) \right) + \int_{\theta_w}^{\theta^i} X^i(\tilde{\theta}) c_\theta(q(\tilde{\theta}), \tilde{\theta})\, d\tilde{\theta} \right] f(\theta^i)\, d\theta^i \right\} \left(\prod_{j \neq i} f(\theta^j) \right) d\Theta^{-i}.$$

Using $X^i(\theta^i) = \mathbb{E}_{\Theta^{-i}}\, x^i(\Theta)$ and integration by parts to eliminate the inner integral, this can be rewritten as

$$\max_{(x^i)_{i=1\dots n}} \sum_{i=1}^{n} \int_{\Theta^{-i}} \left\{ \int_{\underline{\theta}}^{\theta_w} x^i(\Theta) \left[\left(S(q^i(\theta^i)) - c(q^i(\theta^i), \theta^i) \right) - \frac{F(\theta^i)}{f(\theta^i)} c_\theta(q(\theta^i), \theta^i) \right] f(\theta^i)\, d\theta^i \right.$$

$$\left. + \int_{\theta_w}^{\bar{\theta}} \left[x^i(\Theta) \left(S(q^i(\theta^i)) - c(q^i(\theta^i), \theta^i) \right) + \frac{1 - F(\theta^i)}{f(\theta^i)} c_\theta(q(\theta^i), \theta^i) \right] f(\theta^i)\, d\theta^i \right\} \left(\prod_{j \neq i} f(\theta^j) \right) d\Theta^{-i}.$$

Because $c_\theta(q(\theta^i), \theta^i) = 0$ for all $\theta^i \in [\theta_1, \theta_2]$ and as $\lambda(\theta^i) = -F(\theta^i)$ for $\theta^i < \theta_1$ and $\lambda(\theta^i) = 1 - F(\theta^i)$ for $\theta^i > \theta_2$, this is equivalent to

$$\max_{(x^i)_{i=1\dots n}} \sum_{i=1}^{n} \int_{\Theta} \left\{ x^i(\Theta) \left[\left(S(q^i(\theta^i)) - c(q^i(\theta^i), \theta^i) \right) + \frac{\lambda(\theta^i)}{f(\theta^i)} c_\theta(q(\theta^i), \theta^i) \right] \right\} f(\theta^1) \dots f(\theta^n)\, d\Theta.$$

By the linearity of the objective in x^i, it is therefore optimal to contract the firm with the highest virtual valuation. Note that $dVV/d\theta \leq 0$ for $\theta < \theta_w$ as $\lambda(\theta) < 0$ and

$c_\theta(q(\theta),\theta) \geq 0$ for these types (similar derivation as (26)).

Lemma 5 below shows that the solution of the relaxed program is also globally incentive compatible. *Q.E.D.*

Lemma 5. *The relaxed solution in proposition 2 is globally incentive compatible.*

Proof of lemma 5 All $\theta \in [\theta_1, \theta_2]$ produce at $k(\theta)$ which is the quality level at which a type has lower cost than any other type. As these types also have zero profits, no other type can profitably misrepresent as $\theta \in [\theta_1, \theta_2]$. For $\theta \geq \theta_w$ the menu is equivalent to the one described in proposition 1. Therefore, lemma 4 implies non-local IC on this part of the menu. The same proof as for lemma 4 with reversed signs implies that the menu for $\theta < \theta_w$ is non-locally IC.

What remains to be shown is that no type $\theta < \theta_w$ can profitably misrepresent as $\theta' > \theta_w$ (and the other way round). Take such a θ and observe that θ_2 has lower costs at $q(\theta')$:

$$c(q(\theta'),\theta_2) - c(q(\theta'),\theta) = \int_\theta^{\theta_2} c_\theta(q(\theta'),t)\; dt < 0. \tag{32}$$

The inequality follows from the fact that $k(\theta), k(\theta_2) < q(\theta')$ and $c_{q\theta} < 0$. Therefore, the integrand is negative over the whole range. Incentive compatibility for θ requires

$$\begin{aligned} \pi(\theta) &\geq \pi(\theta') + X(\theta')[c(q(\theta'),\theta') - c(q(\theta'),\theta)] \\ &= \underbrace{\pi(\theta') + X(\theta')[c(q(\theta'),\theta') - c(q(\theta'),\theta_2)]}_{\leq 0} + X(\theta')[c(q(\theta'),\theta_2) - c(q(\theta'),\theta)]. \end{aligned}$$

The first term in the last expression is negative because incentive compatibility between θ_2 and θ' is satisfied (see lemma 4 and recall that $\pi(\theta_2) = 0$). The second term is also negative because of equation (32). As $\pi(\theta) \geq 0$, the inequality above and therefore incentive compatibility holds.

The proof for $\theta > \theta_w$ and $\theta' < \theta_w$ works in the same way with θ_1 in place of θ_2.

Q.E.D.

Proof of corollary 1 Consider $\theta' = \underline{\theta}$. Define $\underline{W} = W^{fb}(\underline{\theta}) = W^{sb}(\underline{\theta})$. As $\underline{\theta}$ produces his first best quality and first best welfare is decreasing at $\underline{\theta}$, there are types $\theta > \underline{\theta}$ with lower welfare than $\underline{W}$. By the definition of the (WNM)-case, $W^{fb}(\bar{\theta}) > \underline{W}$.

Taking these two points together and applying the intermediate value theorem yields the existence of a type θ'' such that $W^{sb}(\theta'') = \underline{W}$ and $W^{sb}_{\theta}(\theta'') > 0$.

$$\frac{dW^{sb}(\theta)}{d\theta} = (S_q(q(\theta)) - c_q(q(\theta),\theta))q_\theta(\theta) - c_\theta(q(\theta),\theta) = -\frac{\lambda(\theta)}{f(\theta)}c_{q\theta}(q(\theta),\theta)q_\theta(\theta) - c_\theta(q(\theta),\theta)$$

where the first order condition for q is used for the second equality. We know from proposition 2 and its proof that λ changes sign and c_θ (weakly) changes sign at θ_w. Consequently, $W^{sb}_{\theta}(\theta'') > 0$ implies $\lambda(\theta'') > 0$ and $c_\theta(q(\theta''),\theta'') \le 0$.

The virtual valuation can be written as

$$VV(\theta) = W^{sb}(\theta) + \frac{\lambda(\theta)}{f(\theta)}c_\theta(q(\theta),\theta)$$

and thus $VV(\theta) \le W^{sb}(\theta)$ since λ and c_θ have opposite signs and the inequality is strict if $\lambda(\theta), c_\theta(q(\theta),\theta) \ne 0$.

If $c_\theta(q(\theta''),\theta'') < 0$, it follows that $VV(\underline{\theta}) > VV(\theta'')$. By continuity of W^{sb}, there exist types θ that yield strictly higher welfare than $\underline{\theta}$ but still lose from $\underline{\theta}$ in the procurement.

Now consider the case where $\theta'' \in (\theta_1, \theta_2)$ such that $c_\theta(q(\theta''),\theta'') = 0$. In this case, there are types slightly above $\underline{\theta}$ that lose from types slightly below θ'' although the former yield higher (second best) welfare W^{sb}. *Q.E.D.*

Proof of proposition 3 We will mainly utilize the following obvious result: If the scoring rule implements the optimal mechanism it has to hold that $bid(\theta') = bid(\theta'')$ whenever $VV(\theta') = VV(\theta'')$ under the optimal mechanism.

We first focus on the second score auction. We showed in the main text that $bid_\theta = 0$ in the zero profit interval. Implicitly, this argument relied on the fact that the $arg\,max$ of (21) is also the quality provided by type θ if he wins the auction. To see this, note that the firm maximizes $p - c(q, \theta)$ subject to $s(q, p) = bid^{(2)}$ when winning. If the auction implements the optimal mechanism (in which q depends only on θ), the scoring rule has to be such that the optimal q does not depend on the second highest bid. But then this maximization is clearly solved by the same q as (21), i.e. the solution to both problems satisfies the first order condition $s_q(q, p) = -s_p(q, p)c_q(q, \theta)$.

Take θ_1 and θ_2 as defined in proposition 2.[22] Because all $\theta \in (\theta_1, \theta_2)$ have $q(\theta) = k(\theta)$, it follows that $bid_\theta(\theta) = 0$ for these types and therefore $bid(\theta_1) = bid(\theta_2)$. As virtual valuation and bids are continuous in type, this implies that $VV(\theta_1) = VV(\theta_2)$ has to hold if the scoring rule implements the optimal mechanism: Otherwise, types slightly below θ_1 and slightly above θ_2 have the same bid but different virtual valuations. Since $q(\theta_i) = k(\theta_i)$, the virtual valuation for θ_i is $S(k(\theta_i)) - c(k(\theta_i), \theta_i)$ for $i = 1, 2$. Consequently, the following equation has to hold if the scoring rule implements the optimal mechanism:

$$\int_{\theta_1}^{\theta_2} \frac{d\{S(k(\theta)) - c(k(\theta), \theta)\}}{d\theta}\, d\theta = 0$$

This can be rewritten as

$$\int_{\theta_1}^{\theta_2} \frac{(S_q(k(\theta)) - c_q(k(\theta), \theta))c_{\theta\theta}(k(\theta), \theta)}{-c_{q\theta}(k(\theta), \theta)}\, d\theta = 0. \qquad (33)$$

[22] For the WM case, an analogous argument can be made with $\theta_1 = \underline{\theta}$ and $\theta_2 = \theta_b$ as defined in proposition 1.

(33) uniquely pins down θ_2 for a given θ_1 if $c_{\theta\theta}(k(\theta_2), \theta_2) \neq 0$.[23] For generic cost functions satisfying our assumptions, $c_{\theta\theta}(k(\theta_2), \theta_2) \neq 0$ holds. Hence, generically (33) pins down θ_2 independent of the distribution of types. However, θ_2 is defined by the equation $S_q(k(\theta)) - c_q(k(\theta), \theta) + \frac{1-F(\theta)}{f(\theta)c_{q\theta}(k(\theta),\theta)} = 0$ which depends on $f(\theta_2)$. Hence, slightly perturbing f around θ_2 changes θ_2 but not (33). Consequently, a scoring rule auction cannot implement the optimal mechanism for generic cost and distribution functions.

Second, we analyze the first score auction. To use the same reasoning as above, we have to show the following: In a first score auction implementing the quality and profit schedule of the optimal mechanism, types in the zero profit interval have the same optimal bid. Put differently, we assume that there is a first score auction with score $s(q, p)$ which implements the quality and profit schedule of the optimal mechanism. We then show that all types in a zero profit interval have the same optimal bid. Using the arguments above, this shows that the first score auction does not implement the optimal mechanism.

We denote the profits conditional on winning as $\tilde{\pi}(\theta)$, i.e.

$$\tilde{\pi}(\theta) = max_{p,q} p - c(q, \theta) \qquad s.t.: s(q, p) = bid(\theta).$$

Clearly, the constraint will always be binding (otherwise a firm could get infinite profits). Note that all types in a zero profit interval must have $\tilde{\pi}(\theta) = 0$ which means that the derivative of the Lagrangian

$$L(\theta) = p - c(q, \theta) + \mu(\theta) (s(q, p) - bid(\theta))$$

[23]The reason is that the integrand is negative around θ_1, positive around θ_2 and changes sign only at one type which is between θ_1 and θ_2. This follows from lemma 2.

will equal 0. Using the envelope theorem, we get

$$L_\theta(\theta) = -c_\theta(q(\theta), \theta) - \mu(\theta) bid_\theta(\theta) = 0.$$

Since all types in the zero profit interval have $q(\theta) = k(\theta)$ in the optimal mechanism, the term $-c_\theta(q(\theta), \theta)$ is zero for those types. Since the constraint binds, the Lagrange parameter $\mu(\theta)$ is not zero. Therefore, $bid_\theta(\theta)$ has to be zero which is what we wanted to show. *Q.E.D.*

Proof of proposition 4 For now, assume $VV(\theta_1) \geq VV(\theta_2)$ in the optimal mechanism. We will deal with the opposite case at the end of the proof. It was already shown in the main text that it is optimal for each type to bid the quality that maximizes $S(q) - c(q, \theta) + \Delta(q)$ and the costs the type has at this quality as minimum price. The expression $S(q) - c(q, \theta) + \Delta(q)$ is strictly concave in q and the function Δ was chosen such that the optimal quality schedule solves the first order condition of this maximization problem. Hence, all types find it optimal to bid $(q(\theta), c(q(\theta), \theta))$. By the envelope theorem, $bid_\theta(\theta) = -G'()c_\theta(q(\theta), \theta)$ for $\theta \in [\underline{\theta}, \theta_w]$ and $bid_\theta(\theta) = -c_\theta(q(\theta), \theta)$ for $\theta > \theta_w$. This implies that bids are decreasing on $[\underline{\theta}, \theta_2]$, constant on $[\theta_1, \theta_w]$ and on $(\theta_w, \theta_2]$ and increasing on $[\theta_2, \bar{\theta}]$.

The function G is chosen such that the following two properties are satisfied. First, say there exists a type $\theta' \leq \theta_1$ and a type $\theta'' \geq \theta_2$ such that $VV(\theta') = VV(\theta'')$ in the optimal mechanism (proposition 2). Then G is chosen such that $G\left(S(q(\theta')) - c(q(\theta'), \theta') + \Delta(q(\theta'))\right) = S(q(\theta'')) - c(q(\theta''), \theta'') + \Delta(q(\theta''))$.[24] As $VV_\theta(\theta)$ is negative on $[\underline{\theta}, \theta_1]$ and positive on $(\theta_2, \bar{\theta}]$, G is strictly increasing on the relevant range as required. Second, choose G such that the scores of the optimal bids of types in $(\theta_w, \theta_2]$ according to score B equal the

[24]This will imply that the equilibrium bids of θ' in scoring rule A and θ'' in scoring rule B are equal.

scores of types in $[\theta_1, \theta_w]$ according to score A. As shown above, types in $(\theta_w, \theta_2]$ have the same bid and also types in $[\theta_1, \theta_w]$ have the same bid. Since, scores are decreasing on $[\underline{\theta}, \theta_1]$ and as $VV(\theta_1) \geq VV(\theta_2)$, this property is in line with G being increasing. In fact, G will be discontinuously increasing at the bid of θ_1 if $VV(\theta_1) > VV(\theta_2)$.[25]

Since ties are broken using the virtual valuation of the optimal mechanism (see the main text), the choice of G guarantees that the winner of the auction is indeed the firm chosen under the optimal mechanism.

A firm's expected rent is the same under the dual-score auction with tie breaking and the optimal mechanism. To see this, note that types in $[\theta_1, \theta_2]$ have zero profits in the auction since all such types submit bids leading to the same score while types outside this interval submit bids leading to higher scores. Consequently, a type in $[\theta_1, \theta_2]$ will get a price equal to the minimum price he bid – which was equal to his costs – whenever he wins. Denote the distribution of second highest scores ($score^{(2)}$) in the dual-score auction with tie breaking equilibrium by H and let $p(score^{(2)}, q)$ be the price a winning firm receives when bidding quality q while the second highest score is $score^{(2)}$. Denoting the score associated with the optimal bid of type θ as $score(\theta)$, expected profits for $\theta' > \theta_2$ in the dual-score auction with tie breaking can then be written as

$$\pi^{ds}(\theta') = \int_{score(\theta_2)}^{score(\theta')} p(score^{(2)}, q(\theta')) - c(q(\theta'), \theta')\, dH(score^{(2)}).$$

Since $q(\theta')$ maximizes $p(score^{(2)}, q) - c(q, \theta')$ and as $p(score(\theta'), q(\theta')) = c(q(\theta'), \theta')$, we

[25]Note that a type $\theta_1 + \varepsilon$ will still not want to imitate the much higher bid of a type $\theta_1 - \varepsilon$. While this would increase his chance of winning a lot, he cannot deliver this higher score at non-negative profits, i.e. bidding $(q(\theta_1 + \varepsilon), c(q(\theta_1 + \varepsilon), \theta_1 + \varepsilon))$ gives the highest score type $\theta_1 + \varepsilon$ can deliver at non-negative profits.

get by an envelope argument

$$\pi_\theta^{ds}(\theta') = \int_{score(\theta_2)}^{score(\theta')} -c_\theta(q(\theta'), \theta')\; dH(score^{(2)}) = -X(\theta')c(q(\theta'), \theta').$$

The last equality holds as, for any type vector, the same firm as in the optimal mechanism wins in the dual-score auction with tie breaking. The last equation implies that π_θ is the same in the dual-score auction with tie breaking and the optimal mechanism. A similar derivation holds for $\theta' < \theta_1$ and therefore rents are the same in the dual score auction with tie-breaking and the optimal mechanism. Consequently, the principal's expected payoff is also the same which concludes the proof.

For the case that $VV(\theta_1) < VV(\theta_2)$ in the optimal mechanism, we choose the scoring rules $\tilde{s}_B(q,p) = G(s_B(q,p))$ and $\tilde{s}_A(q,p) = S(q) - p + \Delta(q)$. The same derivation as above goes then through analogously.[26] *Q.E.D.*

[26] G will then be discontinuous at the equilibrium bid of θ_2 instead of the bid of θ_1.

References

Asker, J. and Cantillon, E. “Properties of scoring auctions.” *RAND Journal of Economics*, vol. 39 (2008), pp. 69–85.

Asker, J. and Cantillon, E. “Procurement when price and quality matter.” *RAND Journal of Economics*, vol. 41 (2010), pp. 1–34.

Bagnoli, M. and Bergstrom, T. “Log-concave probability and its applications.” *Economic Theory*, vol. 26 (2005), pp. 445–469.

Branco, F. “The design of multidimensional auctions.” *RAND Journal of Economics*, vol. 28 (1997), pp. 63–81.

Che, Y.K. “Design competition through multidimensional auctions.” *RAND Journal of Economics*, vol. 24 (1993), pp. 668–680.

Council of the European Union. “Sustainable development strategy.” Tech. rep., European Council DOC 10917/06 (2006). `http://register.consilium.europa.eu/pdf/en/06/st10/st10917.en06.pdf`.

DNV. “Pathways to low-carbon coal-fired power generation in europe.” Tech. rep., Abatement potential towards 2020 (2014). `http://www.dnv.com/resources/position_papers/low_carbon_coal-fired_power_generation_paper.asp`.

Ganuza, J.J. and Pechlivanos, L. “Heterogeneity-promoting optimal procurement.” *Economics Letters*, vol. 67 (2000), pp. 105–112.

Johnson, L., Yeh, S. and Hope, C. “The social cost of carbon: implications for modernizing our electricity system.” *Journal of Environmental Studies and Sciences*, vol. 3 (2013), pp. 369–375.

Jullien, B. “Participation constraints in adverse selection models.” *Journal of Economic Theory*, vol. 93 (2000), pp. 1–47.

Laffont, J.J. and Tirole, J. “Auctioning incentive contracts.” *Journal of Political Economy*, vol. 95 (1987), pp. 921–937.

Laffont, J.J. and Tirole, J. *A theory of incentives in procurement and regulation.* MIT Press (1993).

Lewis, T. and Sappington, D. “Countervailing incentives in agency problems.” *Journal of Economic Theory*, vol. 49 (1989), pp. 294–313.

Maggi, G. and Rodriguez-Clare, A. “On countervailing incentives.” *Journal of Economic Theory*, vol. 66 (1995), pp. 238–263.

McAfee, R.P. and Lewis, T.R. *Introduction to Economic Analysis.* http://www.mcafee.cc/Introecon/ (2007).

Myerson, R. “Optimal auction design.” *Mathematics of Operations Research*, vol. 6 (1981), pp. 58–73.

Seierstad, A. and Sydsaeter, K. *Optimal control theory with economic applications*, vol. 24 of *Advanced Textbooks in Economics.* North-Holland Amsterdam (1987).

Steen, M. “Greenhouse gas emissions from greenhouse gas emissions from fossil fuel fired power generation fossil fuel fired power generation.” Tech. rep., European Commission, Joint Research Centre (DG JRC) Institute for Advanced Materials (2015). Http://publications.jrc.ec.europa.eu/repository/bitstream/JRC21207/EUR

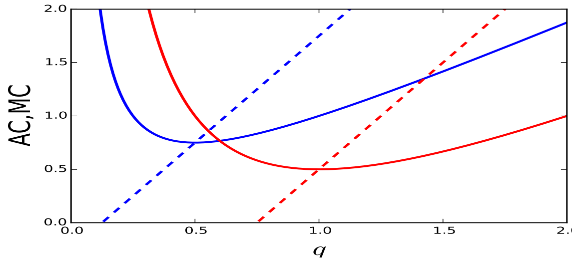
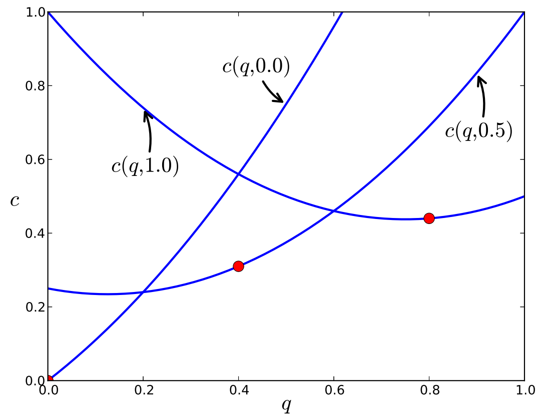
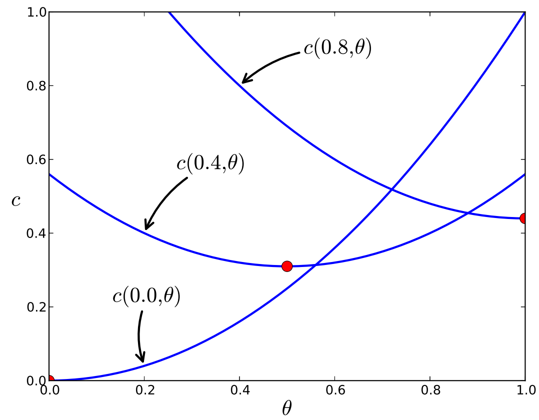


Figure 1: Average costs (AC, solid line) and marginal costs (MC, dashed line) as a function of q for two different firms.



(a) $c(q, \theta)$ for $\theta = 0, 0.5, 1$



(b) $c(q, \theta)$ for $q = 0, 0.4, 0.8$

Figure 2: $c(q, \theta)$ in example 2 both as a function of q (for 3 values of θ) and as a function of θ (for 3 values of q).

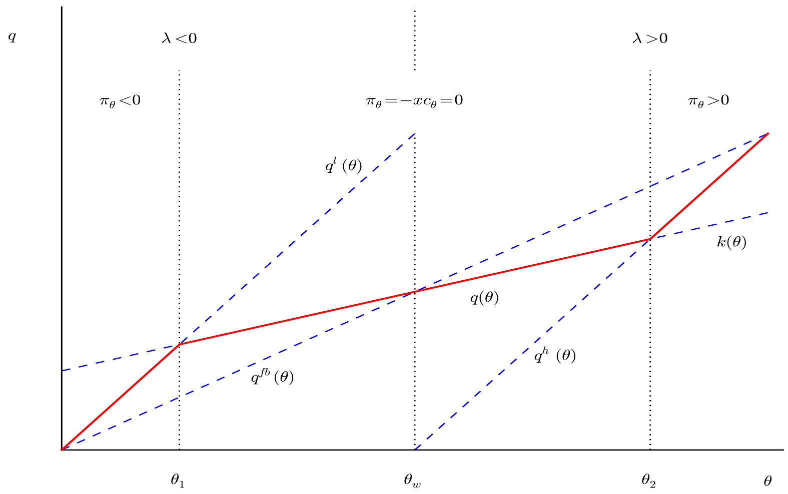


Figure 3: Optimal $q(\theta)$ (solid, red) in the WNM case, together with (dashed) $q^l(\theta), q^{fb}(\theta), k(\theta), q^h(\theta)$.