Regulation of non-marketed outputs and substitutable inputs

AUTHORS

Joachim Bertsch

Simeon Hagspiel

EWI Working Paper, No 15/06

December 2015

Institute of Energy Economics at the University of Cologne (EWI)

www.ewi.uni-koeln.de

Institute of Energy Economics
at the University of Cologne (EWI)

Alte Wagenfabrik
Vogelsanger Straße 321a
50827 Köln
Germany

Tel.: +49 (0)221 277 29-100
Fax: +49 (0)221 277 29-400
www.ewi.uni-koeln.de

CORRESPONDING AUTHOR

Joachim Bertsch
Institute of Energy Economics at the University of Cologne (EWI)
Joachim.Bertsch@ewi.research-scenarios.de

ISSN: 1862-3808

Regulation of non-marketed outputs and substitutable inputs*

Joachim Bertsch and Simeon Hagspiel

Tuesday 1st December, 2015

We study the regulation of a monopolistic firm that provides a non-marketed output based on multiple substitutable inputs. The regulator is able to observe the effectiveness of the provision, but she faces information asymmetries with respect to the efficiency of the firm's activities. Motivated by the example of electricity transmission services, we consider a setting where one input (grid expansion) and the output (uninterrupted electricity transmission) are observable, while another input (sophisticated grid operation) and related costs are not. Multi-dimensional information asymmetries are introduced by discrete distributions for the functional form of the marginal rate of substitution between the inputs as well as for the input costs. For this novel setting, we investigate the theoretically optimal Bayesian regulation mechanism. We find that the first best solution cannot be obtained in case of shadow costs of public funding. The second best solution implies separation of the most efficient type with first best input levels, and upwards distorted (potentially bunched) observable input levels for all other types. Moreover, we compare these results to a simpler non-Bayesian approach and hence, bridge the gap between the academic discussion and regulatory practice. We provide evidence that under certain conditions, a single contract non-Bayesian regulation can indeed get close to the second best of the Bayesian menu of contracts regulation.

JEL classification: D42, D82, L51

Keywords: Regulation, Asymmetric Information, Mechanism Design, Non-Marketed Goods, Substitutable Inputs, Electricity Transmission

*The authors want to thank Felix Höffler and Christian Tode for their helpful comments, Clara Dewes for her support, and the participants of the IAEE 2014 in Rome and the Spring Meeting of Young Economists 2015 in Ghent for valuable discussions. Funding of the German research society DFG through grant HO 5108/2-1 is gratefully acknowledged.

1 Introduction

Numerous goods and services are provided by regulated firms with a monopolistic status. For instance, uninterrupted electricity transmission services – being a textbook example of a natural monopoly – are usually provided by a single firm. Currently, an increasing deployment of renewable energy sources leads to substantially changing requirements to secure an uninterrupted electricity transmission, while multiple substitutable measures may exist to cope with it, such as grid expansion or sophisticated grid operation.[1] Due to the fact that an uninterrupted electricity transmission is crucial for society, the regulator will be well aware of whether or not it has been provided *effectively*.[2] In contrast, however, electricity systems are highly complex, such that interdependent activity levels as well as related cost figures are hard to assess. Hence, it may be difficult for the regulator to judge the *efficiency* of the firm's underlying measures. Technically speaking, this situation may be seen as a production process involving multiple substitutable inputs, incorporating two adverse selection problems: First, the regulator may have a hard time estimating the necessary overall level of the firm's activity, determined by the marginal rate of technical substitution (MRTS), i.e., the isoquant function describing the relation of inputs needed to produce the requested output. Second, the regulator may have difficulties verifying the unit costs of one or multiple inputs. This multi-dimensional asymmetric information increases the complexity of finding an adequate regulation.

In theory as well as in practice, problems of information asymmetry between the regulator and the firm have been tackled by different forms of regulation. Typical approaches in regulatory practice range from cost-based regulation to widely applied incentive regulation (discussed, e.g., in Joskow (2014)), or a linear combination of those two extremes (e.g., Schmalensee (1989)). For instance, the German regulator offers one single contract to electricity transmission firms, dependent on grid expansion, which corresponds to a cost-based regulation of capital.[3] The academic discussion has not yet fully covered the specific multi-dimensional problems of asymmetric information regarding the level and mix of inputs, but more recent theoretical approaches suggest that the best theoretical solution consists of the regulator offering the firm a *menu* of contracts, such that the firm reveals her private information (e.g., Laffont and Tirole (1993)). Even

[1]The German Transmission System Operators estimate the necessary investments into grid reinforcements and expansion to be around 22 bn. for the period 2013-2022 (Netzentwicklungsplan (2013)), which doubles the annual figures for 2012 and quadruples the value for 2006 (Monitoringbericht (2013)).

[2]For instance, in Germany the regulator has defined five observable, quantifiable dimensions for measuring grid quality.

[3]In Germany, transmission system operators formulate a network expansion plan for which they get an allowed investment. In line with economic theory, the chosen levels may be suspected to be inefficiently high (see Footnote 1 for related cost figures). This regulation corresponds to a cost-based regulation for the input factor grid expansion, while neglecting any other possible input, such as better operational measures. Obviously, this triggers some sort of Averch-Johnson-effect and leads to suboptimal distortions of the input levels.

though the dichotomy between such Bayesian models of regulation (which tend to dominate the academic discussion) and simpler non-Bayesian models (which are closer to regulatory practice) is well perceived, corresponding explanations are rather vague. For instance, as Armstrong and Sappington (2007) note, "[...] *regulatory plans that encompass options are 'complicated', and therefore prohibitively costly to implement*".

The goal of this paper is twofold: First, to identify and investigate the optimal Bayesian regulation for the multi-dimensional problem at hand, and second, to bridge the gap between the theoretically optimal solution and simpler regimes applied in regulatory practice.

To derive an optimal regulation strategy, we build on the theory of incentives and contract menus. It is well known that in a simple setting with two types of the firm, the efficient type is incentivized via a contract with first best (price) levels along with some positive rent, while the inefficient type's contract includes prices below the first best and no rent (e.g., Laffont and Tirole (1993)). This analysis has been extended to represent multiple dimensions of information asymmetry in terms of adverse selection, e.g., by Lewis and Sappington (1988b), Dana (1993), Armstrong (1999) or Aguirre and Beitia (2004). While Dana (1993) analyzes a multi-product environment, Lewis and Sappington (1988b), Armstrong (1999) and Aguirre and Beitia (2004) consider two-dimensional adverse selection with only one screening variable. Specifically, the latter three derive optimal regulation strategies in a marketed-good environment (in the sense of Caillaud et al. (1988)) with unknown cost and demand functions. In our paper, unlike Lewis and Sappington (1988b) and Armstrong (1999), we consider shadow costs of public funding instead of distributional welfare preferences. Despite technical differences, this is largely in line with the analysis of Aguirre and Beitia (2004).[4] However, in contrast to all these papers, we solve the two-dimensional adverse selection problem for a non-marketed good environment and a production process that involves two substitutable inputs with an uncertain isoquant and input factor costs.[5]

For the novel setting of multi-dimensional inputs and a non-marketed output, we are able to confirm the general insights from the above literature. We find that expected social welfare necessarily includes positive rents for some types of the firm, such that the first best solution cannot be achieved. While the efficient type is always set to first best input levels, the other contracts' (observable) input levels are distorted upwards.[6] Separation of at least three types is

[4]Aguirre and Beitia (2004) show the difference between shadow costs of public funding and distributional welfare preferences based on a model with continuous probability distribution, while we assume a discrete distribution.

[5]Noticeably, with the (discrete) two-dimensional adverse selection problem, our problem setting is technically closest to the model discussed by Armstrong (1999).

[6]Upwards distorted observable input levels coincide with upwards distorted prices for the inefficient type as shown in Laffont and Tirole (1993). They also agree with the results in a setting with unknown cost and demand functions as long as shadow costs of public funding are considered (Aguirre and Beitia (2004)). Noticeably, the case of prices below marginal costs, as found in Lewis and Sappington (1988b) and Armstrong (1999), is mainly

always possible, while bunching of two types may be unavoidable in case of a very asymmetric distribution of costs or very flat isoquants.

We compare the obtained optimal Bayesian regulation to the results of a non-Bayesian regulation that we obtain by restricting our regulation problem to one single contract. We find that despite the general inferiority a non-Bayesian cost-based regulatory regime may indeed be close to the optimal Bayesian solution for specific circumstances. This especially holds true if the overall input level probably needs to be high, and shadow costs of public funding are large. Considering current circumstances observed in the electricity sector, i.e., substantial changes in the supply structure and ongoing intense discussions about grid tariffs, these conditions may indeed prevail.

The paper is organized as follows: Section 2 introduces the model, Section 3 presents the optimal regulation strategy, Section 4 compares the optimal regulation to simpler regimes, and Section 5 concludes.

2 The model

Consider a single firm that is controlled by a regulator. The firm uses two inputs to provide an output in terms of a good or service level q that is requested by the regulator. The regulator's choice of q could, for instance, result from counterbalancing the economic value of the provided with the related social costs. For simplicity, however, we assume q to be invariant throughout the paper. Although this assumption might seem restrictive at first sight, it may indeed fit a number of relevant cases very well. For instance, due to the very high societal value of uninterrupted electricity transmission, changes in costs will hardly affect the desired level of the transmission service quality q.

In our model, probability μ (respectively $1 - \mu$) leads to a low (high) aggregated input that is necessary to reach the same requested output q. This could, e.g., be an exogenous shock induced by the increased deployment of renewable energies, triggering a changing spatial distribution of supply and hence impacting the necessary overall activity level in the grid sector to achieve secure electricity transmission. From the firm's perspective, an output level q can be provided by means of two different inputs, one of which is observable (x) and one non-observable (y) by the regulator. Stressing again our introductory example of electricity transmission services, x could be the level of grid expansion that is easily observable by the regulator, even by people unfamiliar with the details of electricity transmission. Utilization and measures of sophisticated grid operation, especially as a partial substitute to grid expansion, however, are hardly observable. The tradeoff between those two inputs needed to reach output q is commonly described by

triggered by using a distributive social welfare function instead of shadow costs of public funding.

a production function $q = f(x, y)$ which can be illustrated by means of isoquants. We assume smooth and decreasing marginal returns of both inputs, such that the isoquants are downward sloping, convex and differentiable. Noticeably, two different isoquants can never cross. An example fulfilling these requirements is a Cobb-Douglas-type production function. The inverse production function $g(q, x)$ reflects the necessary level of the non-observable input y needed to reach output q, given a level of x. We will mostly use this inverse function hereafter. Due to the exogenous shock leading to a low (l) or high (h) aggregated input necessary for the envisaged output level q, the inverse function takes one of two possible functional forms, i.e. $g_i(q, x)$, with $i \in [l, h]$ and $g_l(q, x) < g_h(q, x)$.

The optimal rate of substitution between the two inputs minimizing total costs for reaching the requested output depends on the cost functions of the inputs. We consider the cost function $c^x(x)$ of the the observable input to be fixed and common knowledge, while the cost function of the non-observable input $c_j^y(y)$ is subject to a nature draw, which leads with probability ν (respectively $1 - \nu$) to a low (high) cost function (i.e., $j \in [l, h]$). For simplicity, we assume constant factor costs of both inputs, i.e., $c^x(x) = c^x$ and $c_j^y(y) = c_j^y$. The realization of c_j^y influences the isocost line of the two inputs and hence, the optimal rate of substitution.[7] Hence, depending on the two random draws for the isoquant and the costs of the non-observable input, there are four possible first best bundles of inputs, which we denote by $\{x_{ll}^{fb}, y_{ll}^{fb}\}$, $\{x_{lh}^{fb}, y_{lh}^{fb}\}$, $\{x_{hl}^{fb}, y_{hl}^{fb}\}$ and $\{x_{hh}^{fb}, y_{hh}^{fb}\}$. As a last precondition, we require the expansion path, i.e. the curve connecting the optimal input combinations of the different isoquants, to be pointing rightwards as the necessary aggregated input increases.[8] In terms of the first best input levels, this requires $x_{ll}^{fb} > x_{hl}^{fb}$ and $x_{lh}^{fb} > x_{hh}^{fb}$, which again holds true for a wide range of possible production function specifications, including the above mentioned Cobb-Douglas type.

Under optimal Bayesian regulation, the goal of the regulator is to incentivize the firm via a suitable contract framework to choose the welfare-optimizing bundle of inputs, which we will derive based on classic mechanism design entailing truthful direct revelation. Contrary to the firm, the regulator cannot observe the realizations of the two random draws, although the possible realizations as well as the occurrence probabilities are common knowledge. She knows the cost function of the observable input and can observe the corresponding input level. The output is also observable and verifiable.[9] For an optimal regulation, the regulator offers the firm a menu of four contracts, each with a level of the observable input x_{ij} and a corresponding

[7]As it is well known from production theory, the optimal rate of substitution is determined by equating the marginal rate of technical substitution between the factors (i.e., the slope of the isoquant) with the relative factor costs (i.e., the slope of the isocost line).

[8]For an analysis involving continuous variables, this would require the expansion path to behave like a function with a unique function value y for each x, or, in other words, an expansion path that is not bending backwards.

[9]Stochastic deviations due to force majeure are supposed to be detectable and excludable from the contract framework.

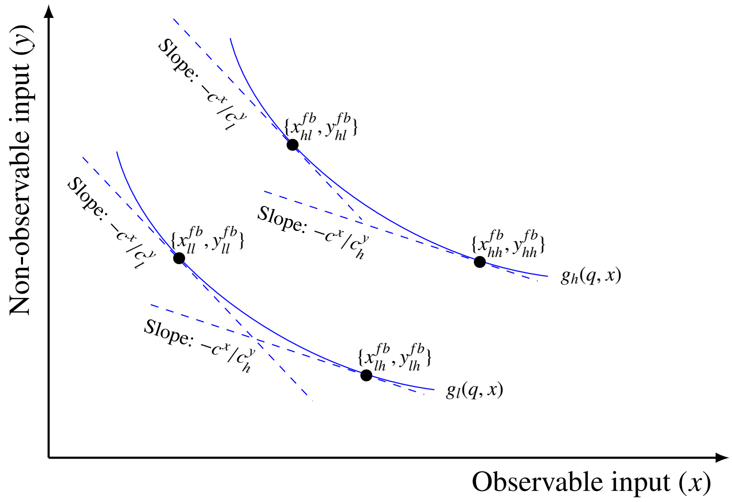

Figure 1: Problem setting with double adverse selection

transfer T_{ij}. Naturally, the contracts can be conditioned on observable parameters only, i.e., the output as well as the amount of the observable input used. Both are enforceable by means of suitably high penalties in case the firm deviates from the requested/contracted level.

The timing – as shown in Figure 2 – is as follows. First, the random draws are realized and the cost function of the non-observable input and the necessary aggregated input relation (isoquant) are observed by the firm. The firm then chooses between several (in our case, four) contracts offered by the regulator. She then realizes the input levels to produce the requested output. The regulator observes one input level (x) and whether the output is as requested; if those are as agreed upon, the contract is executed and the transfer realized.

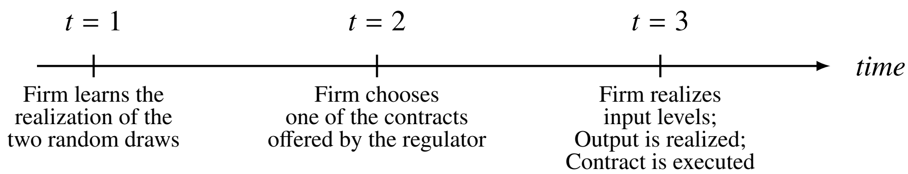

Figure 2: Timing

The rent of the firm R_{ij} given a realization $i \in [l, h]$ and $j \in [l, h]$, results from the transfer T_{ij} minus the private cost of the firm's activities:[10]

$$R_{ij} = T_{ij} - c^x x_{ij} - c^y_j g_i(q, x_{ij}) \tag{1}$$

[10]It goes without saying here that the firm is characterized such that she tries to maximize her rent.

The regulator maximizes expected social welfare, defined as the sum of expected social utility and firm surplus, by adjusting the observables, i.e.:

$$\max_{x_{ij},T_{ij}} W = \mathbb{E}\left[\underbrace{S_q - (1+\lambda)T_{ij}}_{\text{Net social utility}} + \underbrace{(T_{ij} - c^x x_{ij} - c^y_j g_i(q, x_{ij}))}_{\text{Firm's rent } (R_{ij})}\right] \tag{2}$$

where S_q is the gross social utility from reaching output q, and λ denotes the shadow costs of public funding, i.e., the costs due to raising and transferring finances through public channels (for a discussion, see, e.g., Laffont and Tirole (1993)). As discussed previously, we assume q – and hence also gross social utility S_q – to be invariant and independent of the random draws, yielding[11]

$$\max_{x_{ij},T_{ij}} W = S_q - \mathbb{E}\left[\underbrace{(1+\lambda)T_{ij}}_{\text{Transfer costs}} - \underbrace{(T_{ij} - c^x x_{ij} - c^y_j g_i(q, x_{ij}))}_{\text{Firm's rent } (R_{ij})}\right] \tag{3}$$

As an important consequence of Equation (3), we see that the optimization problem of the regulator can be reformulated in terms of a cost-minimization problem, essentially stating that the desired output shall be reached at minimal expected social costs:

$$\min_{x_{ij},T_{ij}} C = \mathbb{E}\left[C_{ij}\right] = \mathbb{E}\left[\lambda \underbrace{R_{ij}}_{\text{Firm's rent}} + (1+\lambda)(\underbrace{c^x x_{ij}}_{\substack{\text{Costs of}\\ \text{observable input}}} + \underbrace{c^y_j g_i(q, x_{ij})}_{\substack{\text{Costs of}\\ \text{non-observable input}}})\right] \tag{4}$$

While choosing x_{ij} and T_{ij} such that social costs are minimized, the regulator is restricted by several participation and incentive constraints for the firm's rent:

$$R_{ij} \geq 0 \qquad \forall i, j \tag{5}$$

$$R_{ij} \geq R_{i'j'} + c^y_{j'} g_{i'}(q, x_{i'j'}) - c^y_j g_i(q, x_{i'j'}) \qquad \forall \text{ pairs } i, j \text{ and } i', j' \tag{6}$$

Equation (5) ensures that all types of firms have a non-negative profit and therefore participate.[12] In line with the revelation principle, Equation (6) provides the firm with the incentive to truthfully report the realized isoquant and non-observable input costs.

[11]This is the reason why q appears as a subscript here. In case of a more complex analysis involving q as a variable, S_q would be replaced by $S(q, x)$ to reflect the counterbalancing of the economic value of the provided output with the related social costs.

[12]Hence, we implicitly assume zero liability for the firm.

Written explicitly, the four participation constraints for the four possible firm types become

$$R_{ll} \geq 0 \tag{7a}$$
$$R_{lh} \geq 0 \tag{7b}$$
$$R_{hl} \geq 0 \tag{7c}$$
$$R_{hh} \geq 0, \tag{7d}$$

and the twelve incentive constraints (each of the four types might be tempted to choose a contract of one of the other three types)

$$R_{ll} \geq R_{lh} + c_h^y g_l(q, x_{lh}) - c_l^y g_l(q, x_{lh}) \tag{8a}$$
$$R_{ll} \geq R_{hl} + c_l^y g_h(q, x_{hl}) - c_l^y g_l(q, x_{hl}) \tag{8b}$$
$$R_{ll} \geq R_{hh} + c_h^y g_h(q, x_{hh}) - c_l^y g_l(q, x_{hh}) \tag{8c}$$
$$R_{lh} \geq R_{ll} + c_l^y g_l(q, x_{ll}) - c_h^y g_l(q, x_{ll}) \tag{8d}$$
$$R_{lh} \geq R_{hl} + c_l^y g_h(q, x_{hl}) - c_h^y g_l(q, x_{hl}) \tag{8e}$$
$$R_{lh} \geq R_{hh} + c_h^y g_h(q, x_{hh}) - c_h^y g_l(q, x_{hh}) \tag{8f}$$
$$R_{hl} \geq R_{ll} + c_l^y g_l(q, x_{ll}) - c_l^y g_h(q, x_{ll}) \tag{8g}$$
$$R_{hl} \geq R_{lh} + c_h^y g_l(q, x_{lh}) - c_l^y g_h(q, x_{lh}) \tag{8h}$$
$$R_{hl} \geq R_{hh} + c_h^y g_h(q, x_{hh}) - c_l^y g_h(q, x_{hh}) \tag{8i}$$
$$R_{hh} \geq R_{ll} + c_l^y g_l(q, x_{ll}) - c_h^y g_h(q, x_{ll}) \tag{8j}$$
$$R_{hh} \geq R_{lh} + c_h^y g_l(q, x_{lh}) - c_h^y g_h(q, x_{lh}) \tag{8k}$$
$$R_{hh} \geq R_{hl} + c_l^y g_h(q, x_{hl}) - c_h^y g_h(q, x_{hl}). \tag{8l}$$

3 Optimal regulation

3.1 Preparatory analysis

As a first preparatory step in the analysis we shall check whether the contract variable x is actually suitable to provide incentives to the firm to reveal her true type. To this end, we investigate whether the incentive to choose another type's contract (motivated by a potential increase in rent) regarding one of the two random draws is impacted by an adjustment of x. This is often referred to as "single crossing" conditions. For the incentive to choose another type's contract regarding the realized input cost, we find that[13]

$$\frac{\partial}{\partial x}(R_{ih}(x) - R_{il}(x)) = (c_l^y - c_h^y)g_i'(q,x) \qquad \text{for} \qquad i = l, h, \tag{9}$$

which is clearly greater than zero due to $c_h > c_l$ and $g_i'(q,x) < 0$. Hence, by an upwards distortion of x, we are able to reduce the incentive for the firm to choose the contract of a high cost type instead of truly revealing the realized low cost type.

Similarly, for the incentive to choose a contract for an isoquant different from the realized one, we find that

$$\frac{\partial}{\partial x}(R_{hj}(x) - R_{lj}(x)) = c_j^y(g_l'(q,x) - g_h'(q,x)) \qquad \text{for} \qquad j = l, h \tag{10}$$

which is greater than zero as long as $g_h'(q,x) < g_l'(q,x)$. Recalling from Section 2 that we have assumed rightwards pointing expansion paths (a property exhibited by a wide range of possible production function specifications, including the Cobb-Douglas type), this condition will always hold true. Hence, upwards distorting x will provide a possibility to reduce the incentive for the firm to choose the contract with a high isoquant instead of truly revealing the realized low isoquant.

The effect of changing incentives following a distortion of x helps us to derive a first characterization of the optimal solution of our regulatory problem. In fact, in order to comply with the incentive constraints (8a)-(8l) (which need to be fulfilled for the optimal solution anyway), input levels x_{ij} need to follow a certain ordering. Note that for each pair of types there are two relevant incentive constraints (e.g., Equations (8a) and (8d) for the types ll and lh). Adding those and using the above single crossing conditions, the necessary ordering can be obtained as follows:[14]

$$x_{ll} \leq x_{lh} \leq x_{hh} \tag{11}$$

$$x_{ll} \leq x_{hl} \leq x_{hh} \tag{12}$$

Moreover, from the incentive constraints (8a) and (8i) it follows that only the participation constraints (7b) and (7d) (i.e., limited liability of the lh and the hh-type) remain relevant for further analyses. In contrast, the other two participation constraints (those of the low-cost types) are implicitly fulfilled if these two incentive constraints hold.

So far unclear from the above analysis, however, is the ordering of the intermediate cases x_{lh} and x_{hl}, which depends on whether the term $R_{hl}(x) - R_{lh}(x)$ is increasing or decreasing in x.

[13]Here and in the following, a prime denotes derivation with respect to x.

[14]For instance, adding Equations (8a) and (8d) yields $(c_l^y - c_h^y)g_l'(q, x_{lh}) \geq (c_l^y - c_h^y)g_l'(q, x_{ll})$, which, together with (9), implies that $x_{lh} \geq x_{ll}$.

Differentiating with respect to x yields

$$\frac{\partial}{\partial x}(R_{hl}(x) - R_{lh}(x)) = (c_h^y g_l'(q,x) - c_l^y g_h'(q,x)) \tag{13}$$

which is increasing in x as long as

$$\frac{c_h^y}{c_l^y} < \frac{g_h(q,x)}{g_l(q,x)}, \tag{14}$$

and decreasing in x otherwise. Together with incentive constraints (8e) and (8h) we infer that if the cost variation is small compared to the isoquant variation, then $x_{lh} \leq x_{hl}$. If the aggregated input level variation is small compared to the cost variation, then $x_{lh} \geq x_{hl}$. For an intuition, recall Figure 1. If the aggregated input level variation and hence the distance between the isoquants is large, x_{hl}^{fb} is larger than x_{lh}^{fb}. If the cost variation, and hence, the vertical distance between the corresponding first best solutions is large, x_{lh}^{fb} is larger than x_{hl}^{fb}.

The results of our preparatory analysis are summarized in the following two Lemmas.

Lemma 1. *Limited liability is only an issue for the high-cost types. Hence, the only relevant participation constraints are* (7b) *and* (7d)*, whereas* (7a) *and* (7c) *are implicitly fulfilled.*

Lemma 2. *In order to reach incentive compatibility, input levels x_{ij} must be ordered as follows:*

(A) If the cost variation is small compared to the isoquant variation, then $R_{hl}(x) - R_{lh}(x)$ is increasing in x and requires

$$x_{ll} \leq x_{lh} \leq x_{hl} \leq x_{hh}. \tag{15}$$

(B) If the cost variation is large compared to the isoquant variation, then $R_{hl}(x) - R_{lh}(x)$ is decreasing in x and requires

$$x_{ll} \leq x_{hl} \leq x_{lh} \leq x_{hh}. \tag{16}$$

3.2 Full information benchmark

If the regulator had no information deficit, she would observe the realized isoquant as well as the realized isocost line. Differentiating all possible realizations of the social cost function C_{ij} with respect to the observable input levels x_{ij} shows that all of them are single-peaked with a unique minimum at $g_i'(q, x_{ij}) = -\frac{c^x}{c_j^y}$, which is necessarily realized at $x_{ij} = x_{ij}^{fb}$. The regulator would easily derive the first best levels of inputs to supply the requested output at minimal social costs, i.e., $\{x_{ij}^{fb}, y_{ij}^{fb}\}$, by equating the known realized marginal rate of technical substitution of the inputs with the realized isocost line. Moreover, she would be able to enforce the implementation of the first best due to the full observability. The corresponding optimal transfers would be $T_{ij}^{fb} = c^x x_{ij}^{fb} + c_j^y y_{ij}^{fb}$, leaving all types of the firm with zero rent. In the case of full information, social costs amount to $C_{ij}^{fb} = (1+\lambda)T_{ij}^{fb} = (1+\lambda)(c^x x_{ij}^{fb} + c_j^y y_{ij}^{fb})$, corresponding to the welfare-optimizing first best solution that could thus be obtained.

3.3 Asymmetric information

In the case of asymmetric information, the only two observables for the regulator are the output q and the observable input x. In addition, she can choose an appropriate level of transfer payment T. As q is invariable and observable, its implementation can be enforced by means of suitably high penalties in case the firm deviates. Hence, x and T are the two variables the regulator will condition her contracts on. The general idea for the regulator's optimal regulation strategy is to offer a menu of contracts with optimized variables $\{x_{ij}^*, T_{ij}^*\}$, such that expected social costs are minimized (as stated in Equation (4)), and participation (Equation (5)) and incentive constraints (Equation 6) fulfilled. Hence, we restrict our attention to incentive compatible contracts ensuring that the firm always reveals her true type. Under these conditions, the revelation principle requires that the solution found (if any) is a Bayesian-Nash equilibrium (Myerson (1979), Laffont and Martimort (2002)).

3.3.1 One-dimensional asymmetric information

We shall first investigate a simplified problem with one-dimensional asymmetric information only, i.e., isoquant *or* cost uncertainty. Eliminating the isoquant uncertainty (by setting $\mu = 0$, $\mu = 1$ or $g_{lj} = g_{hj}$), we are left with two constraints binding: the participation constraint of the high cost type (7b or 7d) and the incentive constraint from the low to the high cost type (8a or 8i). This leads to the simplified cost function:

$$\begin{aligned} C = &\nu \left[\lambda \left(g_i(x_{ih})(c_h^y - c_l^y) \right) + (1+\lambda) \left(c^x x_{il} + c_l^y g_i(x_{il}) \right) \right] \\ &+ (1-\nu) \left[(1+\lambda) \left(c^x x_{ih} + c_h^y g_i(x_{ih}) \right) \right] \end{aligned} \tag{17}$$

Derivating with respect to x_{ij}, $j \in l, h$ yields the following first order conditions:

$$\frac{\partial C}{\partial x_{il}} = 0 \Leftrightarrow g_i'(x_{il}^*) = -\frac{c^x}{c_l^y}, \tag{18}$$

$$\frac{\partial C}{\partial x_{ih}} = 0 \Leftrightarrow \underbrace{\nu\lambda(c_h^y - c_l^y) g_i'(x_{ih}^*)}_{<0} + \underbrace{(1-\nu)(1+\lambda)(c^x + c_h^y g_i'(x_{ih}^*))}_{\substack{=0 \text{ for } x_{ih}=x_{ih}^{fb} \\ <0 \text{ for } x_{ih}<x_{ih}^{fb} \\ >0 \text{ for } x_{ih}>x_{ih}^{fb}}} = 0 \tag{19}$$

Similarly, in case of no cost uncertainty, the observable input levels of the low isoquant types are first best, whereas the high isoquant types are distorted upwards:[15]

Lemma 3. *In case of asymmetric information about either costs or isoquants, the respective l-type is set to first best, while the h-type is distorted upwards compared to its first best.*

Note that the result of an adverse selection problem with one-dimensional information asymmetry on costs is well-known from the literature (e.g., Baron and Myerson (1982) or Sappington (1983)). Also note that the results concerning isoquant uncertainty are strikingly different compared to the one-dimensional demand uncertainty (which essentially corresponds to the isoquant in our setting) studied by Lewis and Sappington (1988a) or Armstrong (1999). In contrast to our model – due to neglecting shadow costs of public funding – they find that the first best can be achieved in the one-dimensional case of demand uncertainty.

3.3.2 Two-dimensional asymmetric information

Solving the full optimal regulation problem requires minimization of social costs, subject to all imposed four participation and twelve incentive constraints. Due to the large number of constraints, we approach the optimization by solving a relaxed problem where only a subset of the constraints is considered. To this end, we need to come up with an educated guess about the binding constraints in the optimum. If we can later show that the remaining constraints are fulfilled at the solution of the relaxed problem, we will have obtained the solution of the full problem.

We already know from Lemma 1 that the participation constraints of the high-cost types are the only relevant ones. Furthermore, it generally seems to be a good approach to assume the "upwards" incentive constraints, i.e., from low to high isoquant, and from low to high costs, to be binding. Moreover, it seems plausible to assume binding incentive constraints from the most efficient to an intermediate type (i.e., *lh* or *hl*), and from an intermediate type to the least efficient type. If we consider the isoquant variation more relevant than the cost variation, assuming the incentive constraints according to the ordering shown in Lemma 2, Case *(A)*, to be binding appears to be the most educated guess we can come up with.[16] Hence, we assume that incentive constraints (8a) ($ll \rightarrow lh$), (8e) ($lh \rightarrow hl$), and (8i) ($hl \rightarrow hh$) are fulfilled with equality. In addition, we assume the participation constraint of the *hh*-type to bind since this is the only type remaining that is not attracted by any other type. Figure 3 illustrates with arrows the binding incentive constraints, such that the former type is not attracted by the latter type-contract. Diamonds mark the binding participation constraints.

[15] Due to the obvious symmetry of the problem, we omit the detailed calculation here.

[16] The ordering and solution of Case *(B)* is reversed, but similar. The corresponding discussion can be found in the

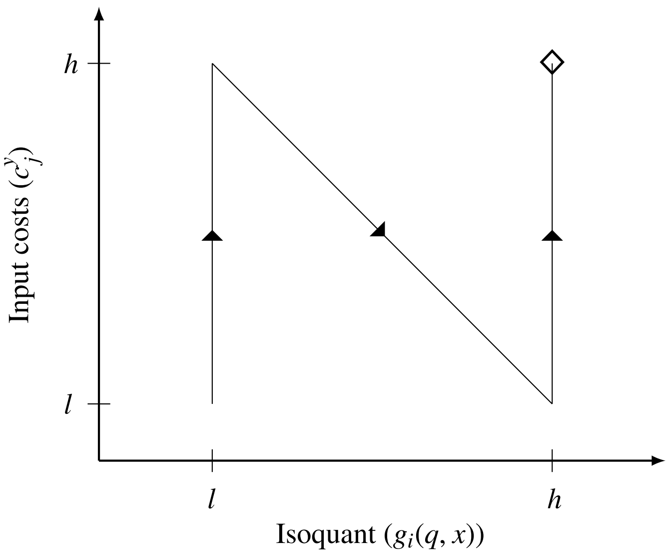

Figure 3: Constraints considered binding for Case *(A)*

We find that this set of assumptions does indeed lead us to the optimal regulation strategy. The results are summarized in the following Proposition 1.

Proposition 1. *For Case (A),*

(i) Optimal regulation is achieved under the following set of observable input levels:

$$x^*_{ll} = x^{fb}_{ll} \tag{20}$$

$$x^*_{lh} \geq x^{fb}_{lh} \tag{21}$$

$$x^*_{hl} \geq x^{fb}_{hl} \tag{22}$$

$$x^*_{hh} \geq x^{fb}_{hh}, \tag{23}$$

while respecting $x^*_{ll} < x^*_{lh} \leq x^*_{hl} \leq x^*_{hh}$.

(ii) The most efficient (ll) type can always be separated. Moreover, separation of at least three types is always possible, while bunching of the lh and hl types is unavoidable in case of $\nu \to 1$. *The hl and hh types may need to be bunched in case of* $g'_l(q, x) \to 0$ *together with* c^y_l *being large.*

Proof. See Appendix. □

appendix.

Corollary 1. *For $\lambda = 0$, the optimal solution is first best. All input levels amount to $x^*_{ij} = x^{fb}_{ij}$, and expected social costs to $C = C^{fb}$.*

Proof. Follows immediately from the solution of Case *(A)* when setting $\lambda = 0$. □

According to Corollary 1, with no shadow costs of public funding, all input levels x^*_{ij} are first best. The regulator optimizes overall welfare, but has no preference regarding the distribution of social surplus. Hence, she can give the firm an arbitrarily high budget at no social costs, and the firm maximizes her rent by setting efficient input levels. In this case, the maximization of the firm and the maximization of social welfare coincide, i.e., there is no problem of aligning the activities of the firm with social interests. Of course, larger parts of the welfare are then given to the firm.

For the general case of $\lambda > 0$, observable input levels of all types besides the *ll*-one are distorted above first best levels, leading to a second best solution only. Naturally, the overall level of inefficiency increases in λ, but also for decreasing μ and ν (i.e., when there is a high probability for "costly" outcomes of the random draws) as well as for $c^y_h - c^y_l$ and $g_h(q, x) - g_l(q, x)$ getting large. In contrast, however, the less significant the cost variation becomes compared to the isoquant variation, the more efficient the solution will be.

Due to keeping the most efficient (*ll*) type at first best level combined with the ordering according to Lemma 2, the type can always be separated in the contract framework. Moreover, we find that at least three types can always be separated, while bunching of two types may be unavoidable in case of vanishing isoquant or cost uncertainties, or if the isoquant variation becomes extremely large. As a last remark, it is worth mentioning that the ordering of rents is (and must be) as depicted in Figure 3, i.e. $0 = R^*_{hh} < R^*_{hl} < R^*_{lh} < R^*_{ll}$.

The results for Case *(B)* are symmetric but structurally identical to Case *(A)*, i.e., the *ll*-type is incentivized to first best input levels while the other types show upwards distortions of x_{ij}. However, roles of isoquants and costs are interchanged, reflected in the inverse occurrence of the terms $g_i \leftrightarrow c^y_j$ and $\mu \longleftrightarrow \nu$. At the same time, as imposed by Lemma 2, Case *(B)*, the sequence of the "intermediate" types is now $hl \rightarrow lh$. Hence, the ordering of observable input levels x_{lh} and x_{hl} as well as rents R_{lh} and R_{hl} need to be reversed to obtain an optimal regulatory contract framework.[17]

4 Comparing the optimal regulation to simpler regimes

In contrast to the optimal Bayesian menu of contracts studied in the previous section, regulatory authorities often apply alternative, simpler approaches. In fact, in the case of electricity trans-

[17] See the appendix for a detailed discussion and the corresponding proposition and proof.

mission grids, it appears that they mostly offer a non-Bayesian, i.e., *single*, contract, while the application of Bayesian contracts in terms of *menus* of contracts, has been very rare.[18] For instance, regulatory practice in Germany is such that TSOs formulate a grid expansion plan, which is then reviewed and approved by the regulator. For the approved measures, the TSOs get their costs reimbursed. This corresponds to a cost-based regulation for the input factor grid expansion, while neglecting any other possible input, such as better operational measures. Meanwhile, driven e.g. by social acceptance issues, the regulator is expected to limit the approval of extensive grid expansion to some "reasonable" level.

Transferring such a simple non-Bayesian approach into our model, we need to limit the set of regulatory choice variables to one single contract with contract variables $\bar{x}$ and $\bar{T}$, such that the objective function of the regulator (in contrast to Equation (4) as in the case of optimal regulation) becomes:

$$\min_{\bar{x},\bar{T}} \bar{C} = \mathbb{E}\left[\lambda \underbrace{\bar{R}_{ij}}_{\text{Firm's rent}} +(1+\lambda)(\underbrace{c^x \bar{x}}_{\substack{\text{Costs of}\\ \text{observable input}}} + \underbrace{c^y_j g_i(q,\bar{x})}_{\substack{\text{Costs of}\\ \text{non-observable input}}})\right] \tag{24}$$

In contrast to the solution of the optimal regulation, this minimization is only subject to the participation constraints (5). With a sole contract and hence, only one observable input $\bar{x}$ for all types, the regulator has no possibility to separate types, which makes the incentive constraints obsolete. As before, the only participation constraint holding with equality is the one of the *hh*-type. Considering that this type gets full cost reimbursement but cannot be distinguished from the other types, it becomes clear that all other types must then necessarily receive a positive rent. The following proposition summarizes the solution of this non-Bayesian regulatory approach.[19]

Proposition 2. *Under a single contract cost-based regulation with quantity restriction, the optimal input level $\bar{x}^*$ represents an expected average of the first best solutions of the four possible types, adjusted by some upwards distortion in case of $\lambda > 0$. As an expected average, it lies between the extreme types' first-best input levels, i.e. $x_{ll}^{fb} < \bar{x}^* < x_{hh}^{fb}$.*

Proof. See Appendix. □

[18]The system operator for England and Wales and the electric distribution companies in the UK are the only two examples for menus of contracts being applied in regulatory practice Joskow (2014).

[19]Note that the solution for a pure cost-based regulation without quantity restriction would simply reimburse the costs of the observable input. This would incentivize the firm to choose infinitely high values of x (known as the gold-plating effect). Assuming that the regulator restricts her set of choices by an upper level of $\bar{x} = x_{hh}^{fb}$ in order to limit excessive (socially costly) rents, all types would then choose this level. In contrast to this *very* simple approach, the regulatory regime considered in this section makes use of being able to use the observable input x as a contracting variable.

Proposition 3. *Compared to a single contract, the regulatory approach based on a menu of contracts is superior with respect to expected social welfare.*

Proof. It is easy to show that the optimal solution of the single contract is a feasible solution of the menu of contracts problem. Due to the fact that the solution for the menu of contracts, as stated in Proposition 1, is both optimal and different from the one in Proposition 2, it must necessarily be superior. □

As stated in Proposition 3, the solution of the single contract regime is always inferior to the one obtained with the menu of contracts. Nevertheless, the characteristics of the different regimes can be compared and deserve a closer look. We contrast the outcome of the optimal menu of contracts with the one of the single contract regime considering three aspects: input levels, cost-efficiency of the input levels, and rents of the firm.

Input levels for the different types have been characterized in Proposition 1 for the menu of contract, stating that all types besides the *ll*-one are distorted above first best levels. According to Proposition 3, the optimal input level for the single contract regime, $\bar{x}^*$, represents an expected average of the first best solution of the four possible types, adjusted by some upwards distortion in case of $\lambda > 0$. Hence, chosen input levels are generally different. However, $\bar{x}^*$ may get close to x^*_{hl} in case of λ being large and μ small (i.e., for a high probability of realizing a high isoquant). At the same time, it will never be as high as x^*_{hh}, due to $\bar{x}^* < x^{fb}_{hh} < x^*_{hh}$.

Cost-efficiency of the input levels is closely connected to the input levels and their deviation from the first-best optimal solution. The optimal menu of contracts approaches first-best cost-efficiency of the input levels for $\lambda \to 0$, as input levels then converge towards first-best levels, i.e., $\{x^*_{ij}, y^*_{ij}\} \to \{x^{fb}_{ij}, y^{fb}_{ij}\}$. In contrast, cost-efficiency is poor for the single contract regime under this condition. However, first-best input levels may also be reached, but only under very restrictive conditions, namely if $\lambda \to 0$ *and* the occurrence probability for one specific type is particularly large (e.g., if $\mu, \nu \to 1$). Type-specific as well as expected cost-efficiency of input levels is (only) then approaching first-best optimality for both contracting frameworks. For the general case of $\lambda \geq 0$, it is clear that cost-efficiency of the input levels is inferior for the *ll* type in the single contract regime, while the ordering is ambiguous for all other types, depending on the optimal choice of $\bar{x}^*$ in comparison to x^*_{ij}.

Regarding **rents of the firm**, remember that they are only an issue for social welfare if there are shadow costs of public funding, i.e., if $\lambda > 0$. Then, however, the well known trade-off for rent-extraction and efficiency becomes relevant. For both contracting regimes, the rent of the inefficient *hh* type is set to zero. Moreover, for both regimes it holds true that $0 = R^*_{hh} < R^*_{hl} \ll R^*_{lh} < R^*_{ll}$ (respectively, $0 = \bar{R}^*_{hh} < \bar{R}^*_{hl} \ll \bar{R}^*_{lh} < \bar{R}^*_{ll}$), if isoquant variation is more relevant than cost variation. For the rent of specific types, we find that $R^*_{hl} < \bar{R}^*_{hl}$, while the ordering of other types' rent is generally ambiguous. Interestingly, however, if $g_i'(x)$ is small in the relevant range, $R^*_{ij} < \bar{R}^*_{ij}$ for all i, j.

Based on the above comparative statics, a singular interesting constellation can be identified for which the two contracting frameworks effectively approach each other.

Proposition 4. *For λ being large and μ small, the performance of the single contract is close to the one of the menu of contracts.*

Proof. For λ being large, $\bar{x}^*$ is distorted upwards (see Proposition 2), while $x^*_{hl} \approx x^{fb}_{hl}$ for μ small. Hence, in this case, $\bar{x}^* \approx x^*_{hl}$. Moreover, due to the fact that we consider Case (A) where cost uncertainty is relatively low, we know that the upwards distortion of x^*_{hh} is low (see Equation (29)), such that x^*_{hl} is not far from x^*_{hh}. Under these conditions, $\bar{C}^* \approx C^*$. □

Transferring Proposition 4 to our example of electricity transmission services and the regulation of the German TSOs, one may indeed come to the conclusion that the practically applied non-Bayesian regulatory approach could be close to the optimal second-best strategy. In fact, a high overall input level appears to be likely due to the strongly changing supply infrastructure, while ongoing intense discussions about the burden of electricity costs and grid tariffs for consumers could indicate high shadow costs of public funding. In the end, however, reasons for the chosen regulation are probably manifold, and might also include an explicit disutility of grid expansion, a commitment problem,[20] or the prohibitively high costs of implementing a 'complicated' regulatory regime (Armstrong and Sappington, 2007).

5 Conclusion

We considered a regulated firm providing a non-marketed output with substitutable inputs. We presented the optimal Bayesian regulation in terms of a menu of contracts when the regulator faces information asymmetries regarding the aggregated input level needed to provide the output as well as the realized optimal marginal rate of substitution between the inputs. Finally, the optimal Bayesian regulation was compared to a simpler non-Bayesian approach which appears to be closer to regulatory practice.

[20]Noticeably, a commitment problem of the regulator might impede the implementation of an incentive-based approach, which would be welfare-superior compared to a cost-based regulation. If the firm gets an unconditional payment representing the pay-off of the hh-type, i.e., $\tilde{T} = c^x x^{fb}_{hh} + c^y_h g_h(q, x^{fb}_{hh})$, she will realize first-best input quantities $\{x^{fb}_{ij}, y^{fb}_{ij}\}$. In this case, the realized rent of the firm becomes $R_{ij} = c^x x^{fb}_{hh} + c^y_h g_h(q, x^{fb}_{hh}) - c^x x^{fb}_{ij} - c^y_j g_i(q, x^{fb}_{ij})$. However, due to the (observable) separation of types via the realized input x, the regulator might be tempted to adjust the regulatory contract ex-post, and hence, jeopardize the regulatory success if the firm anticipates this behavior.

We found that in the optimal Bayesian regulation, the first best solution cannot be achieved under the considered information asymmetries and shadow costs of public funding. This implies a strictly positive rent for the firm. The second best solution that we then characterized depends on the relative importance of the information asymmetries. However, the most efficient type is always set to first best, while the levels of the observable input are distorted upwards for all other types. At least three types can always be separated, while bunching of two types may be unavoidable in case of a very asymmetric distribution of costs or very flat isoquants. These results are structurally similar to the solutions for multi-dimensional adverse selection problems in the literature (e.g. Lewis and Sappington (1988b), Armstrong (1999) or Aguirre and Beitia (2004)). However, in contrast to existing results, our model explains upwards distortions of input levels rather than prices. Hence, we obtained important insights regarding the optimal mechanism design in the context of a regulated monopolistic firm producing a non-marketed good with multi-dimensional inputs.

The comparison to a single contract cost-based approach, as it is often applied in regulatory practice, showed that the menu of contracts is welfare superior. However, there are situations in which the performance of the approaches converge, namely if the overall input level probably needs to be high, and shadow costs of public funding are large. Given our motivating example of electricity transmission services and the current situation, e.g., in Germany, these circumstances may indeed prevail, possibly explaining the gap between the theoretically optimal Bayesian approach and the simpler non-Bayesian regulation applied in practice.

Lastly, we note that our general approach as well as our insights might also be applicable to other industries that show similar characteristics, such as public works or administrative services. Besides investigating such areas of application, future research could relax the limited liability assumption and hence, allow for a shut down of firms. Another expansion could allow the good to be marketed, which would trigger a demand reaction of the regulator (or consumers) and possibly lead to interesting variations of the conclusions derived in this paper.

References

Aguirre, I. and A. Beitia (2004). Regulating a monopolist with unknown demand: Costly public funds and the value of private information. *Journal of Public Economic Theory 6*(5), 693–706.

Armstrong, M. (1999). Optimal regulation with unknown demand and cost functions. *Journal of Economic Theory 84*, 196–215.

Armstrong, M. and D. Sappington (2007). Chapter 27 recent developments in the theory of regulation. In M. Armstrong and R. Porter (Eds.), *Handbook of Industrial Organization*, pp. 1557–1700. Elsevier.

Baron, D. P. and R. B. Myerson (1982). Regulating a monopolist with unknown costs. *Econometrica: Journal of the Econometric Society 50*(4), 911–930.

Caillaud, B., R. Guesnerie, P. Rey, and J. Tirole (1988). Government intervention in production and incentives theory: A review of recent contributions. *The RAND Journal of Economics 29*(1), 1–26.

Dana, J. D. (1993). The organization and scope of agents: Regulating multiproduct industries. *Journal of Economic Theory 59*, 288–310.

Joskow, P. L. (2014). *Incentive Regulation in Theory and Practice: Electric Transmission and Distribution Networks*, Chapter 5. University of Chicago Press.

Laffont, J.-J. and D. Martimort (2002). *The theory of incentives - the principial-agent model*. Princeton University Press.

Laffont, J.-J. and J. Tirole (1993). *A Theory of Incentives in Procurement and Regulation.* MIT Press.

Lewis, T. R. and D. E. Sappington (1988a). Regulating a monopolist with unknown demand. *The American Economic Review 78*(5), 986–998.

Lewis, T. R. and D. E. Sappington (1988b). Regulating a monopolist with unknown demand and cost functions. *The RAND Journal of Economics 19*(3), 438–457.

Monitoringbericht (2013). Bundesnetzagentur, Bundeskartellamt.

Myerson, R. B. (1979). Incentive compatibility and the bargaining problem. *Econometrica: Journal of the Econometric Society 47*(1), 61–73.

Netzentwicklungsplan (2013). Netzentwicklungsplan Strom 2013 - Zweiter Entwurf der Übertragungsnetzbetreiber. 50Hertz Transmission, Amprion, TenneT TSO, TransnetBW.

Sappington, D. (1983). Optimal regulation of a multiproduct monopoly with unknown technological capabilities. *The Bell Journal of Economics 14*(2), 453–463.

Schmalensee, R. (1989). Good regulatory regimes. *The RAND Journal of Economics 20*(3), 417–436.

Appendix

Proof of Proposition 1

Proof. (i) Under the constraints considered binding for Case *(A)* – as discussed and shown in Figure 3 – the social cost function (4) becomes

$$
\begin{aligned}
C = \\
&\mu\nu\left[\lambda\left(g_h(x_{hh})(c_h^y - c_l^y) + c_l^y g_h(x_{hl}) - c_h^y g_l(x_{hl}) + g_l(x_{lh})(c_h^y - c_l^y)\right) + (1+\lambda)\left(c^x x_{ll} + c_l^y g_l(x_{ll})\right)\right] \\
&+ \mu(1-\nu)\left[\lambda\left(g_h(x_{hh})(c_h^y - c_l^y) + c_l^y g_h(x_{hl}) - c_h^y g_l(x_{hl})\right) + (1+\lambda)\left(c^x x_{lh} + c_h^y g_l(x_{lh})\right)\right] \\
&+ (1-\mu)\nu\left[\lambda\left(g_h(x_{hh})(c_h^y - c_l^y)\right) + (1+\lambda)\left(c^x x_{hl} + c_l^y g_h(x_{hl})\right)\right] \\
&+ (1-\mu)(1-\nu)\left[(1+\lambda)\left(c^x x_{hh} + c_h^y g_h(x_{hh})\right)\right]. \qquad (25)
\end{aligned}
$$

To derive the optimal observable input levels, we need to derive the above equation with respect to each of the four possible x_{ij}. Minimizing C with respect to x_{ll} yields

$$g_l'(x_{ll}^*) = -\frac{c^x}{c_l^y}, \qquad (26)$$

which implies that $x_{ll}^* = x_{ll}^{fb}$. Derivations of C with respect to x_{lh}, x_{hl} and x_{hh} take the following forms:

$$\frac{\partial C}{\partial x_{lh}} = \underbrace{\mu\nu\lambda(c_h^y - c_l^y)g_l'(x_{lh})}_{<0} + \underbrace{\mu(1-\nu)(1+\lambda)(c^x + c_h^y g_l'(x_{lh}))}_{\substack{=0 \text{ for } x_{lh}=x_{lh}^{fb} \\ <0 \text{ for } x_{lh}<x_{lh}^{fb} \\ >0 \text{ for } x_{lh}>x_{lh}^{fb}}} \qquad (27)$$

$$\frac{\partial C}{\partial x_{hl}} = \underbrace{\mu\lambda(c_l^y g_h'(x_{hl}) - c_h^y g_l'(x_{hl}))}_{<0} + \underbrace{(1-\mu)\nu(1+\lambda)(c^x + c_l^y g_h'(x_{hl}))}_{\substack{=0 \text{ for } x_{hl}=x_{hl}^{fb} \\ <0 \text{ for } x_{hl}<x_{hl}^{fb} \\ >0 \text{ for } x_{hl}>x_{hl}^{fb}}} \qquad (28)$$

$$\frac{\partial C}{\partial x_{hh}} = \underbrace{(\mu + (1-\mu)\nu)\lambda g_h'(x_{hh})(c_h^y - c_l^y)}_{<0} + \underbrace{(1-\mu)(1-\nu)(1+\lambda)(c^x + c_h^y g_h'(x_{hh}))}_{\substack{=0 \text{ for } x_{hh}=x_{hh}^{fb} \\ <0 \text{ for } x_{hh}<x_{hh}^{fb} \\ >0 \text{ for } x_{hh}>x_{hh}^{fb}}}. \qquad (29)$$

From Equation (27), we see that $\frac{\partial C}{\partial x_{lh}}$ is strictly smaller than 0 for $x_{lh} = x_{lh}^{fb}$ and monotonically increasing in x_{lh}, which implies that $x_{lh}^* > x_{lh}^{fb}$ must always hold. The same logic applies for x^*_{hl} and x^*_{hh}.

(ii) From the fact that $x^{fb}_{ll} < x^{fb}_{lh}$ and the strict upwards distortion of all other types, it follows that the *ll*-type can always be separated. In order to investigate whether the types *lh*, *hl* and *hh* can be separated or need to be bunched, we proceed as follows: For each of the possible pairs $lh - hl$, $hl - hh$ and $lh - hh$, we check the derivative of C with respect to the former type at the optimal level of x^* of the latter type (derived from the first order condition). If the change in C is greater than 0 we can conclude that we have already surpassed the optimal level of the former type, which then must be smaller than the optimal level of the latter type. In other words, we check the level of upwards distortion for the *lh*, *hl* and *hh* types while considering the necessary ordering of the types according to Lemma 2. For the pair *lh*-*hl*, we find that x^*_{lh} may surpass x^*_{hl} in case of $\nu \to 1$, while they are otherwise clearly separated from each. For the pair *hl*-*hh*, bunching may occur for $g'_l(q, x) \to 0$ together with c^y_l being large. Furthermore, we find that *lh*-*hh* can always be separated, implying that at most two types (i.e., either *lh*-*hl* or *hl*-*hh*) may be bunched under certain parameter constellations.

Lastly, it is straightforward to check that the remaining constraints are satisfied under the obtained solution of the relaxed problem. Hence, we have indeed obtained to optimal solution for the full regulatory problem we are facing in Case *(A)*. □

Proof of Proposition 2

Proof. Written explicitly, Equation (24) becomes

$$\begin{aligned}\bar{C} =&\mu\nu\left[\lambda\left(c^y_h g_h(\bar{x}) - c^y_l g_l(\bar{x})\right) + (1+\lambda)\left(c^x\bar{x} + c^y_l g_l(\bar{x})\right)\right]\\ &+ \mu(1-\nu)\left[\lambda\left(c^y_h g_h(\bar{x}) - c^y_h g_l(\bar{x})\right) + (1+\lambda)\left(c^x\bar{x} + c^y_h g_l(\bar{x})\right)\right]\\ &+ (1-\mu)\nu\left[\lambda\left(c^y_h g_h(\bar{x}) - c^y_l g_h(\bar{x})\right) + (1+\lambda)\left(c^x\bar{x} + c^y_l g_h(\bar{x})\right)\right]\\ &+ (1-\mu)(1-\nu)\left[(1+\lambda)\left(c^x\bar{x} + c^y_h g_h(\bar{x})\right)\right]. \end{aligned} \tag{30}$$

Deriving the above with respect to $\bar{x}$ yields, after a few calculations, $\mathbb{E}(g'_i(\bar{x}^*))\mathbb{E}(c^y_j) + c^x + \lambda(c^y_h g'_h(\bar{x}^*) + c^x) = 0$. Hence, for $\lambda = 0$, $\mathbb{E}(g'_h(\bar{x}^*)) = -\frac{c^x}{\mathbb{E}(c^y_j)}$. □

Two-dimensional asymmetric information, Case *(B)*: Cost variation large compared to isoquant variation

To solve the second case following from Lemma 2, we need to apply a different educated guess with respect to the binding constraints. However, we apply a similar reasoning as in Case *(A)*, but take account of the fact that now, cost variation is more relevant than isoquant variation. Hence, we choose a symmetric setting and imply incentive constraints (8b) ($ll \rightarrow hl$), (8h) ($hl \rightarrow lh$) and (8f) ($lh \rightarrow hh$) to be binding. Again, we assume the participation constraint of the *hh*-type to be binding. Figure 4 illustrates this setting.

After having determined the results and checked all remaining constraints, we find the setting of binding constraints as in Figure 4 indeed to be optimal for Case *(B)*. Results are summarized in the following Proposition 5.

Proposition 5. *For case (B),*

(i) Optimal regulation is achieved under the following set of observable input levels:

$$x_{ll}^* = x_{ll}^{fb} \tag{31}$$

$$x_{lh}^* \geq x_{lh}^{fb} \tag{32}$$

$$x_{hl}^* \geq x_{hl}^{fb} \tag{33}$$

$$x_{hh}^* \geq x_{hh}^{fb}, \tag{34}$$

while respecting $x_{ll}^* < x_{hl}^* \leq x_{lh}^* \leq x_{hh}^*$.

(ii) The most efficient (ll) type can always be separated. Moreover, separation of at least three types is always possible, while bunching of the hl and lh types is unavoidable in case of $\mu \rightarrow 1$.*The lh and hh types may be bunched in case of* c_l^y *being small and* $g_h'(q,x)$ *large.*

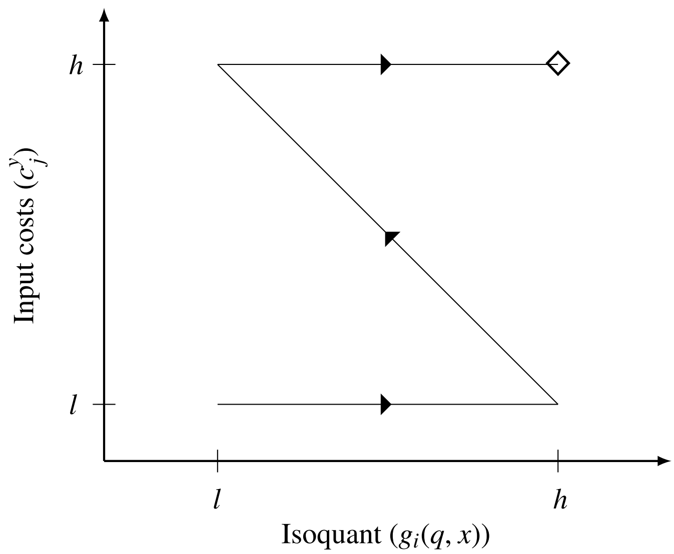

Figure 4: Constraints considered binding for Case *(B)*

Proof. (i) Under the constraints considered binding for Case *(B)* – as discussed and shown in Figure 4 – the social cost function (4) becomes

$$\begin{aligned}
C = \\
&\mu\nu\Big[\lambda\Big(c^y_h(g_h(x_{hh}) - g_l(x_{hh})) + c^y_h g_l(x_{lh}) - c^y_l g_h(x_{lh}) + c^y_l(g_h(x_{hl}) - g_l(x_{hl}))\Big) + (1+\lambda)\Big(c^x x_{ll} + c^y_l g_l(x_{ll})\Big)\Big] \\
&+ \mu(1-\nu)\Big[\lambda\Big(c^y_h(g_h(x_{hh}) - g_l(x_{hh}))\Big) + (1+\lambda)\Big(c^x x_{lh} + c^y_h g_l(x_{lh})\Big)\Big] \\
&+ (1-\mu)\nu\Big[\lambda\Big(c^y_h(g_h(x_{hh}) - g_l(x_{hh}))\Big) + c^y_h g_l(x_{lh}) - c^y_l g_h(x_{lh}) + (1+\lambda)\Big(c^x x_{hl} + c^y_l g_h(x_{hl})\Big)\Big] \\
&+ (1-\mu)(1-\nu)\Big[(1+\lambda)\Big(c^x x_{hh} + c^y_h g_h(x_{hh})\Big)\Big]. \qquad (35)
\end{aligned}$$

Minimizing C with respect to x_{ll} yields

$$g'_l(x^*_{ll}) = -\frac{c^x}{c^y_l}, \qquad (36)$$

which implies that $x^*_{ll} = x^{fb}_{ll}$. Derivation of C with respect to x_{lh}, x_{hl} and x_{hh} yields:

$$\frac{\partial C}{\partial x_{lh}} = \underbrace{\mu\lambda(c^y_h g'_l(x_{lh}) - c^y_l g'_h(x_{lh})}_{<0} + \underbrace{\mu(1-\nu)(1+\lambda)(c^x + c^y_h g'_l(x_{lh}))}_{\substack{=0 \text{ for } x_{lh}=x^{fb}_{lh} \\ <0 \text{ for } x_{lh}<x^{fb}_{lh} \\ >0 \text{ for } x_{lh}>x^{fb}_{lh}}} \qquad (37)$$

$$\frac{\partial C}{\partial x_{hl}} = \underbrace{\mu\nu\lambda(c^y_l g'_h(x_{hl}) - c^y_l g'_l(x_{hl}))}_{<0} + \underbrace{(1-\mu)\nu(1+\lambda)(c^x + c^y_l g'_h(x_{hl}))}_{\substack{=0 \text{ for } x_{hl}=x^{fb}_{hl} \\ <0 \text{ for } x_{hl}<x^{fb}_{hl} \\ >0 \text{ for } x_{hl}>x^{fb}_{hl}}} \qquad (38)$$

$$\frac{\partial C}{\partial x_{hh}} = \underbrace{(\mu + (1-\mu)\nu)\lambda c^y_h(g'_h(x_{hh}) - g'_l(x_{hh}))}_{<0} + \underbrace{(1-\mu)(1-\nu)(1+\lambda)(c^x + c^y_h g'_h(x_{hh}))}_{\substack{=0 \text{ for } x_{hh}=x^{fb}_{hh} \\ <0 \text{ for } x_{hh}<x^{fb}_{hh} \\ >0 \text{ for } x_{hh}>x^{fb}_{hh}}}. \qquad (39)$$

From Equation (37), we see that $\frac{\partial C}{\partial x_{lh}}$ is strictly smaller than 0 for $x_{lh} = x^{fb}_{lh}$ and monotonically increasing in x_{lh}, which implies that $x^*_{lh} > x^{fb}_{lh}$ must always hold. The same logic applies for x^*_{hl} and x^*_{hh}.

(ii) From $x^{fb}_{ll} < x^{fb}_{lh}$ and the strict upwards distortion of all other types, it follows that the ll-type can always be separated. x^*_{hl} may surpass x^*_{lh} in case of $\mu \to 1$. If the low costs

c_l^y are small and $g_h'(q, x)$ becomes large, *lh* and *hh* types may need to be bunched, without impacting the separation of the other types.

The remaining constraints are satisfied under the obtained solution. □

As in Case (A), the first best solution can be obtained for $\lambda = 0$, while the solution is second best and incurring an increasing level of inefficiency for increasing levels of λ. Also again, the most efficient type can always be separated, while bunching of the *hl* and *lh* types (*lh* and *hh* types) may occur for very high occurrence probability of low isoquants, or if $g_h(q, x)$ is very steep and c_l^y small.