# Procurement with specialized firms

Jan Boone and Christoph Schottmüller*

June 15, 2013

**Abstract**

This paper analyzes 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 than higher type firms when providing low quality but more expensive when providing high quality.

Specialization limits the informational distortion for the worst types and leads to zero profits for a mass of types ("profit bunching"). If first best welfare is U-shaped in type, the optimal mechanism is not efficient in the sense that types providing a lower second best welfare win against types providing a higher second best welfare. As standard scoring rule auctions cannot implement the optimal mechanism, we introduce a *dual score auction with tie breaking* that implements the optimal mechanism.

**JEL:** D82, H75, L51

**keywords:** procurement, specialization, deregulation, countervailing incentives

*Boone: Department of Economics, University of Tilburg, P.O. Box 90153, 5000 LE, Tilburg, The Netherlands; Tilec, CentER and CEPR. Schottmüller: Department of Economics, University of Copenhagen. We gratefully acknowledge financial support from the Dutch National Science Foundation (VICI 453.07.003).

# 1. Introduction

The literature on procurement and optimal incentive regulation, see for example Laffont and Tirole (1987), Laffont and Tirole (1993) or Che (1993), assumes that firms have private information with regard to their cost functions. As usual in screening models, this private information is represented by a "type" which is assumed to be a scalar. It is then assumed that higher types are better in the sense that higher types have lower costs. If costs, for example, depend on the quality produced, a higher type has lower costs at every quality level.

The private information of a firm is often interpreted as the production technology it uses. This technology was determined in the past and can therefore be treated as given in the context of one specific procurement contract. Following this interpretation, one should expect that firms chose production technologies that are not obviously inferior to alternative technologies, i.e. there should be some level of quality for which the technology of a firm is efficient. Put differently, firms are specialized in the production of a certain quality level. Such specialization is not covered by standard procurement models which assume that higher types have lower costs at every quality level.

To illustrate the concept of specialization, we describe the market for home care in the Netherlands which was recently liberalized. Local governments now procure home care for their citizens and money saved on the procurement can be used freely by local governments, that is, the money received from the central government to pay for home care is not earmarked. However, the local government does have a duty to provide care of some minimum standard. In the past, regional care offices, which did not have substantial incentives to save costs, contracted local monopolists for providing home care. Due to liberalization, new players have entered the market. For example, cleaning companies considered moving into home care. As these new players have no experience with care–they did not use to hire nurses or other professionals with a medical background– they are seen as low quality players. That is, they can provide simple services like house cleaning and shopping more cheaply than traditional home care companies. It is, however, almost impossible, i.e. very costly, for them to provide high quality care because they have no experience with medical care. In this sense, traditional providers are specialized in high quality production while entrants are specialized in low quality production.[1]

This pattern–where incumbents are specialized in high quality service while entrants are specialized in a low quality (low price) service–is typical after liberalization. Many European countries have liberalized sectors like post, taxis, air transport, railway or local transport. This has led to entry by players who offer lower quality in, for instance, the following sense: only make deliveries twice a week (instead of 6 days a week), drive cars substantially cheaper than a Mercedes (see `http://www.tuktukcompany.nl/` for an example), operate planes with reduced seat pitch and limited on board service as well as offering less connections, use old trains and buses to transport people.

Note that in each of the examples above quality is contractible. The planner can verify whether the person visiting a patient is a qualified nurse (instead of a cleaner with no medical qualification), whether mail is delivered 6 days a week, whether a bus or train is only 3 years old, has air-conditioning etc.

The question we ask is: How should a planner (in the home care example: the municipality) organize the procurement when facing specialized firms? In particular, we are interested in how the optimal procurement mechanism differs from the optimal

[1]To a certain extent this can be resolved through market separation in *care* and *support*. People who do not need medical attention but only someone to clean their home, can be served by cleaning companies. While patients who stay at home and need a nurse can be served by the incumbents. Hence, at the extremes of the home care spectrum, market separation can alleviate the issue. However, many cases in home care are not so clear cut. A nurse helping an elderly woman putting on her clothes in the morning and cleaning the house may recognize the first signs of dementia that would be overlooked by an employee of a cleaning company. In such a separated market, our model applies to the *support* segment of the market.

mechanisms in the procurement literature, i.e. when firms are not specialized but simply differ in efficiency. We are also interested whether the optimal mechanism can be implemented by scoring rule auctions.

We show the following results. Think of low (high) type firms as firms specialized in low (high) quality production. First, if low types (e.g. entrants in the examples above) are worse than high types (incumbents in the examples) with respect to first best welfare, the incumbents do not lose from entrants. Second, only if first best welfare first decreases and then increases in type, types specialized in high quality can lose in the following way: A low quality provider (entrant) can win the procurement even though the high quality provider (incumbent) would provide higher welfare if contracted under the optimal procurement rules. We say the optimal mechanism is "second best inefficient". Third, in this latter case, quality is distorted above first best for some types and below first best for others. Fourth, in both cases an interval of types has zero profits ("profit bunching"). Although all types in this interval have zero profits, they produce different qualities when winning the contract. Put differently, a mass of types will have no economic rents under the optimal contract although types are perfectly separated in equilibrium. Fifth, the optimal mechanism cannot be implemented by standard scoring rule auctions. Therefore, we propose a dual score auction with tiebreaking that can implement the optimal mechanism. The last four results are due to the specialization assumption, that is, these results do not occur if firms differ only in efficiency.

Interestingly, our results seem to relate to post liberalization industries which we used as an example before. In the Dutch home care example, firms complained about low profits after liberalization. Some firms even made losses after being contracted. This situation is reminiscent of our zero profit result. Also complaints that liberalization is bad (for welfare) because of a decrease in quality are often heard. Such a complaint only makes sense if a high quality incumbent would have been willing to provide higher welfare than the winning entrant. These complaints might not always reflect the true situation as incumbents could have an incentive to air such claims even if they are not correct. Yet, our result that high quality players providing higher welfare can lose from low quality players providing lower welfare illustrates that incumbents might have a point. Such an ex post inefficiency is, however, part of a mechanism maximizing ex ante expected welfare.

On a technical level, the paper contributes to the literature by solving a two-dimensional mechanism design problem with countervailing incentives. A technical challenge is that local incentive compatibility is not straightforwardly sufficient for non-local incentive compatibility, i.e. non-local incentive constraints have to be checked explicitly. We give sufficient conditions under which non-local incentive constraints are not binding and derive the optimal mechanism in this case.

The set up of the paper 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 while 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 shows how the model extends to situations in which the assumptions of section 3 are not met and section 8 concludes. Proofs are relegated to the appendix.

## 2. Review of the literature

Our paper 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. newly liberalized industries. We show that relaxing it leads to zero economic rents for a mass of types–producing different quality levels–which is, to our knowledge, a new result in the literature on procurement auctions. We also show that implementation of the optimal mechanism by standard auctions, e.g. scoring rule auctions, is no longer straightforward when firms are specialized.

Asker and Cantillon (2010) are an exception in the procurement literature. They analyze a four type model with a linear cost function. Which of the two middle types has lower costs depends on the quality level, i.e. their model includes some partial specialization although this is not the main focus of their paper. Our paper shares some results with Asker and Cantillon (2010), e.g. quality can be upward and downward distorted. We add by (i) analyzing a situation of pure specialization, (ii) using general cost functions , (iii) having a continuum of types and (iv) proposing 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 connects the literature on competitive procurement with the literature on countervailing incentives, see Lewis and Sappington (1989) for the seminal contribution and Jullien (2000) for the most general treatment. By assuming that firms are specialized, our paper uses a cost function that resembles the utility functions of the countervailing incentives literature. Our result that the participation constraint is binding for a mass of types is also typical for this literature. We contribute by allowing for several agents bidding for the contract while the countervailing incentive literature focuses on settings with one principal and one agent. This makes our problem two-dimensional (quality and probability of winning) while the countervailing incentives literature focusses on one-dimensional settings. In this one dimensional setting, local incentive compatibility constraints are sufficient for non-local incentive compatibility and many of the technical challenges encountered in our paper do not occur. From an applied point of view, having more than one firm leads to the result that optimal procurement auctions can be second best inefficient.

As we solve a mechanism design problem with two variables, i.e. quality and the probability of being contracted, our paper is also related to the literature on multi-dimensional screening as surveyed in Rochet and Stole (2003). We contribute here by analyzing a two-dimensional screening model with countervailing incentives. Other screening models with one-dimensional type and multidimensional decisions include, for example, Matthews and Moore (1987) or Guesnerie and Laffont (1984). These papers feature, in contrast to ours, principal agent models with one agent. Furthermore, type denotes efficiency and not specialization in these models.

## 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 $S(q)$ where we normalize quality in such a way that $S(q) = Sq$ for some $S > 0$.[2] The cost of production is denoted by the three times continuously differentiable cost function $c(q, \theta)$ where a firm's type $\theta$ 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 density $f$.

We make the following assumptions on the cost function $c$ and distribution function $F$.

**Assumption 1.** *We assume that*

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

[2] This is, given our assumptions on the cost function, without loss of generality for weakly concave gross values $S(q)$.

- *for* $q \in I\!R_+$ *it is the case that* $S$ *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 increasing and convex in $q$. Higher $\theta$ implies lower marginal costs $c_q$ (the Spence-Mirrlees condition) and $c$ is convex in $\theta$. It will become clear that this convexity is part of the idea of specialized firms: For each quality level, there is one cost minimizing type, i.e. a type specialized in this quality. The second assumption formalizes the idea in our home care application that the government cannot decide not to provide the service. That is, it is always socially desirable for the service to be supplied. The third part is the monotone hazard rate (MHR) assumption. 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) or Jullien (2000).[3] MHR is relaxed in section 7.

The following assumption states that firms are specialized which is the case we want to analyze in this paper.

**Assumption 2.** *For each* $\theta \in [\underline{\theta}, \bar{\theta}]$*, there exists* $k(\theta) > 0$ *such that*

$$c_\theta(q,\theta) \begin{cases} > 0 & \text{if } q < k(\theta) \\ < 0 & \text{if } q > k(\theta) \end{cases}$$

*Further,*

$$c_{q\theta\theta}(q,\theta) \begin{cases} \leq 0 & \text{if } q < k(\theta) \\ \geq 0 & \text{if } q > k(\theta) \end{cases}$$

$$c_{qq\theta}(q,\theta) \begin{cases} \geq 0 & \text{if } q < k(\theta) \\ \leq 0 & \text{if } q > k(\theta) \end{cases}$$

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

For high values of $q$, a higher type $\theta$ 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 $\theta$ types. To illustrate, high $\theta$ incumbents 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, incumbents might have more expensive but also more qualified workers. Replacing these workers is, especially in Europe, costly because of labor market rigidities and search costs.

Consequently, it is more expensive for incumbents to produce low $q$ than for entrants (and the other way around for high $q$). The function $k(\theta)$ is implicitly defined by $c_\theta(k, \theta) = 0$. By assumption 1, $k(\theta)$ is differentiable and monotonically increasing.

In some sense, our assumption that $c_\theta$ switches sign in $q$ follows naturally from the sorting condition $c_{q\theta} < 0$. However, it is the main departure from the existing literature on procurement which assumes $c_\theta < 0$ or equivalently that $k(\theta) \leq 0$ which implies that $c_\theta < 0$ in the relevant domain. Put differently, the existing literature assumes that types can be ranked in terms of efficiency irrespective of $q$. We allow efficiency advantages to depend on $q$ and therefore firms can be specialized in producing a certain quality.[4]

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).

To ensure that (i) the planner's objective function is concave in $q$ and (ii) quality $q$ increases in type, it is standard in the literature to make assumptions on third derivatives $c_{q\theta\theta}, c_{qq\theta}$. Given that (i) and (ii) are satisfied, a first order approach is valid. If $c_\theta$ does not switch sign, the usual assumption is that these derivatives should not switch sign either. This is different in our case where assumption 2 is needed to ensure (i)

[4]If $q$ is interpreted as quantity, we allow firms to be specialized in a certain scale of production.

and (ii). We discuss in section 7 how the solution changes if these assumptions are not satisfied. Note that we allow for the often used linear-quadratic cost functions where these third derivatives equal zero.

As $c_\theta$ can be positive, it is not clear how first best welfare varies with $\theta$. Below we define the two cases that we consider here. In order to do this, we introduce the following notation. First best output is defined as

$$q^{fb}(\theta) = \arg\max_q Sq - c(q, \theta) \tag{1}$$

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

$$W^{fb}(\theta) = Sq^{fb}(\theta) - c(q^{fb}(\theta), \theta). \tag{2}$$

Our final assumption makes sure that we can focus on two relevant cases only.

**Assumption 3.** *Assume that* $c_{q\theta}^2(k(\theta), \theta) > c_{\theta\theta}(k(\theta), \theta)c_{qq}(k(\theta), \theta)$.

We illustrate how the optimal mechanism changes if this assumption is not satisfied in the supplementary material to this paper. The assumption implies that first best welfare is quasiconvex in $\theta$. Hence, we only need to consider two cases. Either first best welfare is monotone in $\theta$ or it is first decreasing and then increasing in $\theta$. Further, we show that $k(\theta)$ intersects $q^{fb}(\theta)$ at (at most) one type.

**Lemma 1.** *First best welfare* $W^{fb}(\theta)$ *is quasiconvex in* $\theta$*. Furthermore,* $q_\theta^{fb}(\theta) > k_\theta(\theta)$ *at any type* $\theta$ *where* $q^{fb}(\theta) = k(\theta)$.

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 1 there are no other cases). The two cases that we focus on in this paper are therefore:

**Definition 1.** *We consider the two cases*

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

**(WNM)** *where a* $\theta_w$ *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* $W^{fb}(\bar{\theta}) > W^{fb}(\underline{\theta})$*.*

The following two examples give cost and surplus functions that correspond to cases (WM) and (WNM) respectively.

**Example 1.** *Assume* $S(q) = q$ *and* $c(q, \theta) = (q - \theta)^2 + q(1 - \theta/2)$ *where* $\theta$ *is distributed uniformly on* $[0, 1]$*. With these functions* $k(\theta) = 4\theta/5$ *and* $q^{fb}(\theta) = 5\theta/4$*. First best welfare is* $W^{fb}(\theta) = \frac{9}{16}\theta^2$ *which is increasing in* $\theta \in [0, 1]$*.*

The interpretation of this example could be that a firm has the "natural quality level" $\theta$ because of the qualification of its current staff. Producing at different qualities involves adjustment costs that increase with the distance $|q - \theta|$. Additionally, there is a linear cost of quality, e.g. from additional (non-staff) input factors. A high type firm, e.g. a firm that traditionally has had highly qualified staff and therefore is experienced in high quality production, has lower marginal costs of quality.

**Example 2.** *Assume* $S(q) = Sq$ *and* $c(q, \theta) = \frac{1}{2}q^2 - \theta q + \theta k$ *with* $k \in (S + \underline{\theta}, S + \bar{\theta})$*. Thus* $k(\theta) = k$ *in assumption 2. Then we find that* $q^{fb}(\theta) = S + \theta$ *and* $dW^{fb}(\theta)/d\theta = S + \theta - k$*. Hence, with* $(k - S) \in (\underline{\theta}, \bar{\theta})$ *first best welfare increases for* $\theta > k - S$ *and decreases for* $\theta < k - S$*.*

The second example reflects the standard idea that a firm with high fixed costs ($\theta k$) 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.

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

Let $t(\theta)$ denote the expected transfer paid by the planner to a firm of type $\theta$ and $x(\theta)$ the probability that type $\theta$ is contracted. That is, the planner offers a menu of choices for firms and each firm chooses the option that maximizes its profits. The planner's objective is to maximize the expected value of $Sq$ minus the expected transfer payments to all firms. The payoff for a type $\theta$ firm that chooses option $(q, x, t)$ is written as $t - xc(q, \theta)$.[5]

Following Myerson (1981), we use a direct revelation mechanism. That is, we design a menu of choices where $(q(\theta), x(\theta), t(\theta))$ is the choice "meant for" type $\theta$. The menu has to be designed such that it is incentive compatible (IC) for type $\theta$ to choose this option. That is, it is IC for a firm to truthfully reveal its type $\theta$.

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

$$\pi(\hat{\theta}, \theta) = t(\hat{\theta}) - x(\hat{\theta})c(q(\hat{\theta}), \theta). \tag{3}$$

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

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

Using an envelope argument, incentive compatibility requires

$$\pi_\theta(\theta) = -x(\theta)c_\theta(q(\theta), \theta). \tag{4}$$

This equation makes sure that the first order condition for truthful revelation of $\theta$ is satisfied. If this equation is satisfied, we say that local IC is satisfied.

It is well known in the literature (Laffont and Tirole, 1993) that local IC implies global IC if (i) single crossing is satisfied and (ii) the decisions ($q(\theta)$ and $x(\theta)$) are

[5]Note that since firms' and planner's utility is quasilinear in money, it is without loss of generality to assume that transfer payments $t$ are paid without conditioning on being contracted: A price $p$ which is paid only when winning the auction is equivalent to an unconditional transfer $t = px$.

monotone. That is, if types do not want to mimic types close to them, they do not want to mimic any type. However, this result does not apply when firms are specialized.

Intuitively, assumption 2 is similar to a violation of single crossing. Viewing a firm's payoff, $t - xc(q, \theta)$ as a function of $x$, the standard single crossing assumption in a one-dimensional model would require that the derivative of $t - xc(q, \theta)$ with respect to $x$ is monotone in type, i.e. single crossing would require that $c_\theta$ does not change sign. But assumption 2 states exactly the opposite.[6] It is well known that in models without single crossing non-local IC can become relevant, see for example Araujo and Moreira (2010) or Schottmüller (2011).

We will first neglect these 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 1–3. We come back to this in section 7 where we relax 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 $c_\theta$ 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

$$\pi(\theta) \geq 0 \tag{5}$$

where we normalize firms' outside option to zero.

**Lemma 2.** *Given the optimal $x(\cdot)$, quality and profits solve the program*

$$\begin{aligned} \max_{q,\pi} \int_{\underline{\theta}}^{\bar{\theta}} & f(\theta)[x(\theta)(Sq(\theta) - c(q(\theta), \theta)) - \pi(\theta)] \qquad (6) \\ + \quad & \lambda(\theta)\left(\pi_\theta(\theta) + x(\theta)c_\theta(q(\theta), \theta)\right) + \eta(\theta)\pi(\theta)d\theta \end{aligned}$$

[6] Although defining single crossing in multidimensional models is not straightforward (see for example McAfee and McMillan (1988)), we refer to $c_\theta$ switching sign as a "violation of single crossing".

*where $\lambda(\cdot)$ and $\eta(\cdot) \geq 0$ are the Lagrange multipliers of the constraints (4) and (5).*

*Quality $q(\cdot)$ is determined by*

$$f(\theta)(S - c_q(q(\theta), \theta)) + \lambda(\theta)c_{q\theta}(q(\theta), \theta) = 0. \tag{7}$$

*The firm with the highest virtual valuation*

$$VV(\theta) = Sq(\theta) - c(q(\theta), \theta) + \frac{\lambda(\theta)}{f(\theta)}c_\theta(q(\theta), \theta) \tag{8}$$

*is contracted.*

The virtual valuation includes next to the first best welfare ($Sq-c$) a rent extraction term. Roughly speaking, contracting a type with a higher probability, i.e. increasing $x(\theta)$, changes the slope of the rent function $\pi(\theta)$; see equation (4). If, for example, $q(\theta) > k(\theta)$, the rent function is increasing more steeply when $x(\theta)$ is increased. Hence, types above $\theta$ will get a higher rent. $\lambda(\theta)$ is the weight of the types that benefit from this higher rent.

The first order condition for $q(\cdot)$ in (6) can be written as (7). As with the virtual valuation, the condition for optimal $q$ includes a first best welfare term ($S - c_q$) and a rent extraction term. The rent extraction term can point in two directions, depending on whether IC is binding downwards or upwards. Therefore, the following notation proves to be useful. Let $q^h(\theta)$ denote the solution to[7]

$$S - c_q(q(\theta), \theta) + \frac{1 - F(\theta)}{f(\theta)}c_{q\theta}(q(\theta), \theta) = 0 \tag{9}$$

and $q^l(\theta)$ the solution to[8]

$$S - c_q(q(\theta), \theta) - \frac{F(\theta)}{f(\theta)}c_{q\theta}(q(\theta), \theta) = 0. \tag{10}$$

Put differently, $q^h$ is the solution to the first order condition (7) when $\lambda(\theta) = 1 - F(\theta)$ and $q^l$ is the solution to (7) if $\lambda(\theta) = -F(\theta)$.

[7]If several $q$ solve this equation, we denote the highest by $q^h$. By assumption 1 and 2, there can be at most one $q > k(\theta)$ satisfying equation (9).

[8]If the solution to this equation is not unique, let the lowest solution be $q^l$. By assumption 1 and 2, there is at most one $q < k(\theta)$ satisfying equation (10).

## 4. First best welfare monotone

We will now characterize the optimal mechanism for the WM-case. There are two cases to consider. In the first case, the solution (given by equation (9)) is such that specialization does not play a role. 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 $\theta_b$ have zero profits (but with different quality levels) and from $\theta_b \geq \underline{\theta}$ onwards, $q(\theta)$ follows $q^h$, see equation (9). In this case, specialization is relevant: all types below $\theta_b$ are assigned the quality they are specialized in. In both cases, IC binds only downwards, i.e. high $\theta$ types would like to mimic low $\theta$ types (not the other way around).

**Proposition 1.** *There are two cases:*

1. *If $c_\theta(q^h(\underline{\theta}), \underline{\theta}) < 0$, then $q^h(\theta)$ in equation (9) gives the optimal quality for all $\theta \in [\underline{\theta}, \bar{\theta}]$. 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 $\theta_b$ is determined by the unique solution to*

$$S - c_q(k(\theta_b), \theta_b) + \frac{1 - F(\theta_b)}{f(\theta_b)} c_{q\theta}(k(\theta_b), \theta_b) = 0. \tag{11}$$

   *For all $\theta > \theta_b$ quality $q(\theta) = q^h(\theta)$. 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.*

We want to give some intuition for why the optimal quality schedule is different when specialization is relevant. In the first case of proposition 1, the possibility that $c_\theta$ can change sign does not play a role in the relevant range of $q$, i.e. $c_\theta$ is negative for all types under the optimal mechanism. In the second case, $c_\theta$ would be positive for some types in the standard quality menu which is given by (9). In the WM case, $c_\theta \leq 0$ at the first best quality level. Hence, the standard downward distortion of $q$ caused by the rent extraction motive is responsible for having $c_\theta > 0$ for some types under $q^h$. By (4), profits are decreasing at types where $c_\theta > 0$. If $q^h$ was implemented, type $\theta^b$ would therefore have zero profits while lower types would have positive profits. But the principal can do better than $q^h$: By assigning $k(\theta)$ to types below $\theta^b$, the principal (i) saves rents as those types remain at zero profits and (ii) reduces distortion compared to $q^h$. Because each type is most cost efficient at his $k(\theta)$, no other type can profitably misrepresent as $\theta$ if $\theta$ expects zero profits and produces quality $k(\theta)$. Put differently, the incentive constraint is slack for types below $\theta_b$. Therefore, it is not necessary to distort quality further down than $k(\theta)$ for rent extraction purposes. In this sense, specialization leads to "less distortion at the bottom" and more rent extraction.

## 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 he might have positive profits under the optimal mechanism. One can think of the WNM case as having two standard menus. One for lower $\theta$ in which lower types are better, the incenitve constraint is upward binding, profits are decreasing in type and quality is distorted upwards. The other for higher $\theta$ with higher types being better, profits increasing in type, the incentive constraint downward binding and quality distorted downwards.

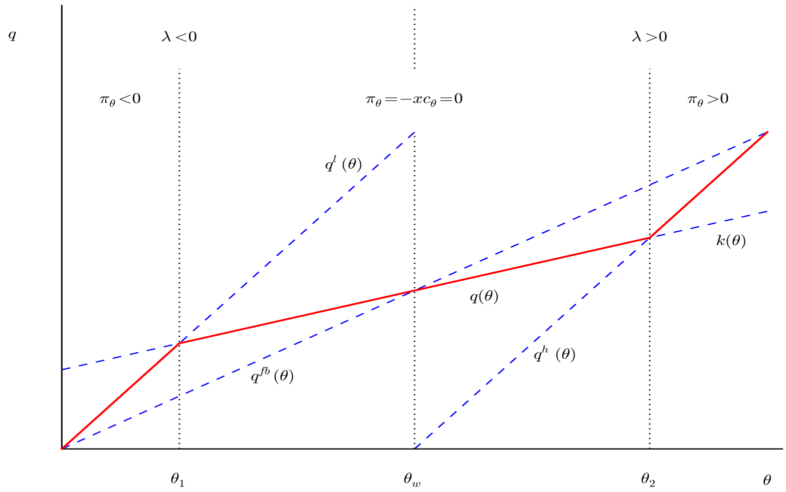


Figure 1: Optimal $q(\theta)$ (solid, red) in the WNM case, together with (dashed) $q^l(\theta), q^{fb}(\theta), k(\theta), q^h(\theta)$.

The way to connect these two standard menus is an interval of types with zero profits (but differing quality levels).[9] Incentive compatibility within the zero profit interval is no problem here: Each zero profit type $\theta$ will produce at 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.** *There exist $\theta_1$ and $\theta_2$, with $\theta_1 < \theta_2$, such that $\theta_1$ and $\theta_2$ are uniquely defined by $q^l(\theta_1) = k(\theta_1)$ and $q^h(\theta_2) = k(\theta_2)$. Quality is determined by*

$$q(\theta) = \begin{cases} q^h(\theta) & \textit{for all } \theta > \theta_2 \\ k(\theta) & \textit{for all } \theta \in [\theta_1, \theta_2] \\ q^l(\theta) & \textit{for all } \theta < \theta_1. \end{cases} \tag{12}$$

[9]Note that this is different from the usual way in which types get bunched. For instance, in Guesnerie and Laffont (1984) or Fudenberg and Tirole (1991, ch. 7) types are bunched on the same $q$ (in our notation) and profits will generally differ between types.

*We have*

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

*Type $\theta_w$, 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*

$$x_\theta(\theta) < 0 \text{ for all } \theta < \theta_w$$

$$x_\theta(\theta) > 0 \text{ for all } \theta > \theta_w.$$

*The relaxed solution is globally incentive compatible.*

Figure 1 illustrates proposition 2. Quality is above first best, i.e. upwards distorted, for low $\theta$ and downwards distorted for high $\theta$. This is a consequence of the U-shaped first best welfare which implies that low types are better around $\underline{\theta}$ and high types are better around $\bar{\theta}$. Quality is not distorted at the (locally) best types $\underline{\theta}$ and $\bar{\theta}$ which resembles the well known "no distortion at the top" result. Quality is also undistorted for the worst type $\theta_w$ which allows a continuous transition from upwards to downwards distortion. As in the WM case, specialization limits the distortion for the worst types.

The boundaries of the zero profit interval $[\theta_1, \theta_2]$ are at those types where the low standard menu ($q^l$) and the high standard menu ($q^h$) intersect $k(\theta)$. In the zero profit interval, each type produces the quality for which he is the cost minimizing type, i.e. $k(\theta)$. Any other quality could not be incentive compatible within a zero profit interval as either types slightly higher or slightly lower would be more efficient. Consequently, they could achieve positive profits by marginally misrepresenting. From $q(\theta) = k(\theta)$, it is evident that misrepresenting as any other type $\hat{\theta} \in [\theta_1, \theta_2]$ cannot be profitable and this is exactly the reason why the zero profit types do not receive any informational rent.

At $\theta_1$ and $\theta_2$, $q(\theta)$ is kinked. At $\theta_1$, for example, the quality according to the standard low menu ($q^l$) would include additional informational distortion pushing quality upwards. Therefore $q^l(\theta) > k(\theta)$ for types slightly above $\theta_1$ while $q(\theta) = k(\theta)$ is necessary to stay in the zero profit interval.

Note that for types above $\theta_w$ the optimal contract is similar to the one derived in proposition 1, i.e. quality and virtual valuation are the same. This is quite intuitive as first best welfare is increasing for those types. In this sense, proposition 2 "extends" proposition 1.

The following proposition formalizes the "grudge" of high $\theta$ (incumbents in our home care example) against low $\theta$ (entrants): although in second best the incumbent generates higher quality and higher welfare than the entrant, it can happen that the entrant wins the procurement contract. Incidentally, the opposite can happen as well: an incumbent wins from an entrant who generates higher (second best) welfare.

**Proposition 3.** *The optimal allocation is not second best efficient in the sense that there exist types $\theta', \theta''$ such that $\theta'$ wins against $\theta''$ while $W^{sb}(\theta'') > W^{sb}(\theta')$.*[10]

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. This decreases the rents A will get by stimulating him to bid more aggressively. In our case, there is only one distribution from which types are drawn. Nevertheless, the intuition is similar. The reason for

[10]We use the term *second best efficient* to describe a situation where the selection rule picks the firm providing the highest $W^{sb}$. $W^{sb}$ is welfare under the optimal quality schedule derived in propositions 1 and 2. A firm *wins* if it is contracted, i.e. has the highest $VV$ of all firms.

discrimination are informational distortions. For the lower standard menu, the relevant term inducing distortion in the virtual valuation is $-F(\cdot)c_\theta(\cdot)$. For high $\theta$, the respective term is $(1-F(\cdot))c_\theta(\cdot)$. While discrimination in Myerson (1981) results from the fact that different distributions govern the distortion, discrimination in our model is due to different parts of the same distribution governing distortion: For low $\theta$, the left tail is relevant and for high types the right tail of the distribution matters for distortion. The reason is that the local incentive constraint is upward binding in the lower standard menu and downward binding in the upper standard menu. On a more intuitive level, by ex ante committing to let a worse low type $\theta' < \theta^w$ win against a better high type $\theta'' > \theta^w$, one can save informational rents for $\theta''$ and all types above him. The reason is that the probability that $\theta''$ wins the auction, i.e. $x(\theta'')$, decreases and therefore the slope of the rent function $\pi_\theta(\theta'') = -x(\theta'')c_\theta(q(\theta''), \theta'')$ decreases. Loosely speaking, one stimulates $\theta''$ and higher types to bid more aggressively.

## 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 in the auction.

Scoring rule auctions are used in pratice 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.[11] Furthermore, Che (1993) shows that the optimal mechanism in a standard procurement model is implementable through a quasilinear second score auction. We will show that this result does not 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 $\theta$ will therefore have the bid

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

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

$$bid_\theta(\theta) = s_p(q(\theta), c(q(\theta),\theta))c_\theta(q(\theta),\theta). \tag{13}$$

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 profits 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 have different virtual valuations and therefore different probabilities of being contracted.

In the appendix, we show that a similar reasoning also holds true in first score auctions which leads to the following result.

[11] The guidelines allow for one alternative to the concept of best economic value: the criterion of lowest price. Such a focus on price is clearly not optimal when quality matters.

**Proposition 4.** *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 do not work whenever specialization matters. We propose an alternative auction mechanism to implement the optimal mechanism. Our *dual-score auction with tie breaking* works in the following way. The principal sets up two scoring rules A and B (and an additional rule translating qualities into prices to be used in the tie breaking round; see below). Each firm bids a score and indicates which of the two scoring rules it wants to use.

As suggested by equation (13), a scoring rule auction can only work if it differentiates correctly between bidders bidding the same score. Assume there are several bidders with the same score and this score is actually the highest one. The auction then proceeds to a tie-breaking round. Firms that chose scoring rule A in the first round will bid a price in round 2 (price auction). Firms that chose scoring rule B in the first round will bid a quality in the second round (beauty contest). Whether the winner of the tie breaking round is the lowest bidder in the price auction or the highest bidder in the beauty contest is then determined by a prespecified rule that translates quality bids into price bids.[12] This rule is chosen such that the firm with the highest VV (8) wins.

For simplicity, we use a second score version of the auction above. The winner has to provide the score of the second highest bidder. The winner's score is measured according to the scoring rule he chose. If a tie breaking round was necessary, the winner still has to provide the score of the second highest bidder (which is now equal to the score he and others bid in the first round) but is also committed to the price or quality he bid in the tie breaking round.

The idea behind the dual-score auction with tie breaking is to divide types into two

[12]If two or more firms are tied in the tie breaking round, the winner is chosen randomly among the tied firms.

groups: Types below $\theta_w$ and types above $\theta_w$. The two scoring rules are designed such that–in equilibrium–all types in $[\underline{\theta}, \theta_w]$ choose scoring rule A and all other types choose scoring rule B. For now, we concentrate on types in $[\underline{\theta}, \theta_w]$ only. Let scoring rule A be $s_A(q,p) = Sq - p + \Delta_A(q)$ where

$$\Delta_A(q) = \int_{q(\underline{\theta})}^{q} \frac{\lambda(q^{-1}(s))}{f(q^{-1}(s))} c_{q\theta}(s, q^{-1}(s))\, ds \qquad \text{for } q \in [q(\underline{\theta}), q(\theta_w)] \tag{14}$$

and $\Delta_A(q) = -\infty$ for $q \notin [q(\underline{\theta}), q(\theta_w)]$ where $q$ is the optimal quality according to proposition 2 and $q^{-1}$ is the inverse function of the optimal quality schedule (mapping into $[\underline{\theta}, \theta_w]$). Bidding the highest score one can provide at non-negative profits is a dominant strategy in the second score auction. Conditional on winning, a type $\theta$ firm maximizes profits by providing the quality that maximizes $Sq - c(q,\theta) + \Delta_A(q)$. The scoring rule is such that a type $\theta$ firm chooses the quality $q(\theta)$ of the optimal mechanism. Furthermore, (13) implies that bids are decreasing on $[\underline{\theta}, \theta_1)$ and constant on $[\theta_1, \theta_w]$.

A price auction tie breaking round will, therefore, only be necessary if all firms (with types below $\theta_w$) have types in $[\theta_1, \theta_w]$. Indeed, a type $\theta \in [\underline{\theta}, \theta_1)$ wins against any type $\theta \in [\theta_1, \theta_w]$. In this case, each firm knows in the tiebreaking round that it can procure its own first round score with non-negative profits only by providing $(q,p) = (k(\theta), c(k(\theta), \theta))$. Therefore, it is a dominant strategy to bid $c(k(\theta), \theta)$ in the tie breaking round. As $c(k(\theta), \theta)$ is increasing in $\theta$ for $\theta \in [\theta_1, \theta_w]$, the winner of the auction is the firm selected by the optimal mechanism.

A similar scoring rule B can be set up for firms with types in $(\theta_w, \bar{\theta}]$. We show in the proof of proposition 5 that the scoring rules A and B can be chosen such that (i) a firm selects scoring rule B if and only if $\theta > \theta_w$, (ii) the scores of the scoring rules are comparable in the sense that the firm with the highest equilibrium bid in the first round has the highest virtual valuation in the optimal mechanism and (iii) the expected payoff of the principal is the same as in the optimal mechanism. This gives the following proposition.

**Proposition 5.** *The optimal mechanism can be implemented by a dual-score auction with tie breaking.*

## 7. Global incentive compatibility

Above we made assumptions on third derivatives of the cost function and the distribution of $\theta$ for ease of exposition. This allowed us to use a first order approach and ignore global IC constraints. In this section, we introduce the global IC constraint. Then we relax the assumptions above such that the solution of the relaxed problem is not necessarily incentive compatible. We show an ironing procedure that can deal with violations of second order incentive compatibility. Finally, we present a family of cost functions for which the first and second order condition for IC (which are both local) imply global IC.

A menu $q(\cdot), x(\cdot), t(\cdot)$ is IC in a global sense if and only if

$$\Phi(\hat{\theta}, \theta) \equiv \pi(\theta, \theta) - \pi(\hat{\theta}, \theta) \geq 0 \tag{15}$$

for all $\theta, \hat{\theta} \in [\underline{\theta}, \bar{\theta}]$; where $\pi(\hat{\theta}, \theta)$ is defined in equation (3).

Equation (4) gives the first order condition for IC. This is not sufficient for (even) a local maximum. The relevant second order condition can be written as follows.

**Lemma 3.** *Second order incentive compatibility requires*

$$x_\theta(\theta)c_\theta(q(\theta), \theta) + x(\theta)c_{q\theta}(q(\theta), \theta)q_\theta(\theta) \leq 0. \tag{SOC}$$

As shown in textbooks like Laffont and Tirole (1993), first and second order conditions for IC imply global IC (as in equation (15)) if $c_\theta < 0$ for all $q \in \mathbb{R}_+$. Because we assume that firms are specialized (assumption 2), local IC does not automatically imply global IC. Hence, we still need to verify global IC even if (4) and (SOC) are satisfied.

Lemmas 4 and 5 in the appendix show that the solutions in propositions 1 and 2 respectively satisfy the global IC constraint (15) (as well as (SOC)) given our assumptions

1–3.

How should the solution to the relaxed problem be adapted if it is not globally IC because our assumptions are not satisfied? For concreteness, we focus here on the WM case and assume that the problems arise because of a violation of the MHR assumption. The cases where third derivatives cause problems with (SOC) are dealt with analogously. In the WM case, the change in $q$ for $\theta > \theta_b$ is given by

$$q_\theta(\theta) = \frac{c_{q\theta}(q(\theta),\theta) - c_{q\theta\theta}(q(\theta),\theta)\frac{1-F(\theta)}{f(\theta)} - c_{q\theta}(q(\theta),\theta)\frac{d\left(\frac{1-F(\theta)}{f(\theta)}\right)}{d\theta}}{-c_{qq}(q(\theta),\theta) + c_{qq\theta}(q(\theta),\theta)\frac{1-F(\theta)}{f(\theta)}}. \tag{16}$$

Now we consider the case where $d((1-F(\theta))/f(\theta))/d\theta > 0$ for $\theta > \theta_b$ in such a way that $q_\theta < 0$ and such that $q_\theta < 0$ causes a violation of (SOC). We first sketch how this is dealt with in general. Then we work out an example.

In the one-dimensional case, a violation of (SOC) is dealt with by bunching types on one quality level $q$; see Guesnerie and Laffont (1984). In the two-dimensional case ($q$ and $x$), it is not necessarily true that a violation of (SOC) leads to bunching of types $\theta$ on the same quality $q$ and probability of winning $x$. Below we do not work with $x$ but with the virtual valuation $VV$ as there is a one-to-one relation between the two (i.e. higher $VV$ implies higher $x$ and the other way around).

Now, we explicitly add constraint (SOC) to the planner's optimization problem (6):

$$\begin{aligned} &\max \int_{\underline{\theta}}^{\bar{\theta}} f(\theta)[x(\theta)(Sq(\theta) - c(q(\theta),\theta)) - \pi(\theta)] \\ + \quad & \lambda(\theta)(\pi_\theta(\theta) + x(\theta)c_\theta(q(\theta),\theta)) \\ - \quad & \mu(\theta)(x_\theta(\theta)c_\theta(q(\theta),\theta) + x(\theta)c_{q\theta}(q(\theta),\theta)q_\theta(\theta)) \\ + \quad & \eta(\theta)\pi(\theta)d\theta \end{aligned} \tag{17}$$

where $\mu(\cdot) \geq 0$ is the Lagrange multiplier (co-state variable) of constraint (SOC). The Euler equation for $q$ can now be written as

$$f(\theta)(S - c_q(q(\theta),\theta)) + \lambda(\theta)c_{q\theta}(q(\theta),\theta) + \mu(\theta)c_{q\theta\theta}(q(\theta),\theta) = -\mu_\theta(\theta)c_{q\theta}(q(\theta),\theta). \tag{18}$$

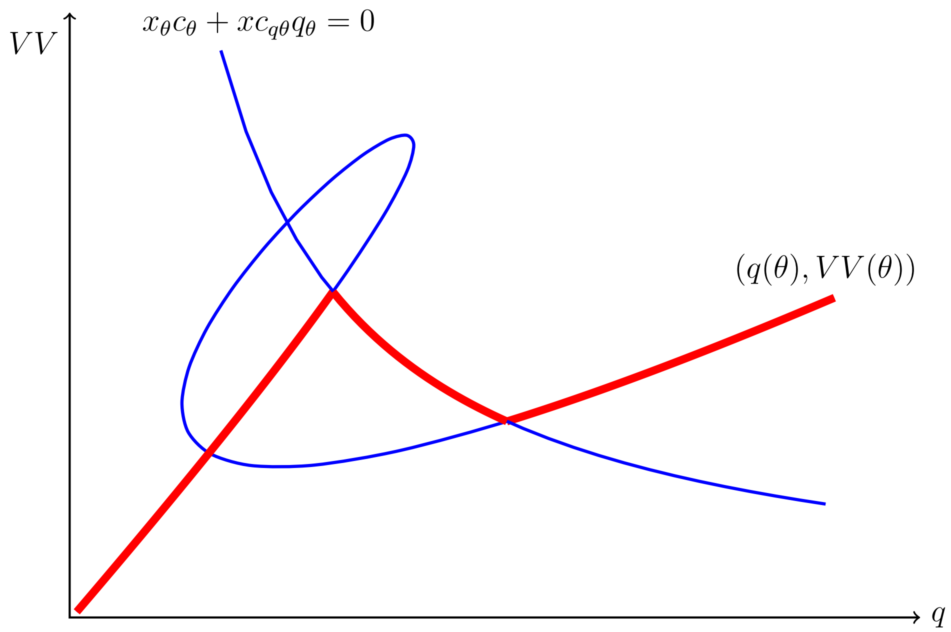


Figure 2: Solution for quality $q(\theta)$ and virtual valuation $VV(\theta)$ for the case where (second order) condition (SOC) is violated.

Consider figure 2 to illustrate the procedure. This figure shows equation (SOC) (where it holds with equality) in $(q, VV)$ space and the solution $(q(\theta), VV(\theta))$ that follows from the planner's optimization problem while ignoring the second order condition; i.e. assuming $\mu_\theta(\theta) = 0$ for all $\theta$. The former curve is downward sloping in the WM case since

$$\frac{dx}{dq} = \frac{x_\theta(\theta)}{q_\theta(\theta)} = -x(\theta)\frac{c_{q\theta}(q(\theta), \theta)}{c_\theta(q(\theta), \theta)} < 0.$$

In the simple case (that we also use in the example below) where $c_{\theta\theta} = 0$, this curve boils down to

$$x(\theta)c_\theta(q(\theta), \theta) = -K < 0 \tag{19}$$

for some constant $K > 0$, as differentiating equation (19) with respect to $\theta$ indeed gives the constraint $x_\theta c_\theta + x c_{q\theta} q_\theta = 0$.

The solution of the relaxed program $(q(\theta), VV(\theta))$ (ignoring the second order constraint!), starts at $\underline{\theta}$ in the bottom left corner and moves first over the thick (red) part of this curve, then follows the thin (blue) part, curving back (i.e. both $q$ and $x$ fall with

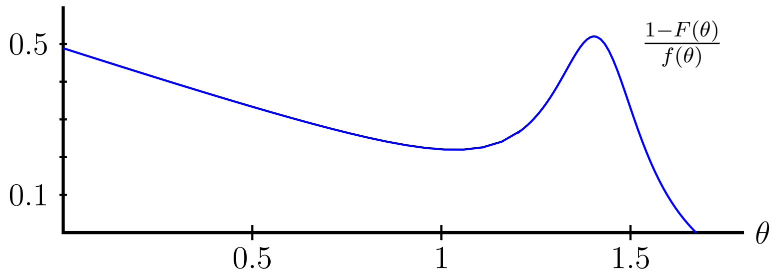


Figure 3: Inverse hazard rate with $f(\theta) = (\theta - a)^2 + 1/50$

$\theta$) then both $q$ and $x$ increase again with $\theta$ and we end on the thick (red) part of the curve. The part of the curve where $q_\theta, x_\theta < 0$ violates equation (SOC).

Hence, we need to find $\theta_a, \theta_b$ where (SOC) starts to bind and $\mu(\theta) > 0$. Then from $\theta_a$ onwards, we follow the binding constraint till we arrive at $\theta_b$, from which point onwards we follow the solution $(q(\theta), VV(\theta))$ again. As shown in figure 2, the choice of $\theta_a$ determines both the trajectory $(\tilde{q}(\theta), \tilde{VV}(\theta))$ satisfying equation (SOC) and the end point of this trajectory $\theta_b$. Since $\mu(\theta) = 0$ both for $\theta < \theta_a$ and for $\theta > \theta_b$, it must be the case that $\int_{\theta_a}^{\theta_b} \mu_\theta(\theta) d\theta = 0$. To illustrate, for the case where $c_{q\theta\theta} = 0$,[13] this can be written as (using equation (18))

$$\int_{\theta_a}^{\theta_b} \frac{f(\theta)(S - c_q(\tilde{q}(\theta), \theta)) + (1 - F(\theta)) c_{q\theta}(\tilde{q}(\theta), \theta)}{c_{q\theta}(\tilde{q}(\theta), \theta)} d\theta = 0. \tag{20}$$

We now illustrate this approach with an example.

**Example 3.** *To violate the monotone hazard rate assumption we use the density* $f(\theta) = (\theta - a)^2 + 1/50$ *with support* $[0, a + 1/4]$ *where* $a$ *has to be approximately* $1.42$ *to satisfy the requirements of a probability distribution. The hazard rate of this distribution is depicted in figure 3.*

*Assume that there are two firms and that* $c(q, \theta) = \frac{1}{2}q^2 - q\theta + \theta$. *Then* $c_\theta(q, \theta) = 1 - q$ *which changes sign at* $q = 1$. *As* $c_{\theta\theta} = 0$, *the binding second order condition takes the*

[13] If $c_{q\theta\theta} \neq 0$, the differential equation (18) has to be solved for $\mu(\theta)$. Although a bit tedious, this is do-able since the differential equation is linear and first order in $\mu(\theta)$.

*form of (19):*

$$x = \frac{K}{q-1}$$

*for some $K > 0$. Note that this equation does not depend on $\theta$. Hence, in this case, "following the constraint" takes the form of bunching $\theta \in [\theta_a, \theta_b]$ on some point*

$$(\tilde{q}, \tilde{VV}) \tag{21}$$

*where $\tilde{VV}$ corresponds to the probability $\tilde{x} = \frac{K}{\tilde{q}-1}$. Choosing $\theta_a$, fixes $\tilde{q} = q(\theta_a)$ and $\theta_b$ since $q(\theta_b) = \tilde{q}$. Writing the dependency of $\tilde{q}, \theta_b$ on $\theta_a$ explicitly, $\theta_a$ solves equation (20):*

$$\int_{\theta_a}^{\theta_b(\theta_a)} f(\theta)(S - (\tilde{q}(\theta_a) - \theta)) - (1 - F(\theta))d\theta = 0. \tag{22}$$

*Since equation (SOC) will already start to bind for $\theta_a$ where $q_\theta(\theta_a) > 0$, it is routine to verify that this equation is downward sloping in $\theta_a$. The unique solution in this example is $\theta_a \approx 1.1685$ which gives a corresponding $\theta_b = 1.428$ and $\tilde{q} = 1.923$.*

While the ironing procedure described above takes care of the local second order condition (SOC), this does not necessarily imply global incentive compatibility. Global constraints are mathematically intractable in general frameworks; see Araujo and Moreira (2010) and Schottmüller (2011) for special examples of how to handle global constraints in a one-dimensional principal agent setup. However, the following proposition establishes that global constraints do not bind for a family of cost functions. This family includes the functions we used in the example and the most commonly used linear-quadratic cost functions.

**Proposition 6.** *If $c_{\theta\theta} = 0$ and the local second order condition (SOC) is satisfied, the solution is globally incentive compatible.*

# 8. 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 many post liberalization situations, incumbents seem to be good at producing high quality while entrants can produce low quality at very low costs. A similar pattern emerges if there are gains from specialization and firms can specialize in either high quality or low costs.

Standard procurement models do not account for this possibility because “type” denotes efficiency and not specialization. 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, specialization limits distortion and helps the principal to extract rents.

If we assume that first best welfare is U-shaped, e.g. there are gains from specializing 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. Contrary to this literature, bidders are drawn from the same distribution in our model. The intuition is that the incentive constraint is first upward and then, for higher types, downward binding. Therefore, different tails of the distribution govern the distortion for low and high types. The commitment to favor some worse types allows the principal to reduce the rents of the best types. Loosely speaking, the better types are incentivized to bid more aggressively. Put differently, competitive pressure can be exerted even by firms that are clearly worse. Further, in this case "gold plating" can be optimal in the sense that some types produce quality levels above their first best levels.

Specialization implies that the optimal mechanism cannot be implemented with a standard scoring rule auction. Instead the procurement agency offers two scoring rules and a firm can decide according to which rule it wants to be evaluated when submitting the bid. To break ties, a second round auction is used where firms bid either quality or price depending on which scoring rule they picked in the first round.

# 9. Appendix

**Proof of lemma 1** From the first order condition for $q^{fb}$, we derive that

$$q_\theta^{fb} = \frac{-c_{q\theta}(q^{fb}(\theta),\theta)}{c_{qq}(q^{fb}(\theta),\theta)} > 0.$$

It follows from $c_\theta(k(\theta),\theta) \equiv 0$ that $c_{q\theta}(k(\theta),\theta)k_\theta(\theta) + c_{\theta\theta}(k(\theta),\theta) = 0$. Hence, $q_\theta^{fb}(\theta) > k_\theta(\theta)$ at a type where $q^{fb}(\theta) = k(\theta)$ if and only if

$$\frac{-c_{q\theta}(q^{fb}(\theta),\theta)}{c_{qq}(q^{fb}(\theta),\theta)} > \frac{c_{\theta\theta}(k(\theta),\theta)}{-c_{q\theta}(k(\theta),\theta)}$$

which holds by assumption 3. Hence, $q^{fb}$ can intersect $k$ at at most one type and only from below. As $W^{fb}(\theta) = -c_\theta(q^{fb}(\theta),\theta)$, 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 assumption 2. This implies quasiconvexity. *Q.E.D.*

**Proof of lemma 2** In principle, the planner does not have to treat all firms symmetrically, e.g. firm 1 could have a higher probability of being contracted when being type $\theta'$ than firm 2 when being type $\theta'$. Given the symmetry of our setup, such asymmetries appear unnatural. Indeed, we show that the symmetric solution is optimal.

The planner maximizes the expected utiltiy $Sq$ minus the transfer paid to firms. If the planner assigns the project to player $i$ with probability $x^i$ where $i$ produces quality $q^i$ and receives transfer $t^i$, the planner's utility from $i$ can be written as $x^i Sq^i - t^i = x^i(Sq^i - c^i) - \pi^i$. The planner's optimization problem including the firm identifier $i$ is

$$\begin{aligned} &\max_{q^i,x^i,\pi^i} \int_{\underline{\theta}}^{\bar{\theta}} \dots \int_{\underline{\theta}}^{\bar{\theta}} \sum_{i=1}^{n} f(\theta^1)\dots f(\theta^n)/f(\theta^i)\left\{f(\theta^i)[x^i(\Theta)(Sq^i(\Theta) - c(q^i(\Theta),\theta^i)) - \pi^i(\Theta)]\right. \\ &+ \lambda^i(\theta^i)(\pi^i_{\theta^i}(\Theta) + x^i(\Theta)c_{\theta^i}(q^i(\Theta),\theta^i)) \qquad (23) \\ &+ \left.\eta^i(\theta^i)\pi^i(\Theta)\right\} + \sigma(\Theta)\left(1 - \sum_i x^i(\Theta)\right) d\theta_1\dots d\theta_n \end{aligned}$$

where $\lambda^i(\cdot)$ and $\eta^i(\cdot) \geq 0$ are the Lagrange multipliers (co-state variables) of the constraints (4) and (5). Here, $x^i(\Theta)$ denotes the probability of firm $i$ being contracted when types are $\Theta = (\theta^1 \dots \theta^n)$. The last constraint ensures that probabilities sum to no more than 1. Because of assumption 1, this constraint will bind and $\sigma(\Theta)$ will therefore be positive. As the objective function is linear in $x^i(\cdot)$ and each $x^i(\Theta)$ has to be nonnegative, we get what is called a "bang-bang" solution in optimal control theory: For any $\Theta$, the firm $i$ such that the derivative of the integrand with respect $x^i$ is highest will be contracted, i.e. $x^i(\Theta) = 1$, while the other firms are not, i.e. $x^j(\Theta) = 0$ for all $j \neq i$.

The Euler equations for $\pi^i$ and $q^i$ are

$$\begin{aligned}
\lambda^i_{\theta^i}(\theta^i) &= -f(\theta^i) + \eta^i(\theta^i) \\
0 &= f(\theta^i)\left(S - c_q(q^i(\Theta), \theta^i)\right) + \lambda^i(\theta^i) c_{q\theta}(q^i(\Theta), \theta^i).
\end{aligned}$$

Note that the maximization over $q^i$ does not depend on $\theta^j$ for $j \neq i$. Therefore, $q^i$ is a function of $\theta^i$ only and we can write $q^i(\theta^i)$ instead of $q^i(\Theta)$. The derivative of the integrand with respect $x^i$ is $VV(\Theta) - \sigma(\Theta)$, where $VV$ only depends on $\theta^i$ (because $q^i$ only depends on $\theta^i$). Hence the planner chooses the firm with the highest $VV(\theta^i)$ and the second result in the lemma follows.

Furthermore, the Euler equations (in fact the optimization problems) for all firms are the same. Hence, $q^i(\theta^i)$ and $\pi^i(\theta^i)$ are the same functions for all $i$ and we can write $q(\theta)$ and $\pi(\theta)$ without the firm identifier. That is, the optimal mechanism treats all firms symmetric: $q(\theta)$ and $x(\theta)$ are the same for all firms.

With this notation and given the optimal $x(\cdot)$, the optimization problem over $q$ and $\pi$ is the one in the lemma. *Q.E.D.*

**Proof of proposition 1** We verify that the conditions of lemma 2 are satisfied by the proposed solution using the sufficient conditions in Seierstad and Sydsaeter (1987, Thm. 1, ch. 5.2).[14] We immediately turn to case 2 of the proposition as the proof of case 1 resembles the proof of case 2 for types $\theta > \theta_b$.

$\lambda(\theta)$ in the proposed solution is $1 - F(\theta)$ for $\theta \geq \theta_b$. $\lambda(\theta)$ is implicitly (and uniquely) defined by (7) and $q(\theta) = k(\theta)$ for $\theta < \theta_b$. Hence, $q(\theta)$ satisfies the first order condition

[14] A more extensive step by step derivation can be found in an earlier working paper version.

of the maximization problem (6). The second order condition is

$$-f(\theta)c_{qq}(q(\theta),\theta)+\lambda(\theta)c_{qq\theta}(q(\theta),\theta)<0$$

which is also satisfied: The first term is negative. The second term is negative for $\theta>\theta_b$ and 0 for $\theta\leq\theta_b$ by assumption 2.

Before continuing to show optimality, we need one intermediate result. Note that the left hand side of (11) is increasing in $\theta_b$ as–by assumption 2–its derivative can be written as

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

where the inequality follows by (MHR) and assumption 3. Hence, $\theta_b$ is uniquely defined by (11). This and the definition of the WM case imply $q^h(\theta)<k(\theta)=q(\theta)<q^{fb}$ for types $\theta<\theta_b$. Therefore, $\lambda(\theta)\in(0,1-F(\theta))$ for $\theta<\theta_b$.

To show that the proposed solution maximizes (6), the Euler equation for $\pi$, i.e.

$$\lambda_\theta=-f(\theta)+\eta(\theta),$$

and the condition $\eta(\theta)\geq 0$ have to be satisfied as well. The function $\eta(\theta)$ in the proposed solution is zero for $\theta\geq\theta_b$ and given by $\eta(\theta)=f(\theta)+\lambda_\theta(\theta)$ for $\theta<\theta_b$. Hence, the Euler equation is satisfied by definition and we only need to check $\eta(\theta)\geq 0$ for types $\theta<\theta_b$. Differentiating (7) with $q(\theta)=k(\theta)$, yields–after plugging in $k_\theta=c_{\theta\theta}/c_{q\theta}$ and using assumption 2–for types $\theta<\theta_b$

$$c_{q\theta}(k(\theta),\theta)\left(-1+\frac{c_{qq}(k(\theta),\theta)c_{\theta\theta}(k(\theta),\theta)}{c_{q\theta}^2(k(\theta),\theta)}+\frac{d\frac{\lambda(\theta)}{f(\theta)}}{d\theta}\right)=0. \tag{24}$$

By assumption 3, the sum of the first two terms in brackets is non-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 $\lambda(\theta)\in(0,1-F(\theta))$.

It remains to show the monotonicity results in proposition 1. Choosing transfers such that $\pi(\underline{\theta}) = 0$ gives together with $\pi_\theta = -x(\theta)c_\theta(q(\theta), \theta)$ the results for $\pi$. 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 (MHR) and assumption 2. The virtual valuation is increasing in type as

$$\frac{dVV}{d\theta} = -c_\theta(q(\theta), \theta)\left(1 - \frac{d\frac{\lambda(\theta)}{f(\theta)}}{d\theta}\right) + c_{\theta\theta}(q(\theta), \theta)\frac{\lambda(\theta)}{f(\theta)} \geq 0 \tag{25}$$

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.*

**Proof of proposition 2** Type $\theta_w$ is determined by the intersection of $q^{fb}$ and $k$ which is unique by lemma 1. We have to show that $\theta_1 < \theta_w < \theta_2$. Assumption 2 and 3 and (MHR) imply that the left hand sides of (9) and (10) are both increasing in $\theta$ if $q(\theta) = k(\theta)$. Hence, $\theta_1$ and $\theta_2$ are unique. As

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

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

The optimal contract of proposition 2 for types above $\theta_w$ is similar to the optimal contract in proposition 1. It is straightforward to check the optimality conditions of lemma 2 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 (7) 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 $(\theta_1, \theta_w)$. Hence, $\lambda(\theta) \in (-F(\theta), 0)$ for $\theta \in (\theta_1, \theta_w)$. From there, all steps of the proof of proposition 1 go through. Just note that $dVV/d\theta \leq 0$ (as derived in (25)) for $\theta < \theta_w$ as $\lambda(\theta) < 0$ and $c_\theta(q(\theta), \theta) \geq 0$ for these types.

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

**Proof of proposition 3** Consider $\theta' = \underline{\theta}$. Define $\underline{W} = W^{fb}(\underline{\theta}) = W^{sb}(\underline{\theta})$. Since $\underline{\theta}$ 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 $\theta''$ such that $W^{sb}(\theta'') = \underline{W}$ and $W^{sb}_{\theta}(\theta'') > 0$.

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

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

The virtual valuation can be written as

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

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

If $c_\theta(q(\theta''), \theta'') < 0$, it follows that $VV(\underline{\theta}) > VV(\theta'')$. By continuity of $W^{sb}$, there exist types $\theta$ 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 $\theta''$ although the former yield higher (second best) welfare $W^{sb}$. *Q.E.D.*

**Proof of proposition 4:** 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. Take $\theta_1$ and $\theta_2$ as defined in proposition 2.[15] Because all $\theta \in (\theta_1, \theta_2)$ have $q(\theta) = k(\theta)$, it follows that $bid_\theta(\theta) = 0$ for these

[15] 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.

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 $\theta_1$ and slightly above $\theta_2$ have the same bid but different virtual valuations. Since $q(\theta_i) = k(\theta_i)$, the virtual valuation for $\theta_i$ is $Sk(\theta_i) - c(k(\theta_i), \theta_i)$ for $i = 1, 2$. Consequently, the following equation has to hold if the scoring rule implements the optimal mechanism:

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

This can be rewritten as

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

Note that this equation uniquely pins down $\theta_2$ for a given $\theta_1$.[16] Furthermore, it does so independent of the distribution of types. However, $\theta_2$ is defined by the equation $S - c_q(k(\theta), \theta) + \frac{1-F(\theta)}{f(\theta)c_{q\theta}k(\theta),\theta)}$ which depends on $f(\theta_2)$. Hence, slightly perturbing $f$ around $\theta_2$ changes $\theta_2$ but not the equation above. Consequently, a scoring rule auction cannot implement the optimal mechanism in a generic sense.

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.

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

[16]The reason is that the integrand is negative around $\theta_1$, positive around $\theta_2$ and changes sign only at one type which is between $\theta_1$ and $\theta_2$. This follows from lemma 1.

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

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

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 5** For now, assume $VV(\theta_1) \leq VV(\theta_2)$ in the optimal mechanism. We will deal with the opposite case below. We start by describing the auction rules consisting of the two scoring rules and the function translating qualities into prices in the tie-breaking round. Then we show the equilibrium of the auction (in which the winner and the quality provided by the winner will be in line with the optimal mechanism). Finally, we show that expected rents in the auction are the same as in the optimal mechanism.

We use the scoring rule in the main text as scoring rule A. Scoring rule B is given by $s_B(q, p) = G(Sq - p + \Delta_B(q))$ where $G$ is a strictly increasing function to be determined later and

$$\Delta_B(q) = \int_{q(\theta_w)}^{q} \frac{\lambda(q^{-1}(s))}{f(q^{-1}(s))} c_{q\theta}(s, q^{-1}(s))\, ds \qquad \text{for } q \in [q(\theta_w), q(\bar{\theta})]$$

and $\Delta_B = -\infty$ for $q \notin [q(\theta_w), q(\bar{\theta})]$ where $q(\cdot)$ is the optimal quality schedule according to proposition 2 and $q^{-1}$ is the inverse of this function mapping into $[\theta^w, \bar{\theta}]$. The definition of the scoring rule ensures that a firm of type $\theta$ winning the second score auction will procure the quality assigned by the optimal mechanism in proposition 2.

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 $x(\theta') = x(\theta'')$ in the optimal mechanism (proposition 2). Then $G$ is chosen such that $Sq(\theta') - c(q(\theta'), \theta') + \Delta_A(q(\theta')) = G(Sq(\theta'') - c(q(\theta''), \theta'') + \Delta_B(q(\theta'')))$.[17] As $x'(\theta)$ is negative on $[\underline{\theta}, \theta_1]$ and positive on $(\theta_1, \bar{\theta}]$, $G$ is strictly increasing on the relevant range as required. Second, choose $G$ such that optimal bids of types in $[\theta_w, \theta_2)$ according to score B equal optimal bids of types in $(\theta_1, \theta_w)$ according to score A. From (13), types in $[\theta_w, \theta_2)$ have the same bid.[18] Since, bids in score A are decreasing on $[\underline{\theta}, \theta_1]$ and as $VV(\theta_1) \leq VV(\theta_2)$, this property is in line with $G$ being increasing. In fact, $G$ will be discontinuously increasing at the bid of $\theta_2$ if $VV(\theta_1) < VV(\theta_2)$.[19]

The function translating qualities in the beauty contest into prices of the price auction (in case a tie breaking round is necessary) is chosen to satisfy a similar property. Call this function $z(q)$ and choose it such that it satisfies two properties. First, for $q \in [k(\theta_w), k(\theta_2)]$, $z(q) = c(k(\theta'), \theta')$ with $\theta' \leq \theta_w$ where $x(\theta') = x(\theta'')$ and $q(\theta'') = q$ under the optimal mechanism. Second, $z$ is strictly decreasing. The two properties are compatible because under the optimal mechanism $x$ is decreasing on $(\theta_1, \theta_w)$ and increasing on $(\theta_w, \theta_2)$ and $q$ is increasing.

In the second score auction, it is a dominant strategy to bid the highest score one can provide at non-negative costs. From (13), bids according to scoring rule B are increasing on $(\theta_2, \bar{\theta}]$ and constant on $[\theta_w, \theta_2]$. If types in $[\theta_w, \theta_2]$ bid in the tie breaking beauty contest, $k(\theta)$ is the only quality they can procure at non-negative costs (given the equilibrium bids in the first round that lead to the tie breaking round). Hence, bidding $k(\theta)$ is a dominant strategy in the beauty contest (for types in $[\theta_w, \theta_2]$). It

[17] This will imly that the equilibrium bids of $\theta'$ in scoring rule A and $\theta''$ in scoring rule B are equal.

[18] Also types in $(\theta_1, \theta_w]$ bidding according to score A have the same bid.

[19] Note that a type $\theta_2 - \varepsilon$ will still not want to immitate the much higher bid of a type $\theta_2 + \varepsilon$. While this would increase his chance of winning a lot, he would not be able to deliver the higher score at non-negative profits; see below.

follows that the auction selects the firm selected by the optimal mechanism if firms selecting scoring rule B have types in $[\theta_w, \bar{\theta}]$. A similar result was already shown in the main text for scoring rule A and types in $[\underline{\theta}, \theta^w]$.

It is an equilibrium for types in $[\underline{\theta}, \theta_w]$ to choose selection rule A and for all other types to choose selection rule B. To see this, note first that types in $[\theta_1, \theta_2]$ make zero profits in this equilibrium. Now take a type $\theta' < \theta_w$ and let him deviate to scoring rule B. Given this deviation, it is still dominant to bid $\max_q s_B(q, c(q, \theta'))$ and to deliver $q^d = \arg\max_q s_B(q, c(q, \theta'))$ in case of winning. As $\Delta_B < 0$ and $q^{fb}(\theta') \leq q(\theta_w) = k(\theta_w)$, it follows that $q^d = k(\theta_w)$ (or the deviation bid is $G(-\infty)$ which cannot be a profitable deviation as it wins with zero probability). $\theta'$ has higher costs than $\theta_w$ at $q^d$. As $\theta_w$ makes zero profits in equilibrium, the deviation will lead to negative profits for $\theta'$ and is therefore not profitable. A similar argument can be made for $\theta' > \theta_w$ deviating to scoring rule A.

A firm's expected rent is the same under dual-score auction with tie breaking and the optimal mechanism. This is clearly true for types in $[\theta_1, \theta_2]$ which have zero profits under both mechanisms. Denote the distribution of second highest bids $b^{(2)}$ in the dual-score auction with tie breaking equilibrium by $H$ and let $p(bid, b^{(2)}, q)$ be the price a winning firm receives when bidding *bid* and providing quality $q$ while the second highest bid is $b^{(2)}$. Expected profits for $\theta > \theta_2$ in the dual-score auction with tie breaking can then be written as

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

Since profits are maximized over the own bid and quality $q$, the envelope theorem gives

$$\pi^{ds}_\theta(\theta) = \int_{bid(\theta_2)}^{bid(\theta)} -c_\theta(q(\theta), \theta)\ dH(b^{(2)}) = -x(\theta)c(q(\theta), \theta).$$

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 $\pi_\theta$ is the same in the dual-score auction with tie breaking and the optimal mechanism.

The allocation and the firms' 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 $\tilde{s}_A(q,p) = G(s_A(q,p))$ and $\tilde{s}_B(q,p) = Sq - p + \Delta_B(q)$. The same derivation as above goes then through analogously.[20] *Q.E.D.*

**Proof of lemma 3** Define the function

$$\Phi(\hat{\theta},\theta) = \pi(\theta,\theta) - \pi(\hat{\theta},\theta) \geq 0$$

By IC this function is always positive and equal to zero if $\hat{\theta} = \theta$. In other words, the function $\Phi$ reaches a minimum at $\hat{\theta} = \theta$. Thus truth-telling implies both

$$\left.\frac{\partial \Phi(\hat{\theta},\theta)}{\partial \hat{\theta}}\right|_{\hat{\theta}=\theta} = 0 \tag{26}$$

and

$$\left.\frac{\partial^2 \Phi(\hat{\theta},\theta)}{\partial \hat{\theta}^2}\right|_{\hat{\theta}=\theta} \geq 0 \tag{27}$$

Since equation (26) has to hold for all $\hat{\theta} = \theta$, differentiating gives

$$\left.\frac{\partial^2 \Phi(\hat{\theta},\theta)}{\partial \hat{\theta}^2}\right|_{\hat{\theta}=\theta} + \left.\frac{\partial^2 \Phi(\hat{\theta},\theta)}{\partial \hat{\theta}\partial\theta}\right|_{\hat{\theta}=\theta} = 0.$$

Then equation (27) implies that

$$\left.\frac{\partial^2 \Phi(\hat{\theta},\theta)}{\partial \hat{\theta}\partial\theta}\right|_{\hat{\theta}=\theta} \leq 0.$$

It follows from the definition of $\Phi$ that

$$\left.\frac{\partial^2 \Phi(\hat{\theta},\theta)}{\partial \hat{\theta}\partial\theta}\right|_{\hat{\theta}=\theta} = x_\theta(\theta)c_\theta(q(\theta),\theta) + x(\theta)c_{q\theta}(q(\theta),\theta)q_\theta(\theta) \leq 0$$

which is the inequality in the lemma. *Q.E.D.*

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

[20] $G$ will then be discontinuous at the equilibrium bid of $\theta_1$ instead of the bid of $\theta_2$.

**Proof of lemma 4** The monotonicity of $x(\theta)$ and $q(\theta)$ together with $c_\theta \leq 0$ and $c_{q\theta} < 0$ imply that the local incentive compatibility constraint (SOC) is satisfied.

For global incentive compatibility we first show that no $\theta$ 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 (4), 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

$$\int_\theta^{\hat{\theta}} \int_t^{\hat{\theta}} x_\theta(s)c_\theta(q(s), t) + x(s)c_{q\theta}(q(s), t)q_\theta(s)\ ds\, dt \leq 0. \tag{28}$$

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. Since $t \leq s$ 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 (28) has to hold.

To show that no $\theta$ gains by misrepresenting as $\hat{\theta} < \theta$ we use the following notation introduced in equation (3):

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

The idea is to define the following cost function

$$\tilde{c}(a, \theta) = \min\{c(q(a), a), c(q(a), \theta)\} \tag{29}$$

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

$$\tilde{\pi}(a, \theta) = t(a) - x(a)\tilde{c}(a, \theta). \tag{30}$$

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

$$
\begin{aligned}
& \pi(\hat{\theta},\theta) - \pi(\theta,\theta) \\
\leq\ & \tilde{\pi}(\hat{\theta},\theta) - \tilde{\pi}(\theta,\theta) \\
=\ & \int_{\theta}^{\hat{\theta}} \frac{\partial \tilde{\pi}(a,\theta)}{\partial a} da \\
=\ & \int_{\hat{\theta}}^{\theta} \left( \left. \frac{\partial \pi(a,\theta)}{\partial a} \right|_{\theta=a} - \frac{\partial \tilde{\pi}(a,\theta)}{\partial a} \right) da && (31) \\
=\ & \int_{\hat{\theta}}^{\theta} x_{\theta}(a)(\tilde{c}(q(a),\theta) - c(q(a),a)) + x(a)(\tilde{c}_a(q(a),\theta) - c_q(q(a),a)q_{\theta}(a))da && (32) \\
\leq\ & 0
\end{aligned}
$$

where the first inequality follows from the definition of $\tilde{c}(\cdot)$ and the observation that $\tilde{\pi}(\theta,\theta) = \pi(\theta,\theta)$. Equation (31) follows because $\left.\frac{\partial \pi(a,\theta)}{\partial a}\right|_{\theta=a} = 0$ by the first order condition of truthful revelation. Equation (32) 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

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

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

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

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

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

because in our solution $c_{\theta}(q(\theta),\theta) \leq 0$ for all $\theta$. *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. Since 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 $\theta$ and observe that $\theta_2$ has lower costs at $q(\theta')$:

$$c(q(\theta'), \theta_2) - c(q(\theta'), \theta) = \int_{\theta}^{\theta_2} c_\theta(q(\theta'), t)\, dt < 0. \tag{33}$$

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 $\theta$ 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 $\theta_2$ and $\theta'$ is satisfied (see lemma 4 and recall that $\pi(\theta_2) = 0$). The second term is also negative because of equation (33). 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 $\theta_1$ in place of $\theta_2$.

*Q.E.D.*

**Proof of proposition 6** As shown in the proof of lemma 4, incentive compatibility between $\theta$ and $\hat{\theta}$ boils down to the inequality

$$\int_{\theta}^{\hat{\theta}} \int_{t}^{\hat{\theta}} x_\theta(s) c_\theta(q(s), t) + x(s) c_{\theta q}(q(s), t) q_\theta(s)\, ds\, dt \leq 0.$$

Now note that $c_{\theta\theta} = 0$ implies

$$x_\theta(s) c_\theta(q(s), t) + x(s) c_{\theta q}(q(s), t) q_\theta(s) = x_\theta(s) c_\theta(q(s), s) + x(s) c_{\theta q}(q(s), s) q_\theta(s).$$

But then global incentive compatibility has to be satisfied as $x_s(s) c_\theta(q(s), s) + x(s) c_{\theta q}(q(s), s)$ $q_s(s) \leq 0$ by the local second order condition. *Q.E.D.*

# References


Araujo, A. and H. Moreira (2010). Adverse selection problems without the Spence-Mirrlees condition. *Journal of Economic Theory 145*(5), 1113–1141.

Asker, J. and E. Cantillon (2008). Properties of scoring auctions. *RAND Journal of Economics 39*(1), 69–85.

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

Bagnoli, M. and T. Bergstrom (2005). Log-concave probability and its applications. *Economic Theory 26*(2), 445–469.

Branco, F. (1997). The design of multidimensional auctions. *RAND Journal of Economics 28*(1), 63–81.

Che, Y.-K. (1993). Design competition through multidimensional auctions. *RAND Journal of Economics 24*(4), 668–680.

Fudenberg, D. and J. Tirole (1991). *Game theory.* MIT Press.

Guesnerie, R. and J.-J. Laffont (1984). A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm. *Journal of Public Economics 25*(3), 329 – 369.

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

Laffont, J.-J. and J. Tirole (1987). Auctioning incentive contracts. *Journal of Political Economy 95*(5), 921–937.

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

Lewis, T. and D. Sappington (1989). Countervailing incentives in agency problems. *Journal of Economic Theory 49*(2), 294–313.

Matthews, S. and J. Moore (1987). Monopoly provision of quality and warranties: An exploration in the theory of multidimensional screening. *Econometrica 55*(2), 441–467.

McAfee, R. and J. McMillan (1988). Multidimensional incentive compatibility and mechanism design. *Journal of Economic Theory 46*(2), 335–354.

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

Rochet, J. and L. Stole (2003). The economics of multidimensional screening. In *Advances in economics and econometrics: Theory and applications, Eighth World Congress*, Volume 1, pp. 150–197. Cambridge University Press.

Schottmüller, C. (2011). Adverse selection without single crossing: The monotone solution. *Center Discussion Paper, no. 2011–123*.

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