The Scandinavian
Journal of Economics

Scand. J. of Economics 124(2), 457–487, 2022
DOI: 10.1111/sjoe.12468

Delegating pollution permits*

Tapas Kundu

Oslo Metropolitan University, NO-0130 Oslo, Norway
tapas.kundu@oslomet.no

Tore Nilssen

University of Oslo, NO-0317 Oslo, Norway
tore.nilssen@econ.uio.no

Abstract

We discuss the decision to delegate the regulation of pollution through sales of permits to a biased expert in a situation where the polluting firm has private information about its technology. We consider, in particular, constrained delegation where the government puts restrictions on the amount of pollution that the expert can sell permits for. We find that, in general, delegation is more likely if the firm is low-cost. This is not in line with the so-called uncertainty principle, which states that there is more delegation the more uncertainty the government faces.

Keywords: Bureaucracy; environmental regulation; pollution standards
JEL classification: *D*82; *H*23; *L*51; *Q*58

1. Introduction

One way to regulate a polluting industry is to require a firm to purchase pollution permits sold by the government in order to be allowed to pollute. With such permits, the government can keep the pollution down, at the same time as the proceeds from the sale of these permits provide the government with needed revenue. The task of regulating polluting firms through permits becomes complicated in situations where firms have private information about their production technology. A standard solution to such an asymmetry of information is for the government to offer a menu of combinations of transfers and pollution permits, such that the firm self-selects in its choice of combination according to its technology, creating a distortion in the pollution level of a high-cost firm and an information rent for a low-cost firm.

In several countries, regulation is left to independent regulatory agencies, to an extent that has been increasing over the last few decades (Gilardi, 2009). Such independent agencies open up for the hiring of

*We are grateful for helpful comments from three anonymous referees, as well as from Svenn Jensen and Igor Letina.

biased experts to carry out dealing with the regulatee. In this paper, we discuss whether leaving the sale of pollution permits to biased experts is an improvement over the government itself doing it. The advantage of hiring an expert is that she has knowledge of the technology of the firm; in the present analysis, we discuss the case where she has perfect such knowledge and therefore does not have to spend resources at all on distorting the pollution level for a high-cost firm or offering an information rent to a low-cost firm that the government would without the expert. The disadvantage of hiring an expert is that she typically has preferences that are not perfectly aligned with those of the government; in the present analysis, we consider the case when the hired expert puts even more weight on having a low pollution level than the government does.[1]

In order to get the most benefits out of delegation, the government will often use the statutes that delegate authority to an independent regulator to also put limits on that regulator's actions. This way, the government creates what Gailmard (2009) calls "windows of discretion", and he mentions examples in the United States from the Interstate Commerce Commission, the Federal Trade Commission, and – of particular interest for us – the Environmental Protection Agency.

We model here this window of discretion as a requirement by the government that the independent bureaucrat, if hired, choose the firm's number of pollution permits from an interval. There are two types of technology for a firm, low-cost and high-cost, which is known to the firm, and to the bureaucrat if she is hired, but which is not known to the government. In this setting, we find that there are two different ways constrained delegation works. One is what we call *weak delegation*, which implies capping the bias of the expert so that she is constrained from selling fewer pollution permits to the firm than the first-best quantity for the low-cost type. By imposing such a restriction, the government ensures that, by hiring the expert, the firm will have the first-best level of pollution if it has low costs. But if the probability of the firm being high-cost is high and/or the distortion from the first-best of the expert's offer to a high-cost firm is large, then weak delegation becomes too costly. We show in our analysis that the government then might choose to resort to *strict delegation*, where in practice the government pinpoints the pollution level of the firm based on an *ex ante* assessment of its cost. In the language of Gailmard

[1]A further amendment, which we do not discuss here, would be to make the pollution permits tradable. There are mainly two reasons for sticking to the case of non-tradable permits. First, our concern is mainly cases of localized pollution, such as pollution of lakes, rivers or coastlines, where a permit to pollute cannot be transferred to other locations, so that tradability is not so crucial. Secondly, as argued by, for example, Lewis (1996), there might be political restrictions that make it difficult to introduce tradable permits.

(2009), the window of discretion can be wide, when weak delegation is the optimum for the government; it can be narrow when strict delegation is the optimum; and it can be closed, when no delegation is the optimum thing for the government to do.

The need to resort to strict delegation comes out of the government's informational disadvantage and becomes a solution in between weak delegation, with few restrictions in the hired expert, and no delegation to an expert at all. Strict delegation has two implications that make our results on delegation of pollution permits stand out relative to related analyses. First, with strict delegation, there is distortion at the top, in the sense that the firm not even when it is low-cost is offered its first-best amount of permits. This aspect of our regulatory regime makes our model contrast with most other models of regulation under asymmetric information, where the most effective type's quantity is not distorted.[2] Also, note that this distortion at the top occurs because delegation, even when strict, turns out to be better for the government than no delegation at all, when there would be no distortion at the top.

Secondly, strict delegation occurs mainly when the probability that the firm is high-cost is high. It follows that the government does not necessarily respond to reduced uncertainty by allowing more delegation. Rather, if the government becomes more certain the firm is high-cost, it is more likely to allow less delegation, that is, either strict or no delegation rather than weak delegation. This contrasts with the so-called uncertainty principle in studies of delegation in government, stating that the government is more prone to delegate tasks when uncertainty is large.

The model we develop to discuss these issues includes a firm with a polluting production that has private information about its production technology; in particular, the firm has either high or low production costs. The firm needs permits to pollute, which are sold by the government. The government, in turn, lacks information about the firm's technology and can offer a menu of contracts (i.e., combinations of number of pollution permits and price for the permits) in order to solve its regulation problem as effectively as possible. Alternatively, it can leave the permits to an expert who has perfect information about the firm's technology but also is less interested in allowing pollution than the government is. In particular, we analyze the following game. First, the government decides whether or not to hire an expert to carry out the sale of pollution permits and, if hiring an expert, which constraints to put on the expert's actions. If an expert is hired, then the expert sells pollution permits to the firm, knowing its

[2]Other reasons found in the literature for a distortion at the top include countervailing incentives (Lewis and Sappington, 1989), renegotiation (Laffont and Tirole, 1990, 1993), and short-term contracts (Laffont and Tirole, 1988, 1993).

technology and therefore not having to resort to offering a contract menu. If an expert is not hired, then the government presents the firm with its menu of contracts. Finally, the firm buys pollution permits from either the expert or the government, depending on whether or not the government has delegated the permit sales to the expert, after which it uses its permits to produce its goods and sells them.

Our work relates to work on regulating a polluting firm under private information; see Lewis (1996) for a review of the early literature. We add to this by discussing how regulation will be changed by the hiring of a biased expert to do the regulation. One important finding in our analysis is that there might be "distortions at the top". This happens if, in optimum, the expert is subjected to strict delegation, when different types have the same pollution level.

Our work is also related to work, particularly in political science, discussing how and when a government delegates tasks to independent bureaucrats; see, for example, Epstein and O'Halloran (1999), Huber and Shipan (2006), and Gailmard and Patty (2012). Two factors emphasized in this literature that are particularly pertinent to our paper are the ally principle and the uncertainty principle. According to the ally principle, the government is more interested in delegating tasks the less biased the bureaucrat is. This shows up also in our framework. The uncertainty principle says that the government is more interested in delegating tasks the more uncertain it is about the effects of its decisions. An important finding in our analysis is that this principle does not show up in the present framework of delegating the task of selling pollution permits. The uncertainty here is with respect to the firm's technology. This uncertainty would be small if the government were almost sure that the firm is either low-cost or high-cost. But we find that delegation is more likely with a low-cost firm.

Our work connects to studies in organizational economics on how to put constraints on a hired biased expert, that is, an agent who is charged with carrying out tasks on a principal's behalf and who is both better informed than the principal and has interests that are misaligned with that of the principal; see, for example, the seminal study of Holmström (1984) and the recent work of Alonso and Matouschek (2008) and Amador and Bagwell (2013), and see Gibbons et al. (2013) for a survey of the literature. Our situation differs from those previously studied in this literature in that the task to be delegated is the design of contracts. Our delegation problem has two distinct features. First, the task to be delegated is two-dimensional, because the contracts we study are two-dimensional. Secondly, the bias of the hired expert is type-dependent: the bureaucrat's interest in avoiding pollution depends on the firm's technology, which is private information. In previous analyses of the delegation of multi-dimensional tasks, such as Frankel (2016), the expert's bias is assumed to be independent of the private information, an assumption that makes it possible to turn the multi-dimensional problem into a one-dimensional one. In our case of delegation of contract design with a type-dependent bias, we are able to obtain this conversion to a one-dimensional problem because the regulated firm's participation constraint is binding for the high-cost type in the optimal contract menu. We note, finally, that the distinction between weak and strict delegation that grows out of our analysis resembles the distinction made by Melumad and Shibano (1991) between communication-dependent and communication-independent decision rules in organizations.

The introduction of experts into models of environmental regulation is also done by Porteiro (2008) and Voss and Lingens (2018). Porteiro (2008) focuses on the role of experts in the acquisition of information. In his analysis, the polluting firm might not know its technology, and he discusses how the presence of an information-collecting expert can affect the chance of such information being acquired. In particular, he compares an unbiased expert, whose costs of information collection has to be covered by the government, and a biased environmentalist who covers such costs herself. In contrast, the expert is already informed in our analysis, and we focus on how to use constrained delegation to balance the knowledge and bias of the expert in order to obtain the most efficient regulation. Voss and Lingens (2018) discuss the incentive contract between the government and a biased regulator. In contrast, our focus is on the regulatory contract between the firm and the government, and we discuss how the government hiring a biased expert through constrained delegation can improve on that contract when incentive contracts for the expert are not feasible.

Finally, Kundu and Nilssen (2020) discuss the delegation of regulation to an independent bureaucrat in a setting that is related to, but still distinct from, the present one. There are especially two ways in which the two analyses differ. First, in Kundu and Nilssen (2020), the regulatory task to be delegated is one of procurement, where the government buys products from the firm, so that transfers are from the government to the producing firm. This means that the firm in that setting likes transfers while the government does not like them. In the situation discussed presently, the firm's production is sold to the market and the firm buys pollution permits from the government. In this case, transfers are from the firm to the government, and we have the opposite relationship: now, the government likes transfers while the firm does not like them.[3] Secondly, in Kundu and Nilssen (2020), the regulatory contract is about the quality of the

[3]The distinction between the frameworks of permits and procurement closely resembles the one made in Caillaud et al. (1988) between marketed and non-marketed goods. The production cost is covered with private funds in the permits framework and with public funds in the procurement

firm's product, rather than the firm's pollution. That quality is affected by both the firm's technology and the firm's efforts, in line with Laffont and Tirole (1993), so that contracts are three-dimensional, whereas in the present model there are no efforts. The bureaucrat, when hired, is biased in favor of high quality in Kundu and Nilssen (2020) and in favor of low pollution in the present work. In our view, it is at the outset not clear how a change of the direction of transfers would affect the analysis. Despite these differences, our findings share some crucial common features: the uncertainty principle does not hold, but at the same time constrained delegation shows up as an effective way of dealing with biased experts in the bureaucracy. One difference between the two analyses concerns the prevalence of no delegation when the probability of the firm having low costs is very high: this prevalence is higher in the present permits setting than in the procurement setting of Kundu and Nilssen (2020); we discuss this issue in more detail towards the end of Section 4.

The paper is organized as follows. In Section 2, we present the model. In Section 3, we analyze the equilibrium of the model for various benchmark cases, such as when the government has complete information, and when the government can delegate but only without constraints. In Section 4, we introduce constrained delegation and the concepts of weak and strong delegation, and we solve for the equilibrium outcome of mode of delegation and regulatory contract. We conclude in Section 5, while all proofs are relegated to the Appendix.

2. The model

We consider the problem of regulating a polluting firm under asymmetric information by requiring that the firm buy permits in order to be allowed to pollute. Following Laffont (1994) and Boyer and Laffont (1999), we assume that the firm's level of pollution is observable and verifiable. The firm provides a good that gives a positive surplus S. The cost of production is given by

$$C(\theta, d) = \theta(K - d),$$

where $K > 0$ is a constant, $d \in [0, K]$ is the observable and verifiable pollution level chosen by the firm, and θ is a cost characteristic that is the firm's private information. With $\partial C/\partial\theta > 0$, a high θ implies a high cost

framework. If the costs of raising public funds and private funds differ, the two frameworks do not yield the same social welfare. In particular, in the presence of positive (and higher) costs of raising public funds, the value of social welfare will be lower in the procurement framework than in the permits framework. This observation is also noted by Caillaud et al. (1988, see their section 4).

and therefore low cost efficiency. For simplicity, we make the assumption that θ can take only two values, $\{\underline{\theta}, \bar{\theta}\}$, with $0 < \underline{\theta} < \bar{\theta} < K$.[4] Let $\nu \in (0, 1)$ be the probability that the firm is low-cost with type $\theta = \underline{\theta}$, so that the expected value of θ is $E_\theta \theta := \nu \underline{\theta} + (1 - \nu)\bar{\theta}$.

The firm produces the good and sells it at a price $p(\theta)$. Because the price is chosen by a firm with private information, it is in principle a function of the firm's type θ. We assume, however, that the firm, whichever type it is, is able to extract the full surplus value when selling it on the market. This implies that $p(\theta) = S$, for $\theta \in \{\underline{\theta}, \bar{\theta}\}$. The government regulates production by issuing pollution permits, which the firm buys in order to be able to pollute. The transaction between the government and the firm on permits can be viewed as the government offering a contract $\alpha = (t, d) \in A := \mathbb{R}_+ \times [0, K]$: by providing a transfer t to the government, the firm can keep its level of pollution at d. With this contract, the firm's payoff is

$$U_P(\theta, \alpha) = p(\theta) - \theta(K - d) - t = S - \theta(K - d) - t. \tag{1}$$

Society is adversely affected by the firm's pollution, with a disutility given by $d^2/2$.[5] Moreover, along the lines of Laffont and Tirole (1993), we assume that the government faces a marginal cost of public funds that exceeds one. The government receiving an amount from the firm means that distortionary taxes can be reduced elsewhere in the economy by that same amount. Specifically, for every unit of transfer, the government benefits $1+\lambda$ from it, where $\lambda > 0$.[6] Thus, the government's payoff from a contract α is the social value of the project:

$$\begin{aligned} U_G(\theta, \alpha) &= \left[S + (1 + \lambda)t - \frac{1}{2}d^2 - p(\theta) \right] + U_P(\theta, \alpha) \\ &= S - \theta(K - d) - \frac{1}{2}d^2 + \lambda t. \end{aligned} \tag{2}$$

We assume that $S \geq K^2/2$, ensuring that the social value of the project is non-negative even at maximum pollution and zero transfer.

[4]Having only two types is a simplification that facilitates the analysis. A continuous-type space would allow a richer analysis. For example, we conjecture that, in a setting with a continuous-type space, forms of delegation in between weak delegation and strict delegation will occur. An exploration of this is left for future research.
[5]The specific functional forms we use for the firm's cost of production and society's cost of pollution are helpful in making the model analytically tractable and the discussion of the roles of weak and strict delegation instructive. We do not believe our results are very dependent on the functional forms used.
[6]The assumption that $\lambda > 0$ is needed to ensure that the equilibrium we study here is unique. For a discussion of the notion that the marginal cost of public funds, here modeled as $1 + \lambda$, is greater than 1, see, for example, Ballard and Fullerton (1992).

The government can delegate the regulatory decision-making to an independent regulator, or a bureaucrat. We assume that the bureaucrat is informed about the firm's cost. She can therefore implement a type-contingent regulatory policy. If the government delegates, then the bureaucrat has the authority to choose a regulatory policy according to her own preferences. We assume that the bureaucrat is intrinsically motivated to keep pollution low. In particular, we assume that her payoff is given by

$$U_B(\theta, \alpha) = U_G(\theta, \alpha) - \beta \frac{d^2}{2}, \tag{3}$$

where $\beta \geq 0$ measures the bureaucratic drift. The higher β is, the more the bureaucrat is concerned about pollution.

The game proceeds as follows.

Stage 1. The government decides whether or not to delegate the decision-making authority to an independent bureaucrat. If it does not delegate, then the authority remains with the government.

Stage 2. The firm learns its type θ, which can be either $\underline{\theta}$ with probability ν or $\bar{\theta}$ with probability $1 - \nu$. The bureaucrat also learns the firm's type at zero cost.

Stage 3. The player with decision-making authority determines the permit contracts.

Stage 4. The firm chooses whether or not to accept an offered contract. If it accepts, then it produces and sells its production. Payoffs are realized. The game ends.

We study the perfect Bayesian equilibrium of the game. We solve the game by backward induction.

3. Analysis

The analysis at stage 4 is trivial. The firm only accepts contracts that satisfy the participation constraint $U_P(\theta, \alpha) \geq 0$ and extracts the full surplus.

At stage 3, the regulatory contract is determined. We first describe the contract $\alpha_{GI}(\theta) = (t_{GI}(\theta), d_{GI}(\theta))$ that the government chooses if it has complete information about θ. The contract for type θ solves the following problem:

$$\max_{\alpha \in A} \; U_G(\theta, \alpha) \quad \text{subject to} \quad U_P(\theta, \alpha) \geq 0. \tag{4}$$

Note that, at the solution, the firm's participation constraint is binding, which gives $U_P(\theta, \alpha) = 0$. After replacing U_P in equation (4), we find that the first-order condition with respect to d is

$$(1 + \lambda)\theta - d = 0. \tag{5}$$

This equation states that, in optimum, marginal social damages from pollution, d, equal marginal social abatement costs, $(1+\lambda)\theta$. Because $\lambda > 0$, this implies that, in optimum, marginal social damages from pollution are larger than marginal private abatement costs, θ. We see from equation (5) that the optimal pollution level is increasing in λ; this is because the government's interest in receiving a transfer from permits increases with λ.

If $\lambda \geq (K/\bar{\theta}) - 1$, then the pollution is at the maximum level K, and there is also a maximum transfer: $\alpha_{GI}(\theta) = (S, K)$. In this case, essentially, the government lets the producer go unregulated, with full pollution and no profit. For the analysis below, we restrict our attention to cases where government does not offer such a no-regulation contract to any type of firms under complete information. Formally, we impose the following restriction.

Assumption 1. $\lambda \leq (K/\bar{\theta}) - 1$.

With this assumption, the government sets the pollution level

$$d_{GI}(\theta) = (1 + \lambda)\theta. \tag{6}$$

From the binding participation constraint, we obtain

$$t_{GI}(\theta) = S - \theta[K - d_{GI}(\theta)]. \tag{7}$$

It follows from equation (6) that a high-cost firm is allowed to pollute more in equilibrium.

The government's expected payoff under full information is given by

$$U_G^{FI} = \nu U_G(\underline{\theta}, \alpha_{GI}(\underline{\theta})) + (1 - \nu)U_G(\bar{\theta}, \alpha_{GI}(\bar{\theta})). \tag{8}$$

3.1. No delegation

With no delegation at stage 1, the uninformed government offers an incentive-compatible pair of contracts $(\underline{\alpha}, \bar{\alpha}) = ((\underline{t}, \underline{d}), (\bar{t}, \bar{d}))$ to the firm at stage 3. The contract pair solves the following problem:

$$\max_{\underline{\alpha}\in A,\bar{\alpha}\in A} \nu U_G(\underline{\theta},\underline{\alpha}) + (1-\nu)U_G(\bar{\theta},\bar{\alpha}) \tag{9}$$

subject to

$$U_P(\bar{\theta},\bar{\alpha}) \geq U_P(\bar{\theta},\underline{\alpha}), \tag{ICH}$$
$$U_P(\underline{\theta},\underline{\alpha}) \geq U_P(\underline{\theta},\bar{\alpha}), \tag{ICL}$$
$$U_P(\bar{\theta},\bar{\alpha}) \geq 0, \tag{IRH}$$
$$U_P(\underline{\theta},\underline{\alpha}) \geq 0, \tag{IRL}$$

where equations (ICH) and (ICL) are the two firm types' incentive-compatibility constraints and equations (IRH) and (IRL) are their individual-rationality constraints. We denote the solution with subscript GN. Define $\triangle\theta := \bar{\theta} - \underline{\theta}$. The following lemma describes the contract pair.

Lemma 1. *Consider the case of no delegation. The optimal incentive-compatible contract pair* $(\alpha_{GN}(\underline{\theta}), \alpha_{GN}(\bar{\theta})) = ((t_{GN}(\underline{\theta}), d_{GN}(\underline{\theta})), (t_{GN}(\bar{\theta}), d_{GN}(\bar{\theta})))$ *that the government offers to the firm is given by*

$$d_{GN}(\underline{\theta}) = (1+\lambda)\underline{\theta}, \tag{10}$$
$$d_{GN}(\bar{\theta}) = \min\left\{(1+\lambda)\bar{\theta} + \lambda\frac{\nu}{1-\nu}\triangle\theta, K\right\}, \tag{11}$$
$$t_{GN}(\underline{\theta}) = S - \underline{\theta}[K - d_{GN}(\underline{\theta})] - \triangle\theta[K - d_{GN}(\bar{\theta})], \tag{12}$$
$$t_{GN}(\bar{\theta}) = S - \bar{\theta}[K - d_{GN}(\bar{\theta})]. \tag{13}$$

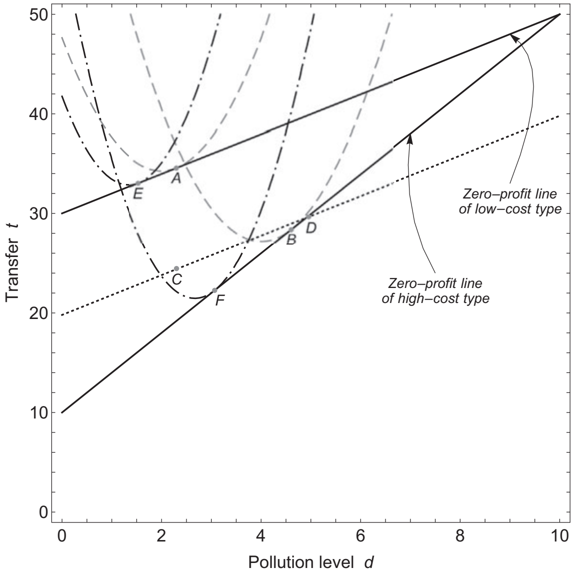
From (11), we have that $d_{GN}(\bar{\theta}) = K$ if ν is sufficiently large, in particular if

$$\nu \geq \nu^* := \frac{K - \bar{\theta} - \lambda\bar{\theta}}{K - \bar{\theta} - \lambda\underline{\theta}} \in [0, 1), \tag{14}$$

which is decreasing in λ. With $d = K$, we have $t = S$ and $C(\cdot, K) = 0$, so that both revenue and cost, and hence profit, equal zero. Thus, while Assumption 1 ensures that this does not happen under complete information, that assumption still allows for it under asymmetric information.

In order to understand the outcome in Lemma 1 better, consider Figure 1, depicting the contract space (i.e., the (d, t) space). The preferences of the firm are towards high pollution and low transfers (i.e., in the direction of the lower-right corner of the figure), while the government's preferences are towards little pollution and high transfers (i.e., in the direction of the upper-left corner of the figure). The straight lines depict contracts that give the firm zero profit, one line for each type; thus, these lines illustrate conditions (IRL) and (IRH) above. The full-information contracts, given in equations (6) and (7), are depicted as the points A and B in Figure 1. Each of them is the optimal contract for the government that satisfies individual

Figure 1. The optimal contracts in (d, t) space



Notes: Specification: $S = 50$, $K = 10$, $\bar{\theta} = 4$, $\underline{\theta} = 2$, $\lambda = 0.15$, $\nu = 0.5$, and $\beta = 0.5$.

rationality on the part of the firm; this is illustrated by the dashed curves in Figure 1, showing indifference curves for the government for contracts to the two firm types. We recognize the feature of the model that a high-cost firm pollutes more than a low-cost firm.

The contracts that prevail under asymmetric information, given in Lemma 1, are depicted as points C and D in Figure 1. Note that there is no distortion at the top, in that information asymmetry does not affect the pollution level of the low-cost type, as a comparison of equations (6) and (10) shows; in Figure 1, this means that point C is vertically below point A. There is, however, an information rent accruing to the low-cost type when $\nu < \nu^*$.[7] From equations (7) and (12), we find that this information rent equals

$$t_{GI}(\underline{\theta}) - t_{GN}(\underline{\theta}) = \triangle\theta[K - d_{GN}(\bar{\theta})].$$

This equation shows a trade-off for the government: it can lower the information rent to the low-cost type, which is costly, only by increasing the

[7]When $\nu \geq \nu^*$, we have $d_{GN}(\bar{\theta}) = K$ and $t_{GN}(\bar{\theta}) = S$, and there is no need to incentivize the low-cost type to keep away from the high-cost contract.

pollution level of the high-cost type, which also is costly. In equilibrium, the government do a bit of this trade-off, which is why the pollution level of the high-cost type is higher under asymmetric information than under complete information: contract D is to the right of contract B on the red zero-profit curve of the high-cost type in Figure 1.

The government's expected payoff under no delegation is given by

$$U_G^{ND} = \nu U_G(\underline{\theta}, \alpha_{GN}(\underline{\theta})) + (1-\nu)U_G(\bar{\theta}, \alpha_{GN}(\bar{\theta})). \tag{15}$$

3.2. Full delegation

If the government delegates at stage 1, then an informed bureaucrat, who has complete information about θ, chooses the contract $\alpha_{BI}(\theta) = (t_{BI}(\theta), d_{BI}(\theta))$ at stage 3. In the case of full delegation, the bureaucrat can choose any contract in A. The contract for type θ solves the following problem:

$$\max_{\alpha \in A} \; U_B(\theta, \alpha),$$
$$\text{subject to} \;\; U_P(\theta, \alpha) \geq 0. \tag{16}$$

We denote the solution with subscript BI. The following lemma, which follows straightforwardly, describes the optimal contract.

Lemma 2. *Assume that the government delegates the decision-making authority to a bureaucrat. The contract $\alpha_{BI}(\theta) = (t_{BI}(\theta), d_{BI}(\theta))$ that the bureaucrat offers to a producer of type $\theta \in \{\underline{\theta}, \bar{\theta}\}$ is given by*

$$d_{BI}(\theta) = \frac{(1+\lambda)\theta}{1+\beta}, \tag{17}$$
$$t_{BI}(\theta) = S - \theta[K - d_{BI}(\theta)]. \tag{18}$$

The bureaucrat's choice of pollution level is always below the government's choice because of her vested interest in reducing pollution. By setting pollution at a lower level, the bureaucrat increases the production cost, and thereby the compensatory transfer. In Figure 1, the bureaucrat's contracts to the low-cost and high-cost firm type are depicted as the points E and F, respectively. Because of the bureaucrat's knowledge of the firm's technology, the two points are on the firm's participation constraints. Because the bureaucrat is biased and dislikes pollution more than the government, points E and F are to the left of points A and B in Figure 1.

The government's *ex ante* expected payoff under full delegation is given by

$$U_G^{FD} = \nu U_G(\underline{\theta}, \alpha_{BI}(\underline{\theta})) + (1-\nu)U_G(\bar{\theta}, \alpha_{BI}(\bar{\theta})). \tag{19}$$

3.3. Comparison between full delegation and no delegation

The condition under which the government prefers no delegation to full delegation at stage 1 is

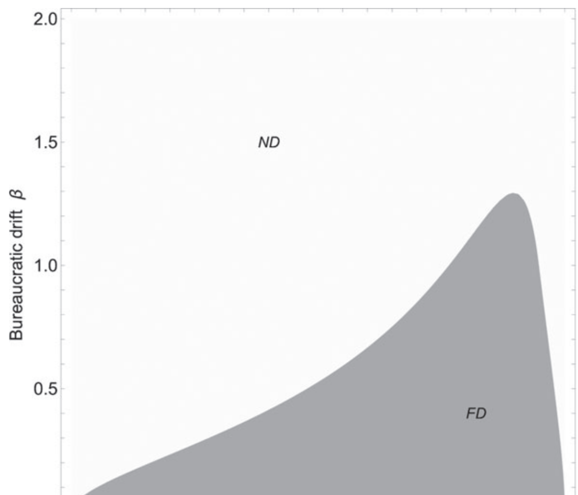
$$\triangle D := U_G^{ND} - U_G^{FD} > 0. \tag{20}$$

The following proposition characterizes the government's preference over no delegation and full delegation.

Proposition 1. *Consider the game in which the government chooses between full and no delegation. For a given ν, the government chooses full delegation in equilibrium if and only if $\beta < \bar{\beta}^{FD}$ for some $\bar{\beta}^{FD} > 0$. Moreover, for a given β, the government chooses full delegation in equilibrium if and only if $\nu \in [\underline{\nu}^{FD}, \bar{\nu}^{FD}]$, which can be a null set, for some $0 < \underline{\nu}^{FD} \leq \bar{\nu}^{FD} < 1$.*

Figure 2 plots government's preferences over full delegation and no delegation in (ν, β) space. The figure illustrates how the result in Proposition 1 is consistent with both the ally principle and the uncertainty principle. Note, first, that the effect of β is consistent with the ally principle, which suggests that the government prefers to give more discretion to better-aligned bureaucrats. This effect arises as the government's payoff under full delegation decreases with β whereas β has no impact on that payoff

Figure 2. No delegation versus full delegation in (ν, β) space



Notes: Specification: $S = 50$, $K = 10$, $\bar{\theta} = 4$, $\underline{\theta} = 2$, and $\lambda = 0.15$.

under no delegation. The intuition is that a bureaucrat with a low bias, as measured by β, is more easily trusted by the government, and therefore will be delegated the regulation of the firm more often than a bureaucrat with a high bias.

Secondly, the effect of ν is consistent with the uncertainty principle, which suggests that the government prefers more bureaucratic discretion in situations with high uncertainty. To see this, observe that the government's no-delegation contract coincides with the full information contract at $\nu = 0$ and $\nu = 1$, and gives her a higher payoff than she receives under full delegation. Further, her expected payoff from full delegation changes linearly with respect to ν, and in Lemma A1 in the Appendix, we show that U_G^{ND} in equation (15) is convex in ν. These observations together imply that $\triangle D$ is convex in ν and is positive as ν approaches 0 or 1. Therefore, $\triangle D$ in equation (20) can take negative values only at an intermediate range of ν (i.e., when the uncertainty about the firm's type is high). It follows that the government's benefit from the bureaucrat's informational advantage is high in situations with high uncertainty. The intuition is that the value of the bureaucrat's knowledge about the firm's technology is greatest when the government's uncertainty about it is the greatest. When ν is close to either 0 or 1, the government is almost certain of the firm's technology and the benefit of the bureaucrat's knowledge is smaller.

4. Constrained delegation

The government can improve its payoff from delegation by restricting the bureaucrat's choice set. As the bureaucrat has an interest in reducing pollution, her preferred pollution level is always below that of the consumer. The government can therefore improve its payoff by imposing a lower bound on the bureaucrat's choice of this level. However, being uninformed, it cannot impose type-dependent bounds. In order to study the government's interest in setting bounds on pollution levels, we will be considering a bureaucrat choosing regulatory contracts $\alpha(\theta) = (t(\theta), d(\theta)) \in A$, $\theta \in \{\underline{\theta}, \bar{\theta}\}$ under the constraint that $d(\theta) \in [d_1, d_2] \subseteq [0, K]$; it is this constraint that we call *constrained delegation*. To study it, we modify stage 1 of the game as follows.

Stage 1 (modified). The government decides whether or not to delegate the decision-making authority to an independent bureaucrat. If it delegates, then the government chooses $0 \leq d_1 \leq d_2 \leq K$ such that the bureaucrat can offer contract $\alpha(\theta) = (t(\theta), d(\theta)) \in A$ with the constraint that $d(\theta) \in [d_1, d_2] \subseteq [0, K]$, $\theta \in \{\underline{\theta}, \bar{\theta}\}$. If it does not delegate, then the authority remains with the government.

Below, we first look at how constrained delegation affects the bureaucrat's choice of regulation contracts. Her optimal contract for type θ solves the following problem:

$$\max_{\alpha \in A} \; U_B(\theta, \alpha),$$
$$\text{subject to } \; U_P(\theta, \alpha) \geq 0 \; \text{ and } \; d \in [d_1, d_2]. \tag{21}$$

We denote the solution with a superscript C and a subscript BI. The following lemma describes the bureaucrat's optimal choice of contracts under constrained delegation.

Lemma 3. *Assume that the government delegates the decision-making authority with the constraint that* $d \in [d_1, d_2] \subseteq [0, K]$. *The bureaucrat's preferred regulation contract for type* $\theta \in \{\underline{\theta}, \bar{\theta}\}$ *is given by* $\alpha^C_{BI}(\theta, d_1, d_2) = (t^C_{BI}(\theta, d_1, d_2), d^C_{BI}(\theta, d_1, d_2))$, *where*

$$d^C_{BI}(\theta, d_1, d_2) = \begin{cases} d_1, & \text{if } d_1 \geq d_{BI}(\theta); \\ d_{BI}(\theta), & \text{if } d_1 < d_{BI}(\theta) < d_2; \\ d_2, & \text{if } d_{BI}(\theta) \geq d_2. \end{cases} \tag{22}$$

$$t^C_{BI}(\theta, d_1, d_2) = S - \theta[K - d^C_{BI}(\theta, d_1, d_2)]. \tag{23}$$

The bureaucrat's choice of contract under constrained delegation coincides with her choice under full delegation if $d_{BI}(\theta)$ in equation (17) lies in the bounded interval $[d_1, d_2]$; otherwise, the optimal choice lies at one of the boundaries. The government can therefore affect her choice by manipulating d_1 and d_2. The government's choice of bounds d_1 and d_2 solves the following problem:

$$\max_{d_1, d_2} \nu U_G(\underline{\theta}, \alpha^C_{BI}(\underline{\theta}, d_1, d_2)) + (1 - \nu) U_G(\bar{\theta}, \alpha^C_{BI}(\bar{\theta}, d_1, d_2)). \tag{24}$$

The following lemma describes the choice for the upper bound.

Lemma 4. *Fix* $d_1 \in [0, K]$. *Suppose the government delegates with a constraint that* $d(\underline{\theta}), d(\bar{\theta}) \in [d_1, d_2]$, *for some* $d_2 \in [d_1, K]$. *The government's payoff is maximized at any* $d_2 \geq \max\{d_1, d_{BI}(\bar{\theta})\}$.

Disregarding the government's indifference, we simply put its choice at $d_2 = \max\{d_1, d_{GI}(\bar{\theta})\} \geq \max\{d_1, d_{BI}(\bar{\theta})\}$. The following lemma describes potential choices for the optimal lower bound.

Lemma 5. *Fix* $d_2 = d_{GI}(\bar{\theta})$. *Suppose the government delegates with a constraint that* $d(\underline{\theta}), d(\bar{\theta}) \in [d_1, d_2]$, *for some* $d_1 \in [0, d_2]$. *If* $d_{BI}(\bar{\theta}) \leq d_{GI}(\underline{\theta})$, *then, among all* $d_1 \in [0, d_2]$, *the government's payoff is maximized at* $d_1 = (1 + \lambda) E_\theta \theta = d_{GI}(E_\theta \theta)$. *If* $d_{BI}(\bar{\theta}) > d_{GI}(\underline{\theta})$, *then, among all* $d_1 \leq d_{BI}(\bar{\theta})$, *the government's payoff is maximized at* $d_1 = d_{GI}(\underline{\theta})$, *while among all* $d_1 \in (d_{BI}(\bar{\theta}), d_2]$, *its payoff is maximized at* $d_1 = (1 + \lambda) E_\theta \theta = d_{GI}(E_\theta \theta)$.

Lemma 5 implies that, if the government delegates with constraints, then two possibilities may arise. In the first case, the government chooses $d_1 = d_{GI}(\underline{\theta})$. In response, the bureaucrat sets $d^C_{BI}(\underline{\theta}, d_1, d_2) = d_{GI}(\underline{\theta})$ and $d^C_{BI}(\bar{\theta}, d_1, d_2) = d_{BI}(\bar{\theta})$. The government implements the full-information regulation contract if the firm is low-cost. There is distortion at the contract offered to a high-cost firm, as $d_{BI}(\bar{\theta}) < d_{GI}(\bar{\theta})$, while the full-information contract is implemented if the firm is low-cost. We refer to this case as weakly constrained delegation, or simply weak delegation (WD), because the government puts rather weak constraints on the bureaucrat's contract choices. The government's expected payoff under weak delegation is given by

$$U^{WD}_G = \nu U_G(\underline{\theta}, \alpha_{GI}(\underline{\theta})) + (1-\nu) U_G(\bar{\theta}, \alpha_{BI}(\bar{\theta})). \tag{25}$$

In the second case, the government chooses $d_1 = d_{GI}(E_\theta \theta) > d_{BI}(\bar{\theta})$. In response, the bureaucrat sets $d^C_{BI}(\underline{\theta}, d_1, d_2) = d^C_{BI}(\bar{\theta}, d_1, d_2) = d_{GI}(E_\theta \theta)$, resulting in a uniform pollution level for both types of firm. The government's choice of d_1 is the optimal uniform pollution level. We refer to this case as strictly constrained delegation, or simply *strict delegation* (SD), since the constraints are stricter than under weak delegation. The government's expected payoff under strict delegation is given by

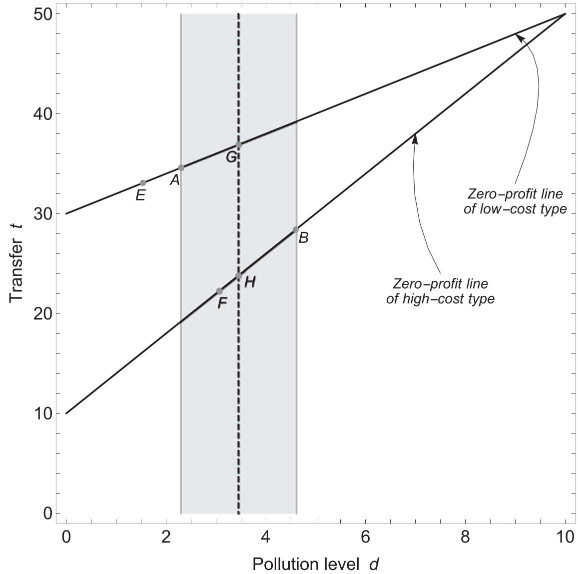
$$U^{SD}_G = \nu U_G(\underline{\theta}, \alpha_{GI}(E_\theta \theta)) + (1-\nu) U_G(\bar{\theta}, \alpha_{GI}(E_\theta \theta)) = U_G(E_\theta \theta, \alpha_{GI}(E_\theta \theta)). \tag{26}$$

In Figure 3, which is comparable to Figure 1, we illustrate the two possibilities that the government can induce through constrained delegation. In our example, $d_{GI}(\underline{\theta}) = 2.3$, $d_{GI}(\bar{\theta}) = 4.6$, and $d_{GI}(E_\theta \theta) = 3.45$. With constrained delegation, the government can either implement contracts A and F by setting d_1 at 2.3 (weak delegation), or implement contracts G and H by setting d_1 at 3.45 (strict delegation). The shaded area shows the bureaucrat's choice set under weak delegation, whereas the dashed vertical line pictures the choice set under strict delegation.

The benefits of weak delegation come from implementing the full-information contract if the firm is low-cost. Therefore, the government prefers weak delegation over strict delegation if the probability of the firm being low-cost is sufficiently high. The costs of weak delegation lie in the distortion at the contract offered to a high-cost firm. This distortion increases with β and so does the government's preference for strict delegation over weak delegation. The following lemma documents how the two key parameters of our model, β and ν, affect the government's preference over the weak and the strict forms of delegation.

Lemma 6. *There exists a threshold $\bar{\nu}^{SD}(\beta)$, which is increasing in β, such that $U^{SD}_G > U^{WD}_G$ if and only if $\nu < \bar{\nu}^{SD}(\beta)$.*

Figure 3. Constrained delegation in the (d, t) space



Notes: Specification: $S = 50$, $K = 10$, $\bar{\theta} = 4$, $\underline{\theta} = 2$, $\lambda = 0.15$, $\nu = 0.5$, and $\beta = 0.5$.

Lemma 6 implies that the government implements either strict delegation or no delegation for $\nu < \bar{\nu}^{SD}(\beta)$ and implements either weak delegation or no delegation for $\nu \in (\bar{\nu}^{SD}(\beta), 1)$.

First, consider the case $\nu < \bar{\nu}^{SD}(\beta)$. Because the government's payoff under strict delegation and no delegation is invariant to β, its choice between these regimes would only depend on ν. When there is no uncertainty about the firm's type, (i.e., when $\nu = 0$ or $\nu = 1$), the government implements the relevant full-information contract and receives the same payoff under both no delegation and strict delegation. In Lemmas A1 and A2 in the Appendix, we show that both U_G^{ND} and U_G^{SD} are increasing and convex in ν. Moreover, as both payoffs are independent of β, we find that, if $U_G^{SD} \geq U_G^{ND}$ for some β, then this relationship holds for all β.

Next, consider the case $\nu \geq \bar{\nu}^{SD}(\beta)$. In this case, the government chooses between weak delegation and no delegation. We find that the difference in the government's expected payoff between the two regimes of no delegation and weak delegation, $U_G^{ND} - U_G^{WD}$, is convex in ν and strictly positive as ν approaches 0, and approaches 0 as ν approaches 1. These observations together imply that $U_G^{ND} - U_G^{WD}$ is negative, if at all, only if ν is above a threshold. The following lemma documents this observation.

Lemma 7. *There exists a threshold $\bar{v}^{ND}(\beta)$, which is increasing in β, such that $U_G^{ND} > U_G^{WD}$ for $v < \bar{v}^{ND}(\beta)$.*

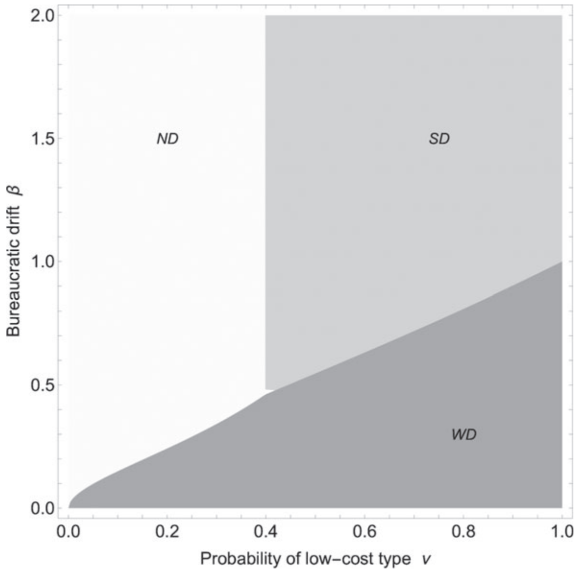
Lemmas 6 and 7 imply that we can observe three possible regimes in equilibrium: weak, strict, or no delegation. The following proposition characterizes how different regimes can arise in equilibrium. The proof follows directly from the above discussion.

Proposition 2. *The government chooses weak delegation in equilibrium if and only if $v \geq \max\{\bar{v}^{SD}(\beta), \bar{v}^{ND}(\beta)\}$, where $\bar{v}^{SD}(\beta)$ and $\bar{v}^{ND}(\beta)$ are defined in Lemmas 6 and 7, respectively. For $v < \max\{\bar{v}^{SD}(\beta), \bar{v}^{ND}(\beta)\}$, the government implements either strict or no delegation. Furthermore, if strict (no) delegation occurs in equilibrium for some β, then no (strict) delegation cannot occur for any β.*

Figure 4 plots the equilibrium delegation regimes in (v, β) space for the numerical example illustrated in Figure 2. In this example, $U_G^{SD} > U_G^{ND}$ if and only if $v \in (0.399, 1)$. Further, for any given β, weak delegation occurs only if v is sufficiently large.

Bureaucratic discretion reduces with bureaucratic drift β. For a given v, weak delegation is dominant for a sufficiently large β and the government

Figure 4. Optimal constrained delegation in (v, β) space



Notes: Specification: $S = 50$, $K = 10$, $\bar{\theta} = 4$, $\underline{\theta} = 2$, and $\lambda = 0.15$.

chooses to give the bureaucrat less discretion, through either strict or no delegation. This observation is in line with the ally principle.

Uncertainty, however, affects bureaucratic discretion non-trivially. The effect depends on how uncertainty changes – whether it happens because the firm is more likely to be low-cost or because it is more likely to be high-cost. It follows from Lemmas 6 and 7 that $\max\{\bar{\nu}^{SD}(\beta), \bar{\nu}^{ND}(\beta)\}$ is increasing in β. If the firm is more likely to be low-cost, so that $\nu \geq \max\{\bar{\nu}^{SD}(\beta), \bar{\nu}^{ND}(\beta)\}$, then the government chooses weak delegation: weak delegation, without an information rent, implements the full-information contract for the low-cost firm, and this is the likely firm type when ν is high. If the firm, instead, is more likely to be high-cost, so that $\nu < \max\{\bar{\nu}^{SD}(\beta), \bar{\nu}^{ND}(\beta)\}$, then the government opts for giving the bureaucrat less discretion, through either strict or no delegation. The uncertainty principle, therefore, does not carry over to a setting where constrained delegation is possible: in our setting, weak delegation is more prevalent the more likely it is that the firm is low-cost. The reason is that strict delegation, which is an alternative to weak delegation, would be based on the government's *ex ante* assessment of the firm's technology, as illustrated in Figure 3. When ν is low, the distortion created in the low-cost type's pollution level is large, with the low-cost strict-delegation contract at G in Figure 3 being far from the low-cost weak-delegation contract at A; on top of that, this distortion is *ex ante* likely, exactly because ν is low. The prevalence of weak delegation at high values of ν is in contrast to the case when only full delegation is available, in which the outcome is consistent with the uncertainty principle, as shown in Section 3.

For $\nu < \max\{\bar{\nu}^{SD}(\beta), \bar{\nu}^{ND}(\beta)\}$, Proposition 2 does not provide a full characterization, except that there will be either no delegation or strict delegation – which of the two dominates does not depend on β. Thus, as illustrated in Figure 4, there will be vertical lines splitting the regions of no delegation and strict delegation. The difficulty in assessing this case stems from the difference $U_G^{ND} - U_G^{SD}$ being non-monotonic when $\nu < \max\{\bar{\nu}^{SD}(\beta), \bar{\nu}^{ND}(\beta)\}$. In the Appendix, we provide a further discussion of the comparison between no delegation and strict delegation, including an example where the government, for $\nu < \max\{\bar{\nu}^{SD}(\beta), \bar{\nu}^{ND}(\beta)\}$, chooses no delegation for low and high ν and strict delegation for ν in a middle range. See, in particular, Figure A1 in the Appendix, illustrating an instance of non-monotonicity of $U_G^{ND} - U_G^{SD}$.

Kundu and Nilssen (2020) discuss delegation in a procurement setting, and do not find any occurrence of no delegation at large values of ν. In fact, they show that, if no delegation at large values of ν does not occur for some specific value λ' of λ, then it cannot occur for any $\lambda > \lambda'$

either.[8] The example we present in Figure A2 in the Appendix is a direct counterexample to that result for the present analysis of delegation in a permits setting: the numerical examples pictured in Figures 4 and A2 are identical except that λ is higher in Figure A2, where no delegation occurs at large values of ν, than in Figure 4, where it does not. This observation indicates that the prevalence of no delegation being chosen for both low and high values of ν is larger in the present permits setting than in the procurement setting of Kundu and Nilssen (2020).

5. Conclusion

In this paper, we have discussed how constrained delegation can improve the sales of pollution permits in situations where the polluting firm has private information about its technology. As documented by Gilardi (2009), there has been an increase in the independence of regulatory agencies over the last few decades. We would like to argue that it is important that this delegation is well understood. Our analysis shows that a carefully designed delegation, putting limits on what the hired biased experts can do, is efficient. In particular, we show how strict delegation, even though giving away very little regulatory decision power to the expert, is better than no delegation at all.

Appendix

We begin with a useful concept that is applied in some of these proofs.

Definition A1. *Define*

$$f(\theta, d) := U_G(\theta, \alpha),$$

where $\alpha = (S - \theta(K - d), d)$*, as the government's full-information payoff when obtaining pollution level* d *from a firm of type* θ*.*

It follows that

$$f(\theta, d) = (1 + \lambda)[S - \theta(K - d)] - \frac{1}{2}d^2. \tag{A1}$$

Observe that $\mathrm{d}f/\mathrm{d}d = (1 + \lambda)\theta - d$ and $\mathrm{d}^2 f/\mathrm{d}d^2 < 0$. It follows that

$$\frac{\mathrm{d}f(\theta, d)}{\mathrm{d}d} \gtreqqless 0 \quad \text{if} \quad d \lesseqqgtr d_{GI}(\theta). \tag{A2}$$

[8] See Lemma B.2 in the online appendix of Kundu and Nilssen (2020), as well as the discussion of that lemma on page 461.

Applying the envelope theorem, we obtain

$$\frac{\mathrm{d}f(\theta, d_{GI}(\theta))}{\mathrm{d}\theta} = -(1+\lambda)[K - d_{GI}(\theta)] < 0, \tag{A3}$$

implying that

$$f(\underline{\theta}, d_{GI}(\underline{\theta})) > f(\bar{\theta}, d_{GI}(\bar{\theta})). \tag{A4}$$

Furthermore,

$$f(\theta, d_1) - f(\theta, d_2) = (d_1 - d_2)\left[(1+\lambda)\theta - \frac{d_1 + d_2}{2}\right], \tag{A5}$$

and

$$f(\theta_1, d) - f(\theta_2, d) = (1+\lambda)(\theta_2 - \theta_1)(K - d). \tag{A6}$$

Proof of Lemma 1: We use the method of substitution to solve for the optimal contracts. This is the standard method used in the screening literature; see, for example, Bolton and Dewatripont (2005, Chapter 2). The government's expected payoff is increasing in transfers, and the firm's payoff is decreasing in transfers. The government would therefore prefer to increase transfers as much as possible, subject to the firm's participation constraint, which can be achieved by having equations (ICL) and (IRH) binding without affecting the other constraints. The low-cost firm can always pretend to be the high-cost firm and receive a payoff of $\bar{t} - \underline{\theta}(K - \bar{d})$. In order to make it choose $(\underline{t}, \underline{d})$, the government therefore shares an information rent of $IR(\bar{d}) := \triangle\theta(K - \bar{d})$. Thus, $\bar{t} = S - \bar{\theta}(K - \bar{d})$ and $\underline{t} = S - \underline{\theta}(K - \underline{d}) - IR(\bar{d})$. Replacing $\bar{t}$ and $\underline{t}$ in equation (9) and using the fact that equations (IRH) and (ICL) together imply equation (IRL), we can rewrite the optimization problem as

$$\begin{aligned} \max_{\underline{d}, \bar{d}} \ & \nu\left[(1+\lambda)S - (1+\lambda)\underline{\theta}(K - \underline{d}) - \frac{1}{2}\underline{d}^2\right] \\ & + (1-\nu)\left[(1+\lambda)S - (1+\lambda)\bar{\theta}(K - \bar{d}) - \frac{1}{2}\bar{d}^2\right] - \nu\lambda IR(\bar{d}), \end{aligned}$$

subject to equation (ICH). (A7)

From the first-order conditions of the unconstrained problem, we see that the pollution levels are given by

$$\begin{aligned} & d_{GN}(\underline{\theta}) = (1+\lambda)\underline{\theta}, \\ & d_{GN}(\bar{\theta}) \min\left\{(1+\lambda)\bar{\theta} + \frac{\nu\lambda\triangle\theta}{1-\nu}, K\right\}. \end{aligned}$$

Note that $d_{GN}(\bar{\theta}) \geq d_{GN}(\underline{\theta})$, which ensures that equation (ICH) is satisfied at the unconstrained solution. The transfers given in the lemma then follow. □

The following lemma documents the convexity property of U_G^{ND}, which is used in the proof of Proposition 1.

Lemma A1. *The government's expected payoff in the no-delegation regime, U_G^{ND}, is increasing and convex in $\nu \in (0, 1)$.*

Proof of Lemma A1: By equation (A7), the government's expected payoff in the no-delegation regime is

$$\begin{aligned} U_G^{ND} &= \max_{\underline{d},\bar{d}} \; \nu f(\underline{\theta}, \underline{d}) + (1-\nu) f(\bar{\theta}, \bar{d}) - \nu\lambda IR(\bar{d}) \\ &= \nu f(\underline{\theta}, d_{GN}(\underline{\theta})) + (1-\nu) f(\bar{\theta}, d_{GN}(\bar{\theta})) - \nu\lambda\triangle\theta[K - d_{GN}(\bar{\theta})], \end{aligned} \tag{A8}$$

where $d_{GN}(\underline{\theta})$ and $d_{GN}(\bar{\theta})$ are given in equations (10) and (11), respectively. Applying the envelope theorem, we find that

$$\begin{aligned} \frac{\mathrm{d}U_G^{ND}}{\mathrm{d}\nu} &= f(\underline{\theta}, d_{GN}(\underline{\theta})) - f(\bar{\theta}, d_{GN}(\bar{\theta})) - \lambda\triangle\theta[K - d_{GN}(\bar{\theta})] \qquad \text{(A9)} \\ &= \left[f(\underline{\theta}, d_{GN}(\underline{\theta})) - f(\underline{\theta}, d_{GN}(\bar{\theta})) \right] \\ &\quad + \left[f(\underline{\theta}, d_{GN}(\bar{\theta})) - f(\bar{\theta}, d_{GN}(\bar{\theta})) \right] - \lambda\triangle\theta[K - d_{GN}(\bar{\theta})] \\ &= \left\{ [d_{GN}(\underline{\theta}) - d_{GN}(\bar{\theta})] \left[(1+\lambda)\underline{\theta} - \frac{d_{GN}(\underline{\theta}) + d_{GN}(\bar{\theta})}{2} \right] \right\} \\ &\quad + \left\{ (1+\lambda)\triangle\theta[K - d_{GN}(\bar{\theta})] \right\} - \lambda\triangle\theta[K - d_{GN}(\bar{\theta})], \end{aligned}$$

by equations (A5) and (A6)

$$= \frac{[d_{GN}(\underline{\theta}) - d_{GN}(\bar{\theta})]^2}{2} + \triangle\theta[K - d_{GN}(\bar{\theta})], \tag{A10}$$

which is strictly positive for all $\nu \in (0, 1)$ and therefore U_G^{ND} is strictly increasing.

To see convexity, first consider $\nu \in (0, \nu^*)$. Then, $d_{GN}(\bar{\theta}) = (1+\lambda)\bar{\theta} + (\nu\lambda\triangle\theta)/(1-\nu)$. Therefore,

$$\begin{aligned} \frac{\mathrm{d}d_{GN}(\bar{\theta})}{\mathrm{d}\nu} &= \frac{\lambda\triangle\theta}{(1-\nu)^2}, \\ \frac{\partial f(\bar{\theta}, d_{GN}(\bar{\theta}))}{\partial d} &= (1+\lambda)\bar{\theta} - d_{GN}(\bar{\theta}) = -\frac{\nu\lambda\triangle\theta}{1-\nu}. \end{aligned}$$

Differentiation of equation (A9) with respect to ν gives

$$\begin{aligned} \frac{\mathrm{d}^2 U_G^{ND}}{\mathrm{d}\nu^2} &= -\frac{\partial f(\bar{\theta}, d_{GN}(\bar{\theta}))}{\partial d} \frac{\mathrm{d}d_{GN}(\bar{\theta})}{\mathrm{d}\nu} + \lambda\triangle\theta \frac{\mathrm{d}d_{GN}(\bar{\theta})}{\mathrm{d}\nu} \\ &= \frac{\nu\lambda^2\triangle^2\theta}{(1-\nu)^3} + \frac{\lambda^2\triangle^2\theta}{(1-\nu)^2} = \frac{\lambda^2\triangle^2\theta}{(1-\nu)^3}, \end{aligned}$$

which is strictly positive for all $\nu < \nu^*$. Therefore, $\mathrm{d}U_G^{ND}/\mathrm{d}\nu$ is strictly increasing in $\nu \in (0, \nu^*)$.

Next, consider $\nu \in (\nu^*, 1)$. Then, $d_{GN}(\bar{\theta}) = K$. From equation (A9),

$$\frac{\mathrm{d}U_G^{ND}}{\mathrm{d}\nu} = f(\underline{\theta}, d_{GN}(\underline{\theta})) - f(\bar{\theta}, K), \tag{A11}$$

which is independent of ν for $\nu \in (\nu^*, 1)$. Further, note that

$$\lim_{\nu \nearrow \nu^*} \frac{\mathrm{d}U_G^{ND}}{\mathrm{d}\nu} = \lim_{\nu \searrow \nu^*} \frac{\mathrm{d}U_G^{ND}}{\mathrm{d}\nu} = f(\underline{\theta}, d_{GN}(\underline{\theta})) - f(\bar{\theta}, K),$$

which implies that $\mathrm{d}U_G^{ND}/\mathrm{d}\nu$ is continuous at $\nu = \nu^*$. Therefore, $\mathrm{d}U_G^{ND}/\mathrm{d}\nu$ is strictly positive and increasing in ν for $\nu < \nu^*$ and continuous at $\nu = \nu^*$, and it remains constant at a strictly positive value at $\nu > \nu^*$. From this, we find that $\mathrm{d}U_G^{ND}/\mathrm{d}\nu$ is strictly positive and weakly increasing in ν for $\nu \in (0, 1)$, which proves convexity of U_G^{ND}. □

Proof of Proposition 1: We prove the first part of the proposition by showing that $\Delta D < 0$ for $\beta = 0$ and $\mathrm{d}\Delta D/\mathrm{d}\beta > 0$ for $\beta > 0$. For $\beta = 0$, $U_G^{FD} = U_G^{FI} > U_G^{ND}$, and therefore, $\Delta D < 0$. To see that $\mathrm{d}\Delta D/\mathrm{d}\beta > 0$, first note that U_G^{ND} is independent of β. By equation (18), the government's expected payoff in the full-delegation regime is

$$\begin{aligned} U_G^{FD} &= \nu \left[(1+\lambda)\{S - \underline{\theta}[K - d_{BI}(\underline{\theta})]\} - \frac{d_{BI}^2(\underline{\theta})}{2} \right] \\ &\quad + (1-\nu) \left[(1+\lambda)\{S - \bar{\theta}[K - d_{BI}(\bar{\theta})]\} - \frac{d_{BI}^2(\bar{\theta})}{2} \right] \\ &= \nu f(\underline{\theta}, d_{BI}(\underline{\theta})) + (1-\nu) f(\bar{\theta}, d_{BI}(\bar{\theta})), \end{aligned} \tag{A12}$$

where $f(\theta, d)$ is given in equation (A1) and $d_{BI}(\theta)$ is given in equation (17). Therefore,

$$\frac{\mathrm{d}U_G^{FD}}{\mathrm{d}\beta} = \nu \left[\frac{\mathrm{d}f(\underline{\theta}, d_{BI}(\underline{\theta}))}{\mathrm{d}d} . \frac{\mathrm{d}d_{BI}(\underline{\theta})}{\mathrm{d}\beta} \right] + (1-\nu) \left[\frac{\mathrm{d}f(\bar{\theta}, d_{BI}(\bar{\theta}))}{\mathrm{d}d} \frac{\mathrm{d}d_{BI}(\bar{\theta})}{\mathrm{d}\beta} \right].$$

We have $\mathrm{d}f(\theta, d_{BI}(\theta))/\mathrm{d}d > 0$ for $\theta \in \{\underline{\theta}, \bar{\theta}\}$, by equation (A2) and the fact that $d_{BI}(\theta) < d_{GI}(\theta)$ for $\beta > 0$. Further, $\mathrm{d}d_{BI}(\theta)/\mathrm{d}\beta = -(1+\lambda)\theta/(1+\beta)^2 < 0$. Together, the two inequalities imply $\mathrm{d}U_G^{FD}/\mathrm{d}\beta < 0$. Therefore,

$$\frac{\mathrm{d}\Delta D}{\mathrm{d}\beta} = \frac{\mathrm{d}U_G^{ND}}{\mathrm{d}\beta} - \frac{\mathrm{d}U_G^{FD}}{\mathrm{d}\beta} > 0.$$

We prove the second part of the proposition by showing that $\Delta D > 0$ for $\nu = 0, 1$, and $\triangle D$ is weakly convex in ν for $\nu \in [0, 1]$. These properties of $\triangle D$ ensure that $\triangle D$ can be negative only at an interval, if at all.

For $\nu = 0, 1$, $U_G^{ND} = U_G^{FI} > U_G^{FD}$, and therefore, $\Delta D > 0$. Further, from equation (A12), U_G^{FD} is linear in ν, and by Lemma A1, U_G^{ND} is weakly convex in ν. Therefore, $\triangle D$ is weakly convex in ν for $\nu \in [0, 1]$.

Proof of Lemma 3: The proof follows from replacing t by $S - \theta(K - d)$ in $U_B(\theta, \alpha)$ and using the first-order condition of equation (21).

Proof of Lemma 4: By Lemma 3, if $d_2 \leq d_{BI}(\bar{\theta})$, then the bureaucrat sets $d_{BI}^C(\bar{\theta}, d_1, d_2) = d_2$, and the government's payoff increases with d_2 in this range. If $d_2 \geq d_{BI}(\bar{\theta})$, then the bureaucrat sets $d_{BI}^C(\bar{\theta}, d_1, d_2) = d_{BI}(\bar{\theta})$, the government's payoff is independent of d_2 in this range, and the payoff is higher than what it receives by setting $d_2 \leq d_{BI}(\bar{\theta})$. Hence, the government's payoff is maximized at any $d_2 \geq d_{BI}(\bar{\theta})$.

Proof of Lemma 5: Consider first the case of $d_{BI}(\bar{\theta}) \leq d_{GI}(\underline{\theta})$. For $d_1 < d_{BI}(\underline{\theta})$, the bureaucrat sets $d_{BI}^C(\underline{\theta}, d_1, d_2) = d_{BI}(\underline{\theta})$ and $d_{BI}^C(\bar{\theta}, d_1, d_2) = d_{BI}(\bar{\theta})$, and the government's payoff is independent of d_1. For $d_{BI}(\underline{\theta}) \leq d_1 \leq d_{BI}(\bar{\theta})$, the bureaucrat sets $d_{BI}^C(\underline{\theta}, d_1, d_2) = d_1$ and $d_{BI}^C(\bar{\theta}, d_1, d_2) = d_{BI}(\bar{\theta})$, and the government's payoff is increasing in d_1, because $d_{BI}(\underline{\theta}) \leq d_{GI}(\underline{\theta})$. For $d_{BI}(\bar{\theta}) \leq d_1 \leq d_2$, the bureaucrat sets $d_{BI}^C(\underline{\theta}, d_1, d_2) = d_{BI}^C(\bar{\theta}, d_1, d_2) = d_1$, resulting in a uniform pollution level for both types of firms. In such a case, the government's payoff increases with d_1 for $d_1 \leq (1 + \lambda)E_\theta\theta$ and decreases thereafter. Note that $(1 + \lambda)E_\theta\theta = d_{GI}(E_\theta\theta) < K$.

Consider next the case of $d_{BI}(\bar{\theta}) > d_{GI}(\underline{\theta})$. For $d_1 < d_{BI}(\underline{\theta})$, the bureaucrat sets $d_{BI}^C(\underline{\theta}, d_1, d_2) = d_{BI}(\underline{\theta})$ and $d_{BI}^C(\bar{\theta}, d_1, d_2) = d_{BI}(\bar{\theta})$, and the government's payoff is independent of d_1. For $d_{BI}(\underline{\theta}) \leq d_1 \leq d_{BI}(\bar{\theta})$, the bureaucrat sets $d_{BI}^C(\underline{\theta}, d_1, d_2) = d_1$ and $d_{BI}^C(\bar{\theta}, d_1, d_2) = d_{BI}(\bar{\theta})$, and the government's payoff increases with d_1 for $d_1 \leq d_{GI}(\underline{\theta})$ and decreases thereafter. For $d_{BI}(\bar{\theta}) \leq d_1 \leq d_2$, the bureaucrat sets $d_{BI}^C(\underline{\theta}, d_1, d_2) = d_{BI}^C(\bar{\theta}, d_1, d_2) = d_1$, resulting in a uniform pollution level for both types of firms. In such a case, the government's payoff increases with d_1 for $d_1 \leq (1 + \lambda)E_\theta\theta$ and decreases thereafter.

The following lemma documents the convexity property of U_G^{SD}, which is used in the proof of Lemma 6.

Lemma A2. *The government's expected payoff in the strict-delegation regime, U_G^{SD}, is increasing and convex in $\nu \in (0, 1)$.*

Proof of Lemma A2: Note that

$$\begin{aligned} U_G^{SD} &= \max_d \, \nu \left\{(1+\lambda)[S - \underline{\theta}(K-d)] - \frac{d^2}{2}\right\} \\ &\quad + (1-\nu)\left\{(1+\lambda)[S - \bar{\theta}(K-d)] - \frac{d^2}{2}\right\} \\ &= \max_d \, \nu f(\underline{\theta}, d) + (1-\nu) f(\bar{\theta}, d) \\ &= \nu f(\underline{\theta}, d_{GI}(E_\theta \theta)) + (1-\nu) f(\bar{\theta}, d_{GI}(E_\theta \theta)). \end{aligned} \tag{A13}$$

Applying the envelope theorem, we find that

$$\begin{aligned} \frac{\mathrm{d}U_G^{SD}}{\mathrm{d}\nu} &= f(\underline{\theta}, d_{GI}(E_\theta \theta)) - f(\bar{\theta}, d_{GI}(E_\theta \theta)) \\ &= (1+\lambda)\triangle\theta[K - d_{GI}(E_\theta\theta)], \text{ by equation (A6)}, \end{aligned} \tag{A14}$$

which is strictly positive. Further, $\mathrm{d}d_{GI}(E_\theta\theta)/\mathrm{d}\nu = -\triangle\theta$, which yields $\mathrm{d}^2 U_G^{SD}/\mathrm{d}\nu^2 = (1+\lambda)\triangle^2\theta > 0$.

Proof of Lemma 6: We prove the lemma in the following steps.

Step 1: U_G^{WD} is increasing and linear in ν and is decreasing in β.

Proof of Step 1: Observe that

$$\begin{aligned} U_G^{WD} &= \nu\left[(1+\lambda)\{S - \underline{\theta}[K - d_{GI}(\underline{\theta})]\} - \frac{d_{GI}^2(\underline{\theta})}{2}\right] \\ &\quad + (1-\nu)\left[(1+\lambda)\{S - \bar{\theta}[K - d_{BI}(\bar{\theta})]\} - \frac{d_{BI}^2(\bar{\theta})}{2}\right] \\ &= \nu f(\underline{\theta}, d_{GI}(\underline{\theta})) + (1-\nu) f(\bar{\theta}, d_{BI}(\bar{\theta})). \end{aligned} \tag{A15}$$

Since $d_{GI}(\underline{\theta})$ and $d_{BI}(\bar{\theta})$ are independent of ν, U_G^{WD} is linear in ν. Further,

$$f(\underline{\theta}, d_{GI}(\underline{\theta})) > f(\bar{\theta}, d_{GI}(\bar{\theta})) > f(\bar{\theta}, d_{BI}(\bar{\theta})),$$

where the first inequality follows from (A4) and the second inequality from (A2). From (17), $d_{BI}(\bar{\theta})$ is decreasing in β. This observation and (A2) together imply that $f(\bar{\theta}, d_{BI}(\bar{\theta}))$ and subsequently U_G^{WD} are decreasing in β.

Step 2: U_G^{SD} is increasing and convex in ν and is independent of β.

Proof of Step 2: The effect of ν follows from Lemma A2. Since $d_{GI}(E_\theta\theta)$ is independent of β, it follows from (A13) that U_G^{SD} is also independent of β.

Step 3: For any $\beta > 0$,

$$\lim_{\nu \to 0}(U_G^{SD} - U_G^{WD}) > 0,$$
$$\lim_{\nu \to 1}(U_G^{SD} - U_G^{WD}) = 0.$$

Proof of Step 3: For $\beta > 0$,

$$\lim_{\nu \to 0} U_G^{SD} = f(\bar{\theta}, d_{GI}(\bar{\theta})) > f(\bar{\theta}, d_{BI}(\bar{\theta})) = \lim_{\nu \to 0} U_G^{WD},$$

where the inequality follows from (A2). Further,

$$\lim_{\nu \to 1} U_G^{SD} = \lim_{\nu \to 1} U_G^{WD} = f(\underline{\theta}, d_{GI}(\underline{\theta})).$$

From Steps 1, 2, and 3, it follows that two possibilities can arise: first, $U_G^{SD} > U_G^{WD}$ for $\nu \in (0,1)$; second, U_G^{SD} intersects U_G^{WD} at most once, at some $\bar{\nu}^{SD} \in [0,1]$. It follows that $U_G^{SD} > U_G^{WD}$ if and only if $\nu < \bar{\nu}^{SD}$. The first possibility arises when $\bar{\nu}^{SD} = 1$. Further, U_G^{SD} intersects U_G^{WD} from above, and therefore $\mathrm{d}U_G^{WD}/\mathrm{d}\nu > \mathrm{d}U_G^{SD}/\mathrm{d}\nu$ at $\nu = \bar{\nu}^{SD}$. To see how $\bar{\nu}^{SD}$ changes with β, write $U_G^{SD} = U_G^{SD}(\nu)$ and $U_G^{WD} = U_G^{WD}(\beta, \nu)$, and differentiate $U_G^{SD}(\bar{\nu}^{SD}) = U_G^{WD}(\beta, \bar{\nu}^{SD})$ with respect to β:

$$\frac{\partial U_G^{SD}(\bar{\nu}^{SD})}{\partial \nu}\frac{\mathrm{d}\bar{\nu}^{SD}}{\mathrm{d}\beta} = \frac{\partial U_G^{WD}(\beta, \bar{\nu}^{SD})}{\partial \nu}\frac{\mathrm{d}\bar{\nu}^{SD}}{\mathrm{d}\beta} + \frac{\partial U_G^{WD}(\beta, \bar{\nu}^{SD})}{\partial \beta}$$
$$\Leftrightarrow \frac{\mathrm{d}\bar{\nu}^{SD}}{\mathrm{d}\beta} = \frac{\partial U_G^{WD}(\beta, \bar{\nu}^{SD})/\partial \beta}{[\partial U_G^{SD}(\bar{\nu}^{SD})/\partial \nu] - [\partial U_G^{WD}(\beta, \bar{\nu}^{SD})/\partial \nu]},$$

which is positive as both the numerator and the denominator are negative. Therefore, $\bar{\nu}^{SD}(\beta)$ is increasing in β.

Proof of Lemma 7: Because U_G^{SD} and U_G^{ND} are both convex and coincide at $\nu = 0$ and $\nu = 1$, the existence and monotonicity properties of $\bar{\nu}^{ND}(\beta)$ follow from arguments similar to the ones we used in characterizing $\bar{\nu}^{SD}(\beta)$ in the proof of Lemma 6.

Comparison between no delegation and strict delegation

Recall from Lemmas A1 and A2 that both U_G^{ND} and U_G^{SD} are convex in ν. Moreover, they coincide at $\nu = 0$ and $\nu = 1$. In general, the difference $U_G^{ND} - U_G^{SD}$ is not monotonic in ν, and U_G^{ND} can intersect U_G^{SD} more than once. To explore their relationship, we study the derivatives U_G^{ND} and U_G^{SD} at the boundaries. From equation (A14), we have

$$\lim_{\nu \to 0} \frac{\mathrm{d}U_G^{SD}}{\mathrm{d}\nu} = (1+\lambda)\triangle\theta[K - d_{GI}(\bar{\theta})],$$

and

$$\lim_{\nu \to 1} \frac{\mathrm{d}U_G^{SD}}{\mathrm{d}\nu} = (1+\lambda)\triangle\theta[K - d_{GI}(\underline{\theta})].$$

Further, from equation (A10), we have

$$\lim_{\nu \to 0} \frac{\mathrm{d}U_G^{ND}}{\mathrm{d}\nu} = \frac{[d_{GI}(\underline{\theta}) - d_{GI}(\bar{\theta})]^2}{2} + \triangle\theta[K - d_{GI}(\bar{\theta})],$$

and

$$\lim_{\nu \to 1} \frac{\mathrm{d}U_G^{ND}}{\mathrm{d}\nu} = \frac{[d_{GI}(\underline{\theta}) - K]^2}{2},$$

which follow from the observations that

$$\begin{aligned} d_{GN}(\underline{\theta}) &= d_{GI}(\underline{\theta}), \\ \lim_{\nu \to 0} d_{GN}(\bar{\theta}) &= d_{GI}(\bar{\theta}), \\ \lim_{\nu \to 1} d_{GN}(\bar{\theta}) &= K. \end{aligned}$$

Therefore,

$$\begin{aligned} \lim_{\nu \to 0} \left(\frac{\mathrm{d}U_G^{ND}}{\mathrm{d}\nu} - \frac{\mathrm{d}U_G^{SD}}{\mathrm{d}\nu} \right) &= \frac{[d_{GI}(\underline{\theta}) - d_{GI}(\bar{\theta})]^2}{2} + \triangle\theta[K - d_{GI}(\bar{\theta})] \\ &\quad - (1+\lambda)\triangle\theta[K - d_{GI}(\bar{\theta})] \\ &= \frac{[d_{GI}(\underline{\theta}) - d_{GI}(\bar{\theta})]^2}{2} - \lambda\triangle\theta[K - d_{GI}(\bar{\theta})], \quad \text{(A16)} \end{aligned}$$

which is positive if

$$\begin{aligned} K - d_{GI}(\bar{\theta}) < \frac{[d_{GI}(\underline{\theta}) - d_{GI}(\bar{\theta})]^2}{2\lambda\triangle\theta} &= \frac{(1+\lambda)^2\triangle\theta}{2\lambda} \\ \Leftrightarrow K < \frac{(1+\lambda)^2\triangle\theta}{2\lambda} + (1+\lambda)\bar{\theta}. & \quad \text{(A17)} \end{aligned}$$

Condition (A17) holds if K is small and $\triangle\theta$ is large; the effect of λ is, however, non-monotonic. As both U_G^{ND} and U_G^{SD} are increasing in ν and have the same value as ν approaches 0, we find that, if equation (A17) holds, then no delegation is preferred to strict delegation as ν approaches 0.

However,

$$\lim_{\nu \to 1}\left(\frac{\mathrm{d}U_G^{ND}}{\mathrm{d}\nu} - \frac{\mathrm{d}U_G^{SD}}{\mathrm{d}\nu}\right) = \frac{[d_{GI}(\underline{\theta}) - K]^2}{2} - (1+\lambda)\triangle\theta[K - d_{GI}(\underline{\theta})]$$
$$= \frac{[K - d_{GI}(\bar{\theta})][K - d_{GI}(\bar{\theta}) - 2(1+\lambda)\triangle\theta]}{2},$$

which is positive if

$$K - d_{GI}(\bar{\theta}) > 2(1+\lambda)\triangle\theta$$
$$\Leftrightarrow K > 2(1+\lambda)(2\triangle\theta + \bar{\theta}). \tag{A18}$$

Condition (A18) holds if K is large, $\triangle\theta$ is small, and λ is small. As both U_G^{ND} and U_G^{SD} are increasing in ν and have the same value as ν approaches 1, we find that, if equation (A18) holds, then strict delegation is preferred to no delegation as ν approaches 1. In Figure 4, we consider the following parameter specification: $S = 50$, $K = 10$, $\bar{\theta} = 4$, $\underline{\theta} = 2$, and $\lambda = 0.15$, which give $d_{GI}(\underline{\theta}) = 2.3$, $d_{GI}(\bar{\theta}) = 4.6$, $K - d_{GI}(\bar{\theta}) = 5.4$, $[(1+\lambda)^2\triangle\theta]/2\lambda = 8.82$, and $2(1+\lambda)\triangle\theta = 4.6$. Therefore, both equations (A17) and (A18) hold true, implying that no delegation is preferred to strict delegation for ν close to 0, and strict delegation is preferred to no delegation for ν close to 1.

From equations (A17) and (A18), it follows that no delegation is preferred to strict delegation for both ν close to 0 and ν close to 1 if

Figure A1. Payoff differences

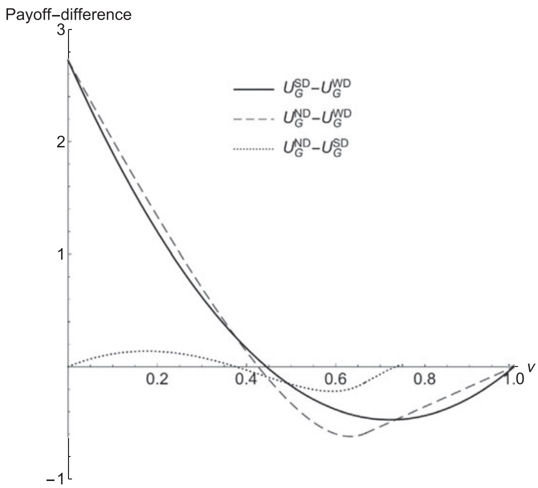
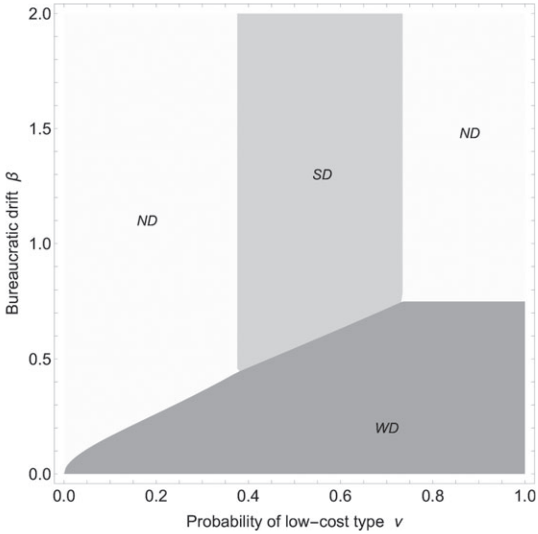


Figure A2. Optimal constrained delegation



$$K - d_{GI}(\bar{\theta}) < \min\left\{\frac{(1+\lambda)^2\triangle\theta}{2\lambda}, 2(1+\lambda)\triangle\theta\right\}. \quad \text{(A19)}$$

Let us continue our numerical example in the text, but now consider $\lambda = 0.75$, which also satisfies Assumption 1, and which gives $d_{GI}(\underline{\theta}) = 3.5$, $d_{GI}(\bar{\theta}) = 7$, $K - d_{GI}(\bar{\theta}) = 3$, $[(1+\lambda)^2\triangle\theta]/2\lambda = 4.08$, and $2(1+\lambda)\triangle\theta = 7$, thus satisfying equation (A19). In Figure A1, the dotted curve plots $U_G^{ND} - U_G^{SD}$ against the probability ν of the firm being low-cost, and it shows that only for an intermediate range of ν is strict delegation preferred to no delegation. Figure A2 illustrates the equilibrium regimes. Observe that, when no delegation is preferred to strict delegation for large values of ν, the equilibrium regime will be a choice between no delegation and weak delegation. For $\nu > \nu^*$, both U_G^{ND} and U_G^{WD} are linear in ν, and they coincide as ν approaches 1. Consequently, $U_G^{ND} - U_G^{WD}$ is linear in ν for $\nu > \nu^*$ and approaches 0 as ν approaches 1, which can be observed from the dashed curve that plots $U_G^{ND} - U_G^{WD}$ against ν in Figure A1.[9] The choice between weak delegation and no delegation is completely determined by the slope of the linear part of $U_G^{ND} - U_G^{WD}$, which is decreasing in β. Therefore, weak (no) delegation is preferred for small (large) β.

[9]Note that, in Figure A1, $U_G^{ND} - U_G^{WD}$ and $U_G^{SD} - U_G^{WD}$ depend on β and have been drawn for the case of $\beta = 0.5$.

References

Alonso, R. and Matouschek, N. (2008), Optimal delegation, *Review of Economic Studies 75*, 259–293.

Amador, M. and Bagwell, K. (2013), The theory of optimal delegation with an application to tariff caps, *Econometrica 81*, 1541–1599.

Ballard, C. L. and Fullerton, D. (1992), Distortionary taxes and the provision of public goods, *Journal of Economic Perspectives 6*, 117–131.

Bolton, P. and Dewatripont, M. (2005), *Contract Theory*, MIT Press, Cambridge, MA.

Boyer, M. and Laffont, J.-J. (1999), Toward a political theory of the emergence of environmental incentive regulation, *RAND Journal of Economics 30*, 137–157.

Caillaud, B., Guesnerie, R., Rey, P., and Tirole, J. (1988), Government intervention in production and incentives theory: a review of recent contributions, *RAND Journal of Economics 19*, 1–26.

Epstein, D. and O'Halloran, S. (1999), *Delegating Powers: A Transaction Cost Politics Approach to Policy Making under Separate Powers*, Cambridge University Press, Cambridge.

Frankel, A. (2016), Delegating multiple decisions, *American Economic Journal: Microeconomics 8*, 16–53.

Gailmard, S. (2009), Discretion rather than rules: choice of instruments to control bureaucratic policy making, *Political Analysis 17*, 25–44.

Gailmard, S. and Patty, J. W. (2012), Formal models of bureaucracy, *Annual Review of Political Science 15*, 353–377.

Gibbons, R., Matouschek, N. and Roberts, J. (2013), Decisions in organizations, in R. Gibbons and J. Roberts (eds), *The Handbook of Organizational Economics*, Princeton University Press, Princeton, NJ, 373–431.

Gilardi, F. (2009), *Delegation in the Regulatory State: Independent Regulatory Agencies in Western Europe*, Edward Elgar, Cheltenham.

Holmström, B. (1984), On the theory of delegation, in M. Boyer and R. Kihlstrom (eds), *Bayesian Models in Economic Theory*, North-Holland, Amsterdam, 115–141.

Huber, J. D. and Shipan, C. R. (2006), Politics, delegation, and bureaucracy, in D. Ritchie and B. Weingast (eds), *Oxford Handbook of Political Economy*, Oxford University Press, Oxford, 256–272.

Kundu, T. and Nilssen, T. (2020), Delegation of regulation, *Journal of Industrial Economics 68*, 445–482.

Laffont, J.-J. (1994), Regulation of pollution with asymmetric information, in C. Dosi and T. Tomasi (eds), *Nonpoint Source Pollution Regulation: Issues and Analysis*, Springer, Berlin, 39–66.

Laffont, J.-J. and Tirole, J. (1988), The dynamics of incentive contracts, *Econometrica 56*, 1153–1175.

Laffont, J.-J. and Tirole, J. (1990), Adverse selection and renegotiation in procurement, *Review of Economic Studies 57*, 597–625.

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

Lewis, T. R. (1996), Protecting the environment when costs and benefits are privately known, *RAND Journal of Economics 27*, 819–847.

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

Melumad, N. D. and Shibano, T. (1991), Communication in settings with no transfers, *RAND Journal of Economics 22*, 173–198.

Porteiro, N. (2008), Pressure groups and experts in environmental regulation, *Journal of Economic Behavior & Organization 65*, 156–175.

Voss, A. and Lingens, J. (2018), What's the damage? Environmental regulation with policy-motivated bureaucrats, *Journal of Public Economic Theory 20*, 613–633.

First version submitted December 2020;
final version received November 2021.