Delegation of Regulation*

Tapas Kundu[†,‡] Tore Nilssen[§]

January 16, 2019

Abstract

We develop a model to discuss a government's incentives to delegate to bureaucrats the regulation of an industry. The industry consists of a polluting firm with private information about its production technology. Implementing a transfer-based regulation policy requires the government to make use of a bureaucracy; this has a *bureaucratic cost*, as the bureaucracy diverts a fraction of the transfer. The government faces a trade-off in its delegation decision: bureaucrats have knowledge of the firms in the industry that the government does not have, but at the same time, they have other preferences than the government, so-called *bureaucratic drift*. We study how the bureaucratic drift and the bureaucratic cost interact to affect the incentives to delegate. Furthermore, we discuss how *partial delegation*, i.e., delegation followed by laws and regulations that restrict bureaucratic discretion, increases the scope of delegation. We characterize the optimal delegation rule and show that, in equilibrium, three different regimes can arise that differ in the extent of bureaucratic discretion. Our analysis has implications for when and how a government should delegate its regulation of industry. We find that bureaucratic discretion reduces with bureaucratic drift but that, because of the nature of the regulation problem, the effect of increased uncertainty about the firm's technology on the bureaucratic discretion depends on how that uncertainty is reduced.

Keywords: Bureaucracy; Delegation; Regulation.

*We are grateful for comments received from Manuel Amador, Clive Bell, Gerard Llobet, Vaiva Petrikaite, Miguel Puchades, Carl Shapiro, and Tina Søreide, and from audiences at PET 2014 in Seattle, EEA 2014 in Toulouse, EARIE 2014 in Milan, RES 2015 in Manchester, ESWC 2015 in Montreal, the Annual Conference on Economic Growth and Development 2015 in Delhi, EPCS 2016 in Freiburg, The Peder Sather Conference 2016 in Bergen, EALE 2016 in Bologna, the Barcelona GSE Summer Forum 2016, Jornadas de Economia Industrial 2016 in Mallorca, and the CESifo 2016 Workshop on Political Economy in Dresden, as well as at seminars at the Universities of Kent, Southern Denmark, and York. This research has received funding from the ESOP Centre at the University of Oslo, with which both authors are associated. ESOP has received support from the Research Council of Norway through its Centres of Excellence funding scheme, project number 179552.

[†]Oslo Business School, Oslo Metropolitan University. Email: tapas.kundu@oslomet.no.

[‡]School of Business and Economics, UiT the Arctic University of Norway.

[§]Department of Economics, University of Oslo. Email: tore.nilssen@econ.uio.no.

1 Introduction

Governments often delegate to bureaucrats to deal with industry. Delegation has contrasting effects, though. On one hand, society benefits from bureaucrats' industry-specific knowledge. On the other hand, society loses control over policy as non-elected officials will be making decisions. So when and how should such regulatory decisions be delegated? Without delegation, an incompletely informed government will have to resort to formulating a menu-based regulatory policy, so that low-cost firms receive an information rent and high-cost firms' production is distorted; see, e.g., Baron [3]. With delegation, regulation is carried out by an informed bureaucrat, and there is no longer a need to provide low-cost firms with an information rent. But the bureaucrat, if she is biased, will distort production for both low-cost and high-cost firms, relative to government's first best. In order to restrain these distortions, while still benefiting from the bureaucrat's knowledge, government can introduce restrictions on the bureaucrat's conduct, which we call *partial delegation*: various laws and rules to go with the bureaucrat's license to deal with industry.

In this paper, we set up a model to discuss which kind of such restrictions government may choose. We find that government will choose one of three options. One is not to delegate, because the bureaucrat's bias is too costly. A second option is what we call *weak delegation*: government puts a cap on the bureaucrat's choice set, so that undistorted production by a low-cost firm is ensured; this happens when the bureaucrat's bias is less costly. But we also find scope for a third option: When it is likely that the firm is high-cost, and/or the distortion that an unrestrained bureaucrat would impose on this firm type is big, the government will choose a stricter cap, which is based on the firm's expected cost; we call this *strict delegation*.

The need for capping the bureaucrat's actions based on expected cost rather than on that of a low-cost firm through strict delegation stems from government's informational disadvantage, leading to cases where capping based on the firm being low-cost is ineffective. The occurrence of strict delegation has two implications that distinguish our model from most of the previous literature. First, strict delegation implies that, when government delegates optimally, there is distortion at the top, in the sense that the firm, when it is low-cost, does not pollute at the first-best level. This is in contrast to most models of regulation under asymmetric information, where distortions are imposed only on firm types other than the most effective one. Secondly, strict delegation occurs mainly when the probability of the firm having high cost is high. This way, the possibility of strict delegation impacts on how government's incentives to delegate respond to changes in uncertainty. In our model, it is not the case that government necessarily responds to reduced uncertainty by allowing more delegation. Rather, we find that, if government becomes more certain that the firm has high cost, it will likely go for strict delegation, while it is more interested in going further into weak delegation if the firm most likely has low cost.

Our model has two key components: a regulated firm with private information about its production technology, and a bureaucrat who, if delegated the power to do so, will carry out the regulation of the firm on behalf of government. Consider first the bureaucracy. Whether or not decision power is delegated, the bureaucracy is there and constitutes a cost for government of regulating the firm.

- The bureaucracy handles any transfer of resources between the firm and the government, and it diverts a fraction of the transferred resources. We refer to this fraction as the *bureaucratic cost* and it exists whether or not regulation decisions are delegated. Bureaucratic leakage like this can appear in many forms in practice. Examples include everything from shirking and empire building, through inflated budgeting, to diversion of public funds for private use as well as outright corruption. Our assumption implies that the bureaucratic cost constrains the effectiveness of any transfer-based regulation policy, irrespective of whether an informed bureaucrat or an uninformed government determines the regulation policy.

When regulation is delegated in our model, it is to a bureaucrat with two features well-known from earlier analyses, particularly in the political-science literature: *bureaucratic expertise*; and *bureaucratic drift*.[1]

- First, a bureaucrat has an informational advantage over government in regulating an industry. A politician can choose a bureaucrat based on her skill and knowledge about the industry in question. Besides, a bureaucrat has a narrower agenda than that of a politician, and therefore she has higher incentives to gather information. We model this informational advantage by assuming that a bureaucrat can freely acquire information about the production technology while a government cannot at any cost. In case it does not delegate, the government faces a standard regulation problem under asymmetric information, leading to a combination of distorted firm behavior and information rents.

- Secondly, bureaucrats are in part motivated by self-interest: when decision-making authority is delegated to a bureaucrat, she pursues an objective different from that of government. In particular, we assume that, in case of delegation, the bureaucrat maximizes a weighted combination of the transfers diverted by bureaucracy, as discussed above, and government's objective. We refer to the weight on the transfers as the *bureaucratic drift*.

Consider next the regulated firm. We model a firm that is able to decrease its production costs by increasing its pollution. Government dislikes pollution and therefore regulates the firm, offering a higher revenue for the firm in return for less pollution. In the benchmark case of full information, pollution would be lower for the low-cost firm than for the high-cost firm. But while pollution is observable, government does not know the firm's production technology. Seen from government, the firm is one of two types: low or high production costs. Without delegation, government solves its lack of information by offering an incentive-compatible menu of regulatory contracts,

[1] See, e.g., Huber and Shipan [24] and Moe [36] for reviews.

i.e., combinations of transfers and pollution levels, subject to a participation constraint; this leads to an undistorted pollution level for the low-cost firm, together with an information rent, and an upwardly distorted pollution level for the high-cost firm. Transfers bring a bureaucratic cost, as discussed above, since a fraction of any amount sent by government to the firm is diverted by the bureaucracy.[2]

In case of delegation, the bureaucrat knows the firm's costs, and so no menu is required. The bureaucrat is biased in favor of transfers diverted by the bureaucracy, though. This means that she is more interested in transfers than government is, and more so the larger is the bias (the bureaucratic drift) and the larger is the diversion (the bureaucratic cost). Interestingly, this means that she prefers a lower level of pollution than does government, since less pollution requires a higher transfer. There is thus a *downward bias* in pollution level in case of delegation.[3] With full, or unrestrained, delegation there will therefore be distortions, relative to government's first-best, in the pollution levels for both the low-cost and the high-cost firm. The task for the government, if full and no delegation are the only feasible options, is to compare the costs of regulating the firm oneself (information rent to the good firm and an upward distortion of the pollution from the high-cost firm) to the cost of delegation (downward distortions in pollution for both types of firm).

In our analysis, we discuss the implications of allowing government to *delegate partially.* In particular, the government can delegate the decision to formulate the regulatory contract to the bureaucrat, but with an instruction to keep the pollution level within an interval. Since the bureaucrat's bias is downward, the crucial question is which proper lower bound on pollution the government should set. There are two options, we find: One is to set the lower bound at the first-best pollution level for the low-cost firm; we call this *weak delegation.* This keeps the bureaucrat in line as far as the low-cost firm goes, but gives her a lot of leeway in the regulation of the high-cost firm. But if the probability of the firm being high-cost is high, and/or if the bureaucrat's distortion of the high-cost firm is large, then there is an alternative that will be preferable: putting the lower bound so high that, in expectation across types, the firm's pollution level is first-best, and the bureaucrat responds by implementing the same pollution level for both types of firm; we call this *strict delegation.* Weak delegation has the feature that the low-cost pollution level is restored to the first best, while the high-cost pollution is lower than what government would prefer, due to the downward bias in the hand of the bureaucrat. With strict delegation, both the low-cost and high-cost pollution levels are distorted relative to first best; thus, when strict delegation occurs, the equilibrium outcome does not feature "no distortion at the top", which is otherwise a very common aspect of screening models with asymmetric information.

[2]This creates a shadow cost of public funds that is related to, but different from, the one modeled by Laffont and Tirole [30], who argue with the existence of distortionary taxes as a rationale for their shadow cost of public funds. In our set-up, there is no distortionary taxation.

[3]As we discuss in section 6.3 below, the same downward bias occurs, with results similar to those in our main analysis, if the bureaucrat has political preferences in favor of little pollution.

As we move from no delegation through strict delegation to weak delegation, there is an increase in the *extent of discretion* offered to the bureaucrat. This extent of discretion is decreasing in the bureaucratic drift and in the bureaucratic cost: the less troublesome the bureaucracy is, the more delegation will take place.

Another factor having an impact on the extent of discretion is government's uncertainty. This uncertainty is at the highest when the two types of the firm are equally likely. We find that there is a lot of discretion when uncertainty is high. However, the effect of reducing the uncertainty depends of which way it is reduced. If uncertainty goes down because the high-cost type gets more likely, then weak delegation performs poorly, and government may want to resort to strict delegation. If, on the other hand, uncertainty goes down because the low-cost type gets more likely, then strict delegation does not have a similar role to play, and government will to a large extent end with weak delegation.

Although we believe that the pollution that the bureaucrat allows in many cases will be easier for government to monitor than the transfers offered to the firm, we also discuss the consequences of having delegation being partial by having government putting bounds on the transfers rather than on pollution. An important difference from the case when bounds are on pollution is that the first-best transfer is not monotone in the firm's technology, while the first-best pollution level is. This difference affects the incentives to delegate. In particular, bureaucratic discretion is no longer monotonic in bureaucratic cost.

The result that there is no strict delegation when the probability of a low-cost firm is high is related to the downward bias of the bureaucrat. The downward bias leads to a need to have a lower bound on pollution in case of delegation. But when it is unlikely that the firm is high-cost, there is little scope for strict delegation to make a difference. This would be different in a case of regulation where the bureaucrat's bias is upward. To see this, we introduce a different set-up for the government-firm relationship. Instead of government regulating the firm, it offers pollution permits in return for transfers from the firm, i.e., transfers run from the firm through the bureaucracy to government, and more transfers mean more pollution. Now, in case of full delegation, we would have an upward bias: more pollution would mean higher fees to be paid and therefore more transfers, and transfers benefit the bureaucrat more than they do government. Our analysis of the permits case leads to an outcome similar to the one above, except when it comes to the occurrence of strict delegation. Now, it is when the probability of a high-cost firm is high that strict delegation has only a little role to play, and so we will see it only when the probability of the firm being low-cost is high.

The paper is organized as follows. In the next section, we present our basic model. While Section 3 contains some preliminary findings, Sections 4 and 5 analyze the full and partial delegation problems, respectively. In Section 6, we discuss how various factors in our model affect the partial delegation rule and the equilibrium regulation policy. Section 7 discusses related literature, and Section 8 concludes. Appendix A contains proofs of results omitted in the main text, and Appendix B includes a detailed analysis of the delegation problem in the case of permits.

2 The model

We consider the problem of regulating a polluting firm under asymmetric information. Following Laffont [27] and Boyer and Laffont [6], we assume that the firm's level of pollution is observable and verifiable. The firm provides a public project with a consumption utility G, and its cost of completing the project is

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

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 which is the firm's private information. With $C_\theta > 0$, a high θ implies a high cost and therefore low cost efficiency. For simplicity, we make the assumption that θ can take only two values, $\{\underline{\theta}, \overline{\theta}\}$, with $0 < \underline{\theta} < \overline{\theta} < K$. Let $\nu \in (0, 1)$ be the probability that the firm is low-cost with type $\theta = \underline{\theta}$.

Consumers are adversely affected by pollution, with a disutility given by $\frac{d^2}{2}$. The social value of the project is therefore

$$V(\theta, d) = G - \theta(K - d) - \frac{1}{2}d^2,$$

where $G \geq \frac{K^2}{2}$, ensuring that the social value of the project is non-negative even at maximum pollution, i.e., that $V(\theta, K) \geq 0$. Note that, for a given θ, the socially efficient pollution level is θ.

With the accounting convention which by now is standard in the literature (see, e.g., Laffont [27]), government is assumed to pay the firm's costs and to receive the proceeds from sales. It follows that a regulatory contract is a combination $\alpha = (t, d) \in A = \mathbb{R}_+ \times [0, K]$: by providing a transfer t to the firm, government compensates the firm for keeping its level of pollution at d. Given a contract $\alpha = (t, d)$, the firm's payoff is

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

We will consider pairs of contracts $(\underline{\alpha}, \overline{\alpha}) = \left((\underline{t}, \underline{d}), (\overline{t}, \overline{d})\right) \in A^2$, for the two types $\underline{\theta}$ and $\overline{\theta}$, that satisfy *incentive-compatibility* constraints:

$$\overline{t} - \overline{\theta}\left(K - \overline{d}\right) \geq \underline{t} - \overline{\theta}\left(K - \underline{d}\right), \tag{ICH}$$

$$\underline{t} - \underline{\theta}\left(K - \underline{d}\right) \geq \overline{t} - \underline{\theta}\left(K - \overline{d}\right). \tag{ICL}$$

We also assume that the firm's participation is voluntary so that contracts must be individually rational. A contract α satisfies the *individual-rationality* constraint if

$$t - \theta (K - d) \geq 0 \ . \tag{IR}$$

A pair of contracts $(\underline{\alpha}, \overline{\alpha})$ satisfy the individual-rationality constraints if

$$\bar{t} - \overline{\theta} \left(K - \overline{d}\right) \geq 0 \ , \tag{IRH}$$

$$\underline{t} - \underline{\theta} (K - \underline{d}) \geq 0 \ . \tag{IRL}$$

We assume that the implementation of the regulation has a bureaucratic cost. For every unit of fund raised, the firm receives only a fraction $(1 - \lambda)$ of it, where $\lambda \in (0, 1)$, and the remaining fraction λ is consumed by the bureaucracy.[4] Therefore, in order to make a transfer of t to the firm, government has to raise public funds of $\frac{t}{1-\lambda}$, out of which $\frac{\lambda t}{1-\lambda}$ is consumed in bureaucracy. The payoff to government from a contract $\alpha = (t, d)$ is thus given by

$$U_C (t, d) \quad = \quad G - \frac{1}{2} d^2 - \frac{t}{1 - \lambda} \ . \tag{2}$$

Government can delegate the regulatory decision-making to an independent regulator, a bureaucrat B. We assume that the bureaucrat is informed about the firm's cost. She can therefore implement a type-contingent regulatory policy. If government delegates, then the bureaucrat has authority to choose a regulatory policy according to her own preferences. We assume that B has a vested interest in that fraction of the transfer that is consumed in bureaucracy. Specifically, the bureaucrat's payoff is a weighted average of this fraction and the payoff to government:

$$\begin{aligned} U_B (\theta, t, d) &= \beta \frac{\lambda t}{1 - \lambda} + (1 - \beta) U_C (t, d) \\ &= (1 - \beta) \left[\left(G - \frac{1}{2} d^2 \right) - \left(1 - \frac{\lambda \beta}{1 - \beta} \right) \frac{t}{1 - \lambda} \right] , \end{aligned} \tag{3}$$

where $\beta \in (0, 1)$ measures the extent of the bureaucrat's rent-seeking motivation.

If $\frac{\lambda\beta}{1-\beta} > 1$, or equivalently $\beta > \frac{1}{1+\lambda}$, then the bureaucrat's payoff is increasing in t. In this case, B will maximize her payoff by choosing an infinitely high transfer, without adversely affecting the firm's participation constraint. We impose a restriction on how misaligned the bureaucrat and government can be to ensure that this does not happen:

Assumption 1. $\beta \leq \frac{1}{1+\lambda}$.

The game proceeds as follows.

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

[4] Public transfer often involves other forms of distortionary cost, such as distortion caused by taxation. We disregard such costs and focus on the bureaucratic cost.

government.

- Stage 2. The firm learns its type θ, which can be either $\underline{\theta}$ with probability ν or $\overline{\theta}$ with probability $1-\nu$. B also learns the firm's type.
- Stage 3. The player with decision making authority – government or bureaucrat – determines the regulatory policy.
- Stage 4. Production takes place. Payoffs are realized. The game ends.

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

3 Preliminary analysis

We first describe the regulatory contract that government chooses if it has complete information about θ. The contract for type θ solves the following problem:

$$\max_{\alpha} G - \frac{1}{2}d^2 - \frac{t}{1-\lambda} \tag{4}$$
$$\text{subject to (IR).}$$

We denote the solution with subscript CI. Clearly, the firm's participation constraint is binding, so we write $t = \theta(K-d)$. Replacing t in (4), we find from the first-order condition that the optimal contract $\alpha_{CI}(\theta) = (t_{CI}(\theta), d_{CI}(\theta))$ is given by

$$\begin{aligned} d_{CI}(\theta) &= \min\left\{\frac{\theta}{1-\lambda}, K\right\}, \\ t_{CI}(\theta) &= \theta\left[K - d_{CI}(\theta)\right]. \end{aligned} \tag{5}$$

Observe that, when $\lambda \geq 1 - \frac{\theta}{K}$, the pollution is at the maximum and there is no transfer: $\alpha_{CI}(\theta) = (0, K)$. In this case, essentially, government lets the producer go unregulated, with no transfer 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 2. $\lambda \leq 1 - \frac{\overline{\theta}}{K}$.

As we will see below, this Assumption still allows for no-regulation contracts under asymmetric information.

With Assumption 2 in place, we can write

$$d_{CI}(\theta) = \frac{\theta}{1-\lambda}.$$

As $d_{CI}(\theta)$ is increasing in θ, the cost efficient low-cost firm is also efficient in reducing pollution.[5] The actual cost of production, and consequently, the compensating transfer $t_{CI}(\theta)$ is however not monotone in θ. We have $t_{CI}(\underline{\theta}) \leq t_{CI}(\overline{\theta})$ if and only if $\lambda \leq 1 - \frac{\overline{\theta}+\underline{\theta}}{K}$. The low-cost firm receives less (more) transfer than the high-cost firm for low (high) values of λ.

3.1 Regulation by government under asymmetric information

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

$$\max_{\underline{\alpha},\overline{\alpha}} \nu \left[G - \frac{1}{2}\underline{d}^2 - \frac{\underline{t}}{1-\lambda} \right] + (1-\nu) \left[G - \frac{1}{2}\overline{d}^2 - \frac{\overline{t}}{1-\lambda} \right] \tag{6}$$

subject to (IRH), (IRL), (ICH), and (ICL).

We denote the solution with subscript CN. Define $\triangle\theta := \overline{\theta} - \underline{\theta}$. The following Lemma describes the contract pair.

Lemma 1. *Consider the case of no delegation. The contract pair*

$$\left(\alpha_{CN}(\underline{\theta}), \alpha_{CN}(\overline{\theta})\right) = \left(\left(t_{CN}(\underline{\theta}), d_{CN}(\underline{\theta})\right), \left(t_{CN}(\overline{\theta}), d_{CN}(\overline{\theta})\right)\right)$$

that government offers to the firm is given by

$$\begin{aligned} d_{CN}(\underline{\theta}) &= \frac{\underline{\theta}}{1-\lambda}, \\ d_{CN}(\overline{\theta}) &= \min\left\{ \frac{1}{1-\lambda}\left(\overline{\theta} + \frac{\nu}{1-\nu}\triangle\theta\right), K \right\}, \\ t_{CN}(\underline{\theta}) &= \underline{\theta}\left(K - d_{CN}(\underline{\theta})\right) + \triangle\theta\left(K - d_{CN}(\overline{\theta})\right), \\ t_{CN}(\overline{\theta}) &= \overline{\theta}\left(K - d_{CN}(\overline{\theta})\right). \end{aligned} \tag{7}$$

Proof. In Appendix A. □

In absence of complete information, government shares an information rent with the low-cost firm. In addition, there is a distortion in the pollution level set for high-cost firms.

Note that, from (7), we have $d_{CN}(\overline{\theta}) = K$ if ν is sufficiently large, in particular if

[5] This feature follows from our assumptions on the cost function: $C_\theta > 0$ and $C_{\theta d} < 0$. The increasing property of $d_{CI}(\theta)$ can be reversed with alternative assumptions: $C_\theta > 0, C_{\theta d} > 0$, as in $C(\theta, d) = K - d/\theta$; or $C_\theta < 0, C_{\theta d} < 0$, as in $C(\theta, d) = K - \theta d$. Our assumptions are, however, consistent with previous related work, such as Boyer and Laffont [6]. As they point out, "with a one-dimensional asymmetry of information, the positive correlation between ability to produce and to reduce pollution seems more compelling than the alternative assumption" (p. 140). It is, however, worth noting that the monotonicity of $d_{CI}(\theta)$ does not depend on a specific sign of $C_{\theta d}$, rather it follows from $C_{\theta d}$ having the same sign for all θ and d.

$$\nu \geq \nu^* := \frac{(1-\lambda) K - \overline{\theta}}{(1-\lambda) K - \underline{\theta}} \in [0,1], \tag{8}$$

which is decreasing in λ. With $d = K$, we have $t = 0$ and $C(\cdot, K) = 0$, so that both revenue and cost, and hence the firm's profit, equal zero; in effect, government lets the high-cost firm go unregulated. Thus, while Assumption 2 ensures that this does not happen under complete information, that Assumption still allows for it under asymmetric information.

3.2 Regulation by bureaucrat

Consider a case where an informed bureaucrat regulates the firm without any restrictions. The contract for type θ solves the following problem:

$$\begin{aligned} \max_{\alpha} \; & (1-\beta)\left[\left(G - \frac{1}{2}d^2\right) - \left(1 - \frac{\lambda\beta}{1-\beta}\right)\frac{t}{1-\lambda}\right], \\ \text{subject to } \; & \text{(IR)}. \end{aligned} \tag{9}$$

We denote the solution with subscript BI. The following Lemma describes the contract.

Lemma 2. *Assume that 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 $\theta \in \{\underline{\theta}, \overline{\theta}\}$ is given by*

$$\begin{aligned} d_{BI}(\theta) &= \left(1 - \frac{\lambda\beta}{1-\beta}\right)\frac{\theta}{1-\lambda} \\ t_{BI}(\theta) &= \theta\left(K - d_{BI}(\theta)\right). \end{aligned} \tag{10}$$

Proof. Under Assumption 1, the bureaucrat's objective function is decreasing in t, so that the firm's participation constraint will be binding and we can write $t = \theta(K - d)$. Replacing t in (9) and solving the maximization problem, we obtain the result. □

The bureaucrat's choice of pollution level is always below government's full-information choice because of her vested interest in transfer. By setting pollution at a lower level, the bureaucrat can increase the production cost, and thereby the compensatory transfer. Note that Assumption 1 ensures that her optimal pollution level is always non-negative.

4 Full delegation

Full delegation refers to a case where government delegates decision-making authority to a bureaucrat without imposing any restriction on the latter's choice set. The condition under which government prefers no delegation to full delegation is

$$\triangle D := \quad E_\theta U_C\left(\alpha_{CN}\left(\theta\right)\right) - E_\theta U_C\left(\alpha_{BI}\left(\theta\right)\right) \quad > 0, \tag{11}$$

where, for an arbitrary function $g\left(\cdot\right)$ of θ, we let

$$E_\theta g\left(\theta\right) := \nu g\left(\underline{\theta}\right) + \left(1-\nu\right) g\left(\overline{\theta}\right).$$

Below, we discuss how the sign of $\triangle D$ changes with respect to β, λ, and ν.

The effect of β is straightforward. Government's payoff under full delegation decreases with β whereas β has no impact on that payoff under no delegation. Therefore, government prefers no delegation to full delegation if and only β is above a threshold. This finding is similar to the *ally principle* (e.g., Huber and Shipan [24]), which suggests government prefers to give more discretion to more aligned bureaucrats.

The bureaucratic cost λ has two contrasting effects on $\triangle D$: on one hand, the adverse effect of policy distortion from delegation increases with λ; on the other hand, the positive effect of not sharing information rent in delegation reduces with λ. The two effects interact in a way that can change $\triangle D$ non-monotonically. However, it can be shown that $\triangle D$ changes its sign only once. Furthermore, for large values of λ, the first effect dominates the second one, and government therefore prefers no delegation to full delegation if and only if λ is above a threshold.

How does ν affect delegation? Since $\triangle D$ is weakly convex in ν and takes positive values at $\nu = 0$ and $\nu = 1$, $\triangle D$ can only take negative values at an intermediate range of ν, i.e, when the uncertainty about the firm's type is high. This is because government's benefit from the bureaucrat's informational advantage is high in situations with high uncertainty. Our result is consistent with the *uncertainty principle*, which suggests that government prefers more bureaucratic discretion in situations with high uncertainty (Huber and Shipan [24]).

The following Proposition documents the above findings.

Proposition 1. *Consider the game in which government chooses between the alternatives of full delegation and no delegation. The equilibrium is characterized as follows:*

(i) For given λ and ν, there exists a threshold $\overline{\beta}$ such that no delegation occurs if and only if $\beta \geq \overline{\beta}$.

(ii) *For given β and ν, there exists a threshold $\overline{\lambda}$ such that no delegation occurs if and only if $\lambda \geq \overline{\lambda}$.*

(iii) *For given λ and β, there exist $0 < \underline{\nu}^{FD} \leq \overline{\nu}^{FD} < 1$ such that full delegation occurs if and only if $\underline{\nu}^{FD} \leq \nu \leq \overline{\nu}^{FD}$. The interval $\left[\underline{\nu}^{FD}, \overline{\nu}^{FD}\right]$ can be a null set for large values of β.*

Proof. In Appendix A. □

Below we present a numerical example.

Example 1. Consider an example with $G = 50$, $K = 10$, $\overline{\theta} = 4$, and $\underline{\theta} = 2$. The feasible range of λ, satisfying Assumption 2, is $[0, 0.6]$. Figure 1 plots government's preferences over full delegation and no delegation in (β, λ) space, with $\nu = 0.5$, $0 \leq \lambda \leq 0.6$, and $0 \leq \beta \leq \frac{1}{1+\lambda}$. The FD area represents parameter values for which government prefers full delegation to no delegation. Figure 2 plots government's preferences over full delegation and no delegation in (ν, λ) space, with $\beta = 0.5$, $0 \leq \lambda \leq 0.6$, and $0 \leq \nu \leq 1$.

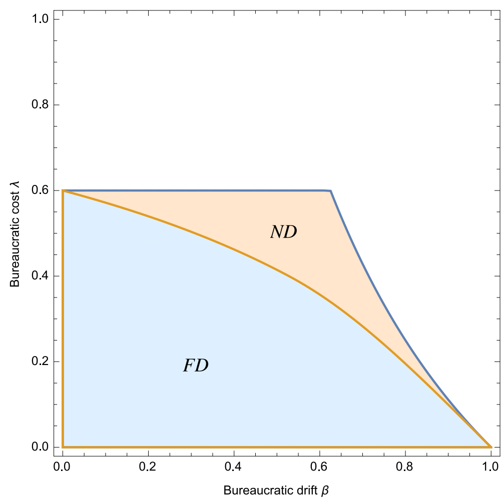

Figure 1. No delegation (ND) vs full delegation (FD) in (β, λ) space

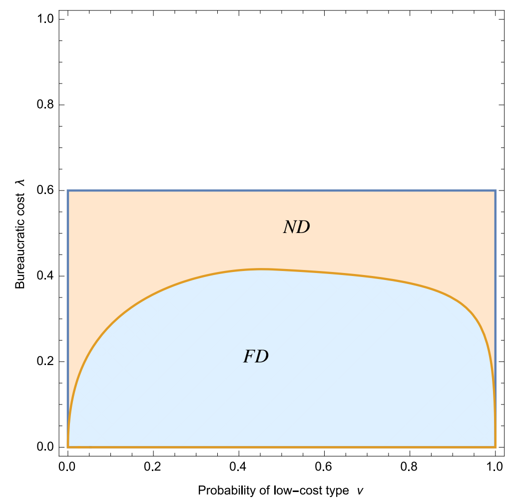

Figure 2. No delegation (ND) vs full delegation (FD) in (ν, λ) space

5 Partial delegation

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

This notion of partial delegation resembles interval delegation (see, e.g., Alonso and Matouschek [1] and Amador and Bagwell [2]). Since the task to be delegated is one of regulation, we have on one hand a multi-dimensional action space and on the other hand a two-type information issue. Here, we do not consider whether interval regulation, or any multidimensional equivalent (such as, e.g., in Frankel [14]), is optimal and limit our attention to the above notion of partial delegation.[6]

[6] The alternative form of partial delegation, putting constraints on transfers rather than on pollution, is discussed in Section 6.1 below. A discussion of doing both, that is, having constraints on both pollution and transfers,

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

$$\max_{\alpha} \ (1-\beta)\left[\left(G-\frac{1}{2}d^2\right)-\left(1-\frac{\lambda\beta}{1-\beta}\right)\frac{t}{1-\lambda}\right], \tag{12}$$
$$\text{subject to} \quad (\text{IR}), \ \text{ and } \ d\in[d_1,d_2].$$

We denote the solution with a superscript P and a subscript BI. The following Lemma describes the bureaucrat's optimal choice of contracts under partial delegation.

Lemma 3. *Assume that government delegates the decision making authority with the restriction that $d\in[d_1,d_2]\subseteq[0,K]$. The bureaucrat's preferred regulation contract for type $\theta\in\{\underline{\theta},\overline{\theta}\}$ is given by $\alpha_{BI}^P(\theta,d_1,d_2)=\left(t_{BI}^P(\theta,d_1,d_2),d_{BI}^P(\theta,d_1,d_2)\right)$, where*

$$\begin{aligned}
d_{BI}^P(\theta,d_1,d_2) &= \begin{cases} d_1, & \text{if } d_1\geq d_{BI}(\theta);\\ d_{BI}(\theta)=\left(1-\frac{\lambda\beta}{1-\beta}\right)\frac{\theta}{1-\lambda}, & \text{if } d_1<d_{BI}(\theta)<d_2;\\ d_2, & \text{if } d_{BI}(\theta)\geq d_2.\end{cases}\\
t_{BI}^P(\theta,d_1,d_2) &= \theta\left(K-d_{BI}^P(\theta,d_1,d_2)\right).
\end{aligned}$$

Proof. Follows from replacing t by $\theta(K-d)$ in (12) and using the first-order condition of the optimization problem. □

The bureaucrat's choice of contract under partial delegation coincides with her choice under full delegation if the latter lies in the bounded interval $[d_1,d_2]$; otherwise, the optimal choice lies at the boundaries. The government can therefore affect her choice by manipulating d_1 and d_2.

Next, we study government's choice of d_1 and d_2. The following lemma describes the choice for the upper bound.

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

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

Disregarding government's indifference, we simply put its choice at $d_2=\max\{d_1,d_{CI}(\overline{\theta})\}$. Recall that $d_{CI}(\overline{\theta})$ is government's preferred pollution level for the high-cost firm under full information, and that $d_{CI}(\overline{\theta})>d_{BI}(\overline{\theta})$. The following lemma describes potential choices for the optimal lower bound.

would require enriching the present model to have more than two firm types and is left for future research.

Lemma 5. *Fix* $d_2 = \max\{d_1, d_{CI}(\overline{\theta})\}$*. Suppose government partially delegates with a restriction that* $d(\underline{\theta}), d(\overline{\theta}) \in [d_1, d_2]$*, for some* $d_1 \in [0, d_2]$*. If* $d_{BI}(\overline{\theta}) \leq d_{CI}(\underline{\theta})$*, then, among all* $d_1 \in [0, d_2]$*, government's payoff is maximized at* $d_1 = \frac{E_\theta \theta}{1-\lambda} = d_{CI}(E_\theta \theta)$*. If* $d_{BI}(\overline{\theta}) > d_{CI}(\underline{\theta})$*, then, among all* $d_1 \in [0, d_{BI}(\overline{\theta})]$*, it is maximized at* $d_1 = d_{CI}(\underline{\theta})$*, while among all* $d_1 \in (d_{BI}(\overline{\theta}), d_2]$*, it is maximized at* $d_1 = \frac{E_\theta \theta}{1-\lambda} = d_{CI}(E_\theta \theta)$*.*

Proof. In Appendix A. □

Lemmata 3, 4, and 5 together imply that, when government partially delegates, two possibilities may arise in equilibrium.

- *Weak Delegation* (WD): In this regime, government chooses $d_1 = d_{CI}(\underline{\theta})$ and $d_2 = d_{CI}(\overline{\theta})$. In response, the bureaucrat sets $d_{BI}^P(\underline{\theta}, d_1, d_2) = d_{CI}(\underline{\theta})$ and $d_{BI}^P(\overline{\theta}, d_1, d_2) = d_{BI}(\overline{\theta})$. Government implements the full-information contract if the firm is low-cost. There is distortion at the contract offered to a high-cost firm, as $d_{BI}(\overline{\theta}) < d_{CI}(\overline{\theta})$. We refer to this regime as weak delegation.

- *Strict Delegation* (SD): In this regime, government chooses $d_1 = d_{CI}(E_\theta \theta) > d_{BI}(\overline{\theta})$. In response, the bureaucrat sets $d_{BI}^P(\underline{\theta}, d_1, d_2) = d_{BI}^P(\overline{\theta}, d_1, d_2) = d_{CI}(E_\theta \theta)$, resulting in a uniform pollution level for both types of firm. Government's choice of d_1 is the optimal uniform pollution level. Also note that the bureaucrat's choice of pollution does not depend on the upper bound d_2 for any $d_2 \geq d_{CI}(E_\theta \theta)$. In this case, we can therefore write government's optimal delegation strategy as setting $d_1 = d_2 = d_{CI}(E_\theta \theta)$, and the bureaucrat's discretion is limited to setting the transfer. We refer to this regime as strict delegation.

The following proposition documents the above findings.

Proposition 2. *Suppose that government partially delegates with a restriction that* $d(\underline{\theta}), d(\overline{\theta}) \in [d_1, d_2] \subseteq [0, K]$*. Government's partial-delegation rule takes one of the following two forms:*

(i) $d_1 = d_{CI}(\underline{\theta})$ *and* $d_2 = d_{CI}(\overline{\theta})$*, in which case government implements the full-information contract to a low-cost firm and there is distortion to the contract offered to a high-cost firm;*

(ii) $d_1 = d_2 = d_{CI}(E_\theta \theta)$*, in which case government implements a uniform pollution level.*

These findings are similar to those in Melumad and Shibano ([34], MS hereafter) and Alonso and Matouschek ([1], AM hereafter). Both MS and AM consider the delegation problem in a framework with a one-dimensional action space, a quadratic loss function, and state-dependent biases. Our model reduces to a one-dimensional set-up when the producer's participation constraint binds. In this reduced set-up, the players' payoff functions share common features with those studied in MS and AM. When the participation constraint binds, we can rewrite the respective payoff functions of the bureaucrat and the government as

$$U_B = F_B(\theta) - \frac{1-\beta}{2}\left(d - (1-b)\, d_{CI}(\theta)\right)^2,$$
$$U_C = F_C(\theta) - \frac{1}{2}\left(d - d_{CI}(\theta)\right)^2,$$

where

$$F_B(\theta) = (1-\beta)\left(G - K(1-b)\, d_{CI}(\theta) + \frac{1}{2}(1-b)^2 d_{CI}^2(\theta)\right),$$
$$F_C(\theta) = G - K d_{CI}(\theta) + \frac{1}{2} d_{CI}^2(\theta),$$

and $b = \frac{\lambda\beta}{1-\beta}$ is a measure of the conditional preference divergence.[7] Our descriptions of the two delegation regimes also have interpretations similar to those found in MS. MS define a decision rule as communication independent if it does not distinguish between environments in different states and as communication dependent otherwise. A weak (strict) delegation regime in our framework is qualitatively similar to a communication-dependent (communication-independent) rule as, in this regime, the implemented pollution level varies (does not vary) with the state.

Below, we continue the numerical example to illustrate the two possibilities.

Example 2. We continue with the same parametric specification ($G = 50$, $K = 10$, $\overline{\theta} = 4$, and $\underline{\theta} = 2$) as used in Example 1. In addition, assume $\lambda = 0.25$, $\beta = 0.65$, and $\nu = 0.5$. Figure 3 plots the full-information regulation contract, government's preferred contract under no delegation, and the bureaucrat's preferred contract under full delegation, in (d, t) space. The straight lines represent the firm's individual-rationality constraints, with the steeper one corresponding to a high-cost firm, and the firm's profit increasing in the top-right direction. The dashed curves and the dot-dashed curves represent the indifference curves of government and the bureaucrat, respectively, with utility increasing in the bottom-left direction. The points A, B, C, D, E, and F represent the regulatory contracts $\alpha_{CI}(\underline{\theta})$, $\alpha_{CI}(\overline{\theta})$, $\alpha_{CN}(\underline{\theta})$, $\alpha_{CN}(\overline{\theta})$, $\alpha_{BI}(\underline{\theta})$, and $\alpha_{BI}(\overline{\theta})$, respectively. In Figure 4 (which is comparable to Figure 3), we illustrate the two possibilities that government can induce through partial delegation. In our example, $d_{CI}(\underline{\theta}) = 2.67$, $d_{CI}(\overline{\theta}) = 5.33$, and $d_{CI}(E_\theta \theta) = 4$. With partial delegation, government can either implement contracts A and F by setting d_1 at 2.67 (weak delegation), or implement contracts G and H by setting d_1 at 4 (strict delegation). The shaded area in Figure 4 shows the bureaucrat's choice set under weak delegation whereas the dashed vertical line in Figure 4 corresponds to the choice set under strict delegation. In this example, government's expected value turns out to be higher at the pair $\{G, H\}$, i.e., strict delegation, implying the optimal $d_1 = 4$.

[7] While MS consider a form of divergence of preferences similar to ours, AM allow for more general forms of divergence. Also, note that the action variable d and the state variable θ appear in a non-additive way in the expression of bias, which is typically measured by $U_C - U_B$, the difference between the payoffs of the two players.

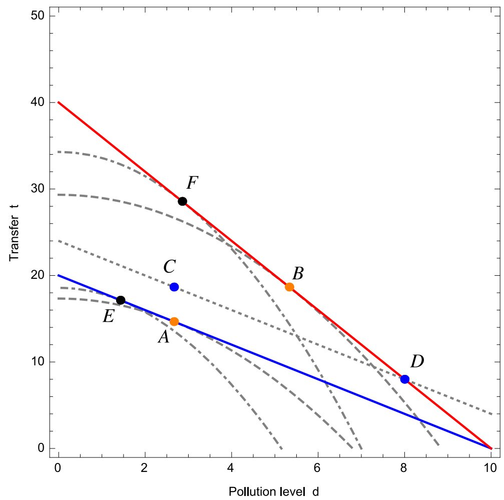

Figure 3. The regulation contracts in (d, t) space

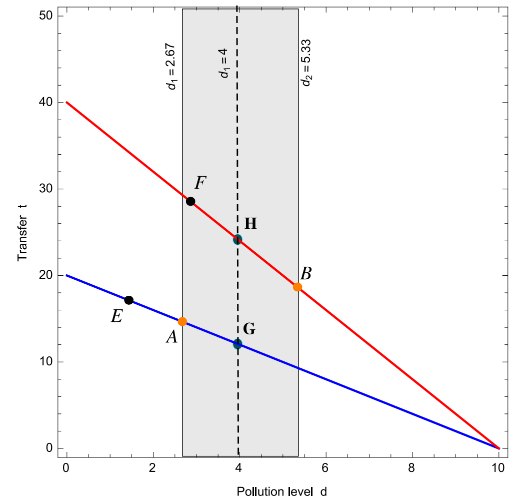

Figure 4. Partial delegation in (d, t) space

Based on government's expected payoffs in the various cases, we observe three possible regimes in equilibrium − weak (WD), strict (SD), and no delegation (ND). The following proposition fully characterizes how different regimes can arise in equilibrium.

Proposition 3. *Consider the game in which government chooses between partial delegation and no delegation. The equilibrium regime is characterized as follows:*

(i) Fix ν. There exists a threshold $\lambda^D(\beta)$, decreasing in β, and a constant λ^{ND} such that weak delegation occurs if $\lambda < \lambda^D(\beta)$; no delegation occurs if $\lambda \geq \max\left\{\lambda^D(\beta), \lambda^{ND}\right\}$; and strict delegation occurs if $\lambda^D(\beta) < \lambda \leq \lambda^{ND}$.

(ii) *Fix λ and β. Define $\hat{\nu} := \min\left\{\left(\frac{\lambda\beta\overline{\theta}}{(1-\beta)\triangle\theta}\right)^2, 1\right\} \in [0, 1]$. Government prefers strict delegation to weak delegation if and only if $\nu \leq \hat{\nu}$. For $\nu \leq \hat{\nu}$, there exists a threshold $\underline{\nu}^{SD} \in [0, \hat{\nu}]$ such that strict delegation occurs in equilibrium if $\nu \leq \underline{\nu}^{SD}$; and no delegation occurs in equilibrium if $\nu \in \left[\underline{\nu}^{SD}, \hat{\nu}\right]$. For $\nu \geq \hat{\nu}$, there exist threshold values $\underline{\nu}^{WD}$ and $\overline{\nu}^{WD}$, with $\hat{\nu} \leq \underline{\nu}^{WD} \leq \overline{\nu}^{WD} \leq 1$, such that weak delegation occurs in equilibrium if and only if $\underline{\nu}^{WD} \leq \nu \leq \overline{\nu}^{WD}$; and no delegation occurs in equilibrium otherwise.*

Proof. In Appendix A. □

Bureaucratic discretion is higher under weak delegation than under strict delegation, and there is, of course, zero bureaucratic discretion when no delegation happens. Taken together, Propositions 2 and 3 detail the extent of bureaucratic discretion that government chooses. Government's benefit from weak delegation is the ability to implement the first-best contract if the firm is low-cost. However, with weak delegation, there is a distortion in the contract offered to a high-cost firm. This distortion depends on the bureaucratic drift and the bureaucratic cost. It follows that when the high-cost type is likely and/or the distortions are high, government is better off giving up the first-best contract for the low-cost firm and instead opt for strict delegation. This is illustrated in Figure 4.

In particular, the following observations based on Proposition 3 are worth noting. First, bureaucratic discretion reduces with bureaucratic cost λ and also with bureaucratic drift β, as the threshold $\lambda^{D}(\beta)$ is decreasing in β. With either strict or no delegation, B has no discretion in determining the pollution level. Therefore government's payoff is independent of the bureaucratic drift β. At $\lambda = \lambda^{ND}$, government is indifferent between strict and no delegation. We observe strict delegation in equilibrium only if $\lambda^{ND} > \lambda^{D}(\beta)$.

Secondly, bureaucratic discretion is also affected by uncertainty. However, the nature of the equilibrium delegation rule depends on how uncertainty changes $-$ whether the firm becomes more likely to be low-cost or high-cost. In particular, if the firm is more likely to be low-cost (so that $\nu > \hat{\nu}$), then government chooses either weak or no delegation. This is because, in both these regimes, it implements the full-information regulation contract for the low-cost firm. The thresholds $\underline{\nu}^{WD}$ and $\overline{\nu}^{WD}$ can coincide, though, in which case we do not observe weak delegation in equilibrium for any ν. The thresholds can also take boundary values, in which case we observe only one regime, either weak delegation or no delegation for every $\nu > \hat{\nu}$. In contrast, if a firm is more likely to be high-cost (so that $\nu \leq \hat{\nu}$) , then government chooses strict delegation rather than no delegation if ν is lower than a threshold value. The threshold can also take boundary values 0 or $\hat{\nu}$, in which case we observe only one regime, either strict delegation or no delegation, for every $\nu \leq \hat{\nu}$.

Below we present a numerical example to demonstrate the equilibrium regimes.

Example 3. We continue with the same parametric specification ($G = 50$, $K = 10$, $\overline{\theta} = 4$, and $\underline{\theta} = 2$) as used in Example 1. The feasible range of λ, satisfying Assumption 2, is $[0, 0.6]$. Figure 5 plots the equilibrium regimes in (β, λ) space, with $\nu = 0.5$, $0 \leq \lambda \leq 0.6$ and $0 \leq \beta \leq \frac{1}{1+\lambda}$. Figure 6 plots the equilibrium regimes in (ν, λ) space, with $\beta = 0.5$, $0 \leq \lambda \leq 0.6$ and $0 \leq \nu \leq 1$. In this example, $\hat{\nu} = 4\lambda^2$. Therefore, government prefers strict delegation to weak delegation if $\nu \leq 4\lambda^2$. Moreover, $\nu^* = \frac{3-5\lambda}{4-5\lambda}$ in this example, so that $\nu^* < \hat{\nu}$ if and only if $\lambda > 0.37$. Thus, in the comparison between no delegation and weak delegation for $\nu > \hat{\nu}$, we have $u_{CN}\left(\bar{\theta}\right) = K$ as long as $\lambda > 0.37$. Note that, at $u_{CN}\left(\bar{\theta}\right) = K$, the sign of the payoff difference between no delegation and weak delegation does not depend on ν.[8] This explains why, in Figure 6, the curve delineating weak delegation and no delegation is a horizontal line. The dot-dashed curves in the two Figures show, from Figures 1 and 2, the parameter values at which C is indifferent between full delegation and no delegation in the full-delegation game. The scope of delegation increases with the possibility of partial delegation.

We are now in a position to summarize how the equilibrium regulation policy changes across the various delegation regimes. With both weak and strict delegation, the bureaucrat implements type-contingent regulation contracts. With no delegation, government implements a pair

[8] This can be seen from equation (A10) in Appendix A.

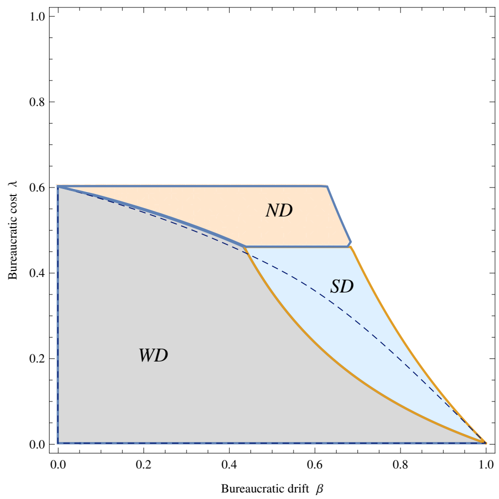

Figure 5: Optimal partial delegation in (β, λ) space

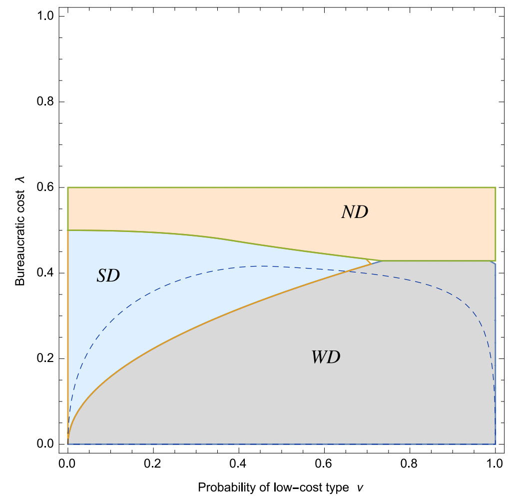

Figure 6: Optimal partial delegation in (ν, λ) space

of incentive-compatible contracts that allow the high-cost firm to over-pollute and leave an information rent with the low-cost firm. The low-cost firm faces the following regulation contract:

$$\alpha(\underline{\theta}) = (t(\underline{\theta}), d(\underline{\theta})) = \begin{cases} (t_{CI}(\underline{\theta}), d_{CI}(\underline{\theta})), & \text{with weak delegation;} \\ (\underline{\theta}(K - d_{CI}(E_\theta \theta)), d_{CI}(E_\theta \theta)), & \text{with strict delegation;} \\ \left(t_{CI}(\underline{\theta}) + \triangle\theta\left(K - d_{CN}\left(\overline{\theta}\right)\right), d_{CI}(\underline{\theta})\right), & \text{with no delegation.} \end{cases}$$

The high-cost firm faces the following regulation contract:

$$\alpha\left(\overline{\theta}\right) = \left(t\left(\overline{\theta}\right), d\left(\overline{\theta}\right)\right) = \begin{cases} \left(t_{BI}\left(\overline{\theta}\right), d_{BI}\left(\overline{\theta}\right)\right), & \text{with weak delegation;} \\ \left(\overline{\theta}(K - d_{CI}(E_\theta \theta)), d_{CI}(E_\theta \theta)\right), & \text{with strict delegation;} \\ \left(t_{CN}\left(\overline{\theta}\right), d_{CN}\left(\overline{\theta}\right)\right), & \text{with no delegation.} \end{cases}$$

Interestingly, when strict delegation is chosen by government, the classic result of regulation theory, that the optimum contract features no distortion at the top, no longer holds: with high bureaucratic cost and/or drift and also a high probability of the firm being high-cost, government prefers putting such a strict cap on the bureaucrat's activities that contracts are distorted for both firm types. In this way, allowing for partial delegation opens up for novel theoretical predictions on how government tackles the challenge of regulating industry.

To see how equilibrium contracts change with parameters, consider the effect of uncertainty. As Proposition 3 shows, ν affects which of the three regimes will occur in equilibrium. When ν is low, in particular, when $\nu \leq \hat{\nu}$, government chooses between strict and no delegation. We have strict delegation if $\nu \leq \min\left[\hat{\nu}, \underline{\nu}^{SD}\right]$, when an increase in ν lowers $E_\theta \theta$ and therefore lowers the pollution level of both firm types, $d(E_\theta \theta)$, until $\nu = \min\left[\hat{\nu}, \underline{\nu}^{SD}\right]$; this is the case when there is distortion at the top. If $\nu \in \left[\underline{\nu}^{SD}, \hat{\nu}\right]$, then we have no delegation, the pollution level of the low-cost firm at its first-best and independent of ν, and that of the high-cost firm increasing in ν until it reaches K. When ν increases further, so that it becomes $\nu > \hat{\nu}$, we might still have no delegation, with the features just described. But if ν enters the range $\left[\underline{\nu}^{WD}, \overline{\nu}^{WD}\right]$, then we have weak delegation with the pollution level of the low-cost firm being at the first-best level and that of the high-cost at the bureaucrat's preferred level, lower than the first-best level, both being independent of ν.

6 Discussion

Here, we extend the analysis in three directions. First, in Section 6.1, we discuss the consequences of having partial delegation by putting constraints on transfers rather than on pollution. Thereafter, in Section 6.2, we discuss the consequences of having a different regulation problem, in particular, one where the bureaucrat's bias in terms of pollution is an upward one rather than a downward one. Finally, in Section 6.3, we discuss other bases for bureaucratic drift than a preference for high transfers.

6.1 Restrictions on transfers

We now relax our assumption of no constraints on transfer. The pollution level, on the other hand, can be chosen by the bureaucrat without any constraints.[9] Specifically, assume now that the bureaucrat chooses regulatory contracts $\alpha(\theta) = (t(\theta), d(\theta)), \theta \in \{\underline{\theta}, \overline{\theta}\}$, under the constraint that $t(\theta) \in [t_1, t_2]$. Her optimal contract for type θ solves, instead of the one in (12), the following problem:

$$\max_{\alpha}\ (1-\beta)\left[\left(G - \frac{1}{2}d^2\right) - \left(1 - \frac{\lambda\beta}{1-\beta}\right)\frac{t}{1-\lambda}\right], \tag{13}$$
$$\text{subject to (IR), and } t \in [t_1, t_2].$$

We denote the solution with a superscript R and a subscript BI. Recall, from (10), that the bureaucrat's choice of transfer, in the absence of a transfer restriction, is

$$t_{BI}(\theta) = \theta\left(K - \frac{\theta}{1-\lambda} + \frac{\lambda\beta\theta}{(1-\lambda)(1-\beta)}\right).$$

Given the concave payoff function of the bureaucrat, it can be easily shown that (13) has an interior solution at $t_{BI}(\theta)$ whenever the constraints are not binding. The result is described in the following claim, the proof of which parallels that of Lemma 3 and is therefore skipped.

Claim. Assume that government delegates the decision-making authority with the restriction that $t \in [t_1, t_2]$. The bureaucrat's preferred regulation contract for type $\theta \in \{\underline{\theta}, \overline{\theta}\}$ is given by

[9] Having constraints on transfers and pollution at the same time would not be possible to discuss in any interesting way without allowing for more than two firm types.

$\alpha^R_{BI}(\theta, d_1, d_2) = \left(t^R_{BI}(\theta, d_1, d_2), d^R_{BI}(\theta, d_1, d_2)\right)$, where

$$
\begin{aligned}
t^R_{BI}(\theta, t_1, t_2) &= \begin{cases} t_1, & \text{if } t_1 \geq t_{BI}(\theta); \\ t_{BI}(\theta) = \theta\left(K - \frac{\theta}{1-\lambda} + \frac{\lambda\beta\theta}{(1-\lambda)(1-\beta)}\right), & \text{if } t_1 < t_{BI}(\theta) < t_2; \\ t_2, & \text{if } t_{BI}(\theta) \geq t_2. \end{cases} \\
d^R_{BI}(\theta, t_1, t_2) &= K - \frac{t^R_{BI}(\theta, d_1, d_2)}{\theta}.
\end{aligned}
$$

Next, consider government's preferred bounds on transfers. Because of her vested interest in transfers, B prefers a higher level of transfer than what government does. Government therefore sets a cap on the transfer level, and this upper bound on transfer is typically binding for at least one type of firm. Similarly to the case of partial delegation with restrictions on the pollution level, three regimes can arise in equilibrium $-$ weak delegation, strict delegation, and no delegation. However, there is a crucial difference in outcomes between the two cases of restrictions on pollution and restrictions on transfer. The difference is driven by the fact that, unlike $d_{CI}(\theta)$, the transfer $t_{CI}(\theta)$ is not monotone in θ. From (5), we have:

$$
t_{CI}(\underline{\theta}) < t_{CI}(\overline{\theta}) \quad \Leftrightarrow \tfrac{\underline{\theta}+\overline{\theta}}{1-\lambda} < K \quad \Leftrightarrow \lambda < 1 - \frac{\underline{\theta}+\overline{\theta}}{K}, \tag{14}
$$

which is likely to hold for low λ and high K. Therefore, with weak delegation, the upper bound on transfer is set either at $t_{CI}(\overline{\theta})$ or at $t_{CI}(\underline{\theta})$, depending on which is higher. Consequently, the full-information regulation contract is offered either to a high-cost firm or to a low-cost firm. Below, we summarize our main observations.

- If $\lambda < 1 - \frac{\underline{\theta}+\overline{\theta}}{K}$, then, in the weak-delegation regime, government chooses the bounds $[t_1, t_2] = \left[t_{CI}(\underline{\theta}), t_{CI}(\overline{\theta})\right]$. In response, the bureaucrat offers the following menu of contracts to the firm:

$$
\begin{aligned}
\alpha(\underline{\theta}) &= \left(\min\left\{t_{BI}(\underline{\theta}), t_{CI}(\overline{\theta})\right\}, K - \frac{\min\left\{t_{BI}(\underline{\theta}), t_{CI}(\overline{\theta})\right\}}{\underline{\theta}}\right), \\
\alpha(\overline{\theta}) &= \left(t_{CI}(\overline{\theta}), d_{CI}(\overline{\theta})\right).
\end{aligned}
$$

  Government thus implements the full-information regulation contract if the firm is high-cost. There is, on the other hand, a distortion at the contract offered to a low-cost firm, as $K - \frac{\min\left\{t_{BI}(\underline{\theta}), t_{CI}(\overline{\theta})\right\}}{\underline{\theta}}$ is less than $d_{CI}(\underline{\theta})$.

- If $\lambda \geq 1 - \frac{\underline{\theta}+\overline{\theta}}{K}$, however, then government chooses the bounds $[t_1, t_2] = \left[t_{CI}(\overline{\theta}), t_{CI}(\underline{\theta})\right]$ in the weak-delegation regime. In response, the bureaucrat offers the following menu of contracts:

$$\begin{aligned}\alpha(\underline{\theta}) &= (t_{CI}(\underline{\theta}), d_{CI}(\underline{\theta})),\\ \alpha(\overline{\theta}) &= \left(\min\left\{t_{BI}(\overline{\theta}), t_{CI}(\underline{\theta})\right\}, K - \frac{\min\left\{t_{BI}(\overline{\theta}), t_{CI}(\underline{\theta})\right\}}{\overline{\theta}}\right).\end{aligned}$$

Government thus implements the full-information regulation contract if the firm is low-cost. There is a distortion at the contract offered to a high-cost firm, as $K - \frac{\min\left\{t_{BI}(\overline{\theta}), t_{CI}(\underline{\theta})\right\}}{\overline{\theta}}$ is less than $d_{CI}(\overline{\theta})$.

- In the strict-delegation regime, government chooses the bounds $[t_1, t_2] = [t_{unif}, t_{unif}]$ where $t_{unif} = \left(\frac{\nu K}{\underline{\theta}} + \frac{(1-\nu)K}{\overline{\theta}} - \frac{1}{1-\lambda}\right) / \left(\frac{\nu}{\underline{\theta}^2} + \frac{(1-\nu)}{\overline{\theta}^2}\right)$.[10] In response, the bureaucrat offers the following menu of contracts:

$$\begin{aligned}\alpha(\underline{\theta}) &= \left(t_{unif}, K - \frac{t_{unif}}{\underline{\theta}}\right),\\ \alpha(\overline{\theta}) &= \left(t_{unif}, K - \frac{t_{unif}}{\overline{\theta}}\right).\end{aligned}$$

Government thus implements a uniform transfer for both types of firms.

- Note that t_{unif} always lies in between $t_{CI}(\underline{\theta})$ and $t_{CI}(\overline{\theta})$. In the knife-edge case of $\lambda = 1 - \frac{\underline{\theta}+\overline{\theta}}{K}$, we have $t_{unif} = t_{CI}(\underline{\theta}) = t_{CI}(\overline{\theta})$. In this special case, through strict delegation, government can implement first-best contracts for both types of firms. As its payoff changes continuously with respect to the parameters, strict delegation is preferred to other alternatives if λ is sufficiently close to $1 - \frac{\underline{\theta}+\overline{\theta}}{K}$, or equivalently, if $t_{CI}(\underline{\theta})$ is sufficiently close to $t_{CI}(\overline{\theta})$. This result is in sharp contrast to our results on partial delegation with restrictions on pollution. To see this, recall the effect of the uncertainty parameter ν on the delegation rule with constraints on pollution: If the uncertainty reduces such that the firm is more (less) likely to be low-cost, then government prefers weak delegation more (less) than strict delegation. However, in the present case of delegation with restrictions on transfers, we find that government prefers strict delegation to weak delegation for any uncertainty level if $t_{CI}(\underline{\theta})$ is sufficiently close to $t_{CI}(\overline{\theta})$.

We illustrate the above observations with a numerical example.

Example 4. We continue with the same parametric specification ($G = 50$, $K = 10$, $\overline{\theta} = 4$, and $\underline{\theta} = 2$) as used in Example 1. The feasible range of λ, satisfying Assumption 2, is $[0, 0.6]$. Consider the game in which C decides on delegation with restrictions on transfers, rather than

[10] t_{unif} is the government's preferred uniform level of transfer under complete information. Specifically, it solves the following optimization problem: $\max_t \nu\left[G - \frac{1}{2}\underline{d}^2 - \frac{t}{1-\lambda}\right] + (1-\nu)\left[G - \frac{1}{2}\overline{d}^2 - \frac{t}{1-\lambda}\right]$, subject to $t = \underline{\theta}(K - \underline{d}) = \overline{\theta}(K - \overline{d})$.

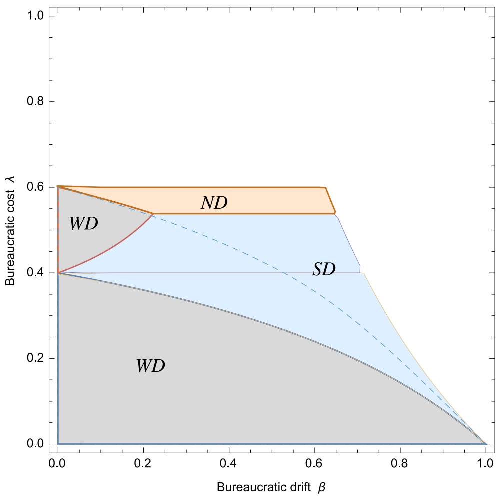

Figure 7: Optimal delegation with restricted transfer in (β, λ) space

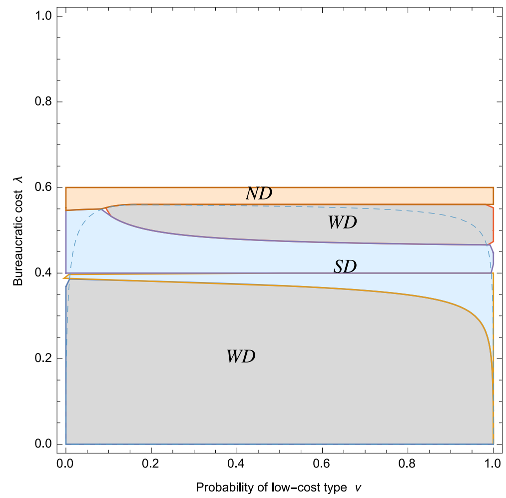

Figure 8: Optimal delegation with restricted transfer in (ν, λ) space

on pollution. Figure 7 plots the equilibrium regimes in (β, λ) space, with $\nu = 0.5$, $0 \leq \lambda \leq 0.6$ and $0 \leq \beta \leq \frac{1}{1+\lambda}$. In this example, (14) holds if and only if $\lambda < 0.4$. Figure 8 plots the equilibrium regimes in (ν, λ) space, with $\beta = 0.15$, $0 \leq \lambda \leq 0.6$, and $0 \leq \nu \leq 1$.[11] Observe that, when λ is close to 0.4, which is the level of λ at which $t_{CI}\left(\underline{\theta}\right) = t_{CI}\left(\overline{\theta}\right)$, government prefers strict delegation to other alternatives. The dot-dashed curves show the parameter values at which government is indifferent between full delegation (without any restriction on transfer) and no delegation in the full-delegation game. Like before, allowing for delegation being partial, in this case with restrictions on transfers, increases the scope for delegation.

6.2 The direction of transfer

In the regulation setting discussed above, the transfer goes from government to the firm to compensate the firm for its production costs. Government can, however, interact with the industry in more than one way. We consider an alternate framework in which the direction of transfer is opposite to what we find in the regulation setting. We call this setting *permits.* The two settings differ in how bureaucratic cost influences the equilibrium outcome.

In the permits framework, the firm sells its product in the market. Bureaucratic cost does not affect the market transaction, and therefore does not entail any loss of public funds. Production has, however, a social externality, and government regulates production by issuing pollution-contingent permits. The permit fee is a transfer that goes from the firm to government and is affected by bureaucratic cost.

To illustrate the difference between the two frameworks with an example, consider the case of building a road. In the regulation setting discussed above, the government procures the

[11]Note that the value of β here differs from that used in the previous examples; this is done to get as clear a picture as possible of this case of restrictions on transfers.

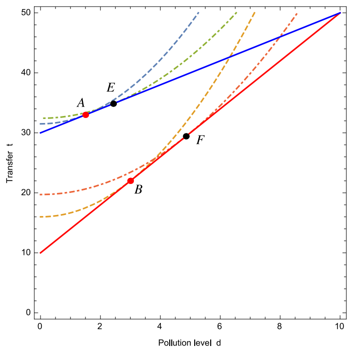

Figure 9: Permit contracts in (d, t) space.

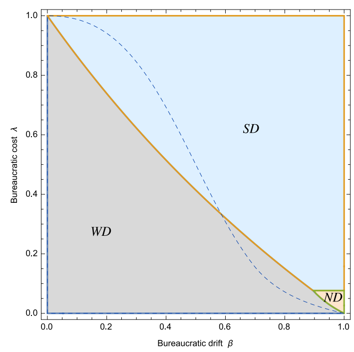

Figure 10: Permits setting: optimal partial delegation in (β, λ) space.

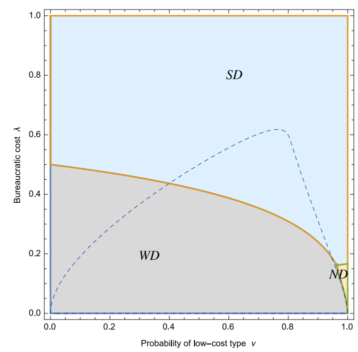

Figure 11: Permits setting: optimal partial delegation in (ν, λ) space.

construction of the road from the firm and compensates it for its production cost. The firm is regulated as its production causes pollution. In the permits setting, the firm produces on its own and charges a toll directly on users of the road. Government regulates by charging a fee for permitting the firm to use a polluting production technology to build the road. It follows that, in both cases, government's contracts consist of combinations (t, d) of transfers and pollution levels, the difference being the direction of transfers. It can be shown that the two settings produce the same outcome when there is no bureaucratic cost.

We discuss here our findings on the permits case; details of the analysis are in Appendix B. Figure 9 illustrates the construction of contracts in the permits setting; this Figure can be compared to Figure 3 for the regulation setting discussed above.[12] Government still has preferences for little pollution, but would want transfers to be high, since the government is now in the receiving end of the transfers. For the bureaucrat, it is, as before, important that there is a flow of transfer through the regulatory system that she controls. The more transfers that are involved, and the more easily such transfers can be diverted, the more can a bureaucrat get out of her position. In the permits setting, therefore, the bureaucrat would like a higher pollution level than what is optimal for the government, since this allows for more permits sold and thereby higher transfers. As in Figure 3, the blue and red straight lines in Figure 9 are the participation constraints for the low-cost and high-cost firm, respectively. The A and B points are the contracts offered by government were it informed, whereas the E and F points are the contracts offered by a bureaucrat; for the sake of clarity, the contract menu offered by an uninformed government is not represented in this Figure.

The bureaucrat now being upward biased in terms of pollution has implications for government's optimal partial delegation rules. The government sets an upper bound on the pollution level when it partially delegates in the permits setting. Consequently, the equilibrium regulation contracts differ between the two settings. Specifically, under weak delegation, government implements the full-information regulation contract only for one type of firms −while that is the

[12] In fact, the parametric specification is the same for the two Figures: $G = 50$, $K = 10$, $\overline{\theta} = 4$, $\underline{\theta} = 2$, $\lambda = 0.25$, $\beta = 0.65$, and $\nu = 0.5$.

low-cost type before, it is the high-cost type in the current permits setting. The implication is that the *uncertainty principle* still needs to be modified in the permits setting. But now strict delegation is more prevalent for high values of ν, i.e., when the probability of the firm being low-cost is high.

Another feature of the permits case that distinguishes it is that pollution now is decreasing in the bureaucratic cost, measured by λ, while it is increasing in the regulation case. The reason is that, in the regulation setting, an increase in λ makes it more expensive for government to pay for a decrease in pollution, so that d increases; while in the permits setting, an increase in λ makes it more expensive for the firm to pay for the right to pollute, so that d decreases.[13] The consequence is that, in the case of permits, strict delegation dominates no delegation when λ is high, which is the opposite of the situation in the regulation case.

While the precise statements of our findings on the permits setting are found in Appendix B, Figures 10 and 11 illustrate the equilibrium outcome in this case. As the Figures illustrate, no delegation is the outcome only when β and ν are high and λ is low. Note, first, that Assumptions 1 and 2 do not apply to the permits setting. In this setting, the worry in terms of Assumption 1 would be that the bureaucrat would want to set a negative transfer, which will not happen for any β, λ, $\nu \in (0, 1)$.[14] And in terms of Assumption 2, an informed government would never go for a no-regulation permit contract for any β, λ, $\nu \in (0, 1)$.[15] Secondly, note that the parametric specification here differs from that of Example 1, which is used elsewhere in this paper, in that we have lowered K from 10 to 5, so that our parameters are set at $G = 50$, $K = 5$, $\overline{\theta} = 4$, and $\underline{\theta} = 2$. In Figure 10, we put $\nu = 0.5$ and show how delegation varies in (β, λ) space, while in Figure 11, we put $\beta = 0.5$ and show how delegation varies in (ν, λ) space. The change in parameters is done in order to have any prevalence of no delegation at all; as seen from part (i) of Proposition B.2 in Appendix B, a necessary condition for No Delegation to occur is that $\overline{\theta}$ is closer to K than to $\underline{\theta}$, i.e., that $\overline{\theta} \geq \frac{K+\underline{\theta}}{2}$.[16] The dotted curves in the two Figures show the prevalence of delegation when only full delegation is feasible, with full delegation being government's choice below the curves.

In many situations, the setting, either regulation or permits, is exogenously given. In other situations, government may influence the way it deals with industry, e.g., by being able to choose between a regulation setting and a permits setting. Differences in the way delegation works in the two settings affect government's preference over them. When government regulates industry without delegation, it is constrained by an information problem. Therefore, in that case, government's preference is influenced by how the information rent is shared with the low-cost firm in the two settings. On the other hand, when government partially delegates to informed bureaucrats to interact with the industry, it is constrained by the distortion that the

[13] Compare, e.g., equation (5) with equation (B5) in Appendix B.
[14] This can be seen from problem (B8) in Appendix B.
[15] This can be seen from equation (B5) in Appendix B.
[16] Put $\nu = 1$ in the condition, in there, for $\lambda^{NDT} = 0$ to get this necessary condition.

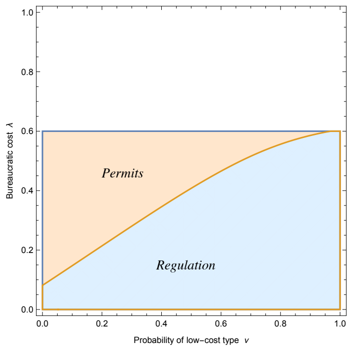

Figure 12: Regulation vs Permits in the full-information case

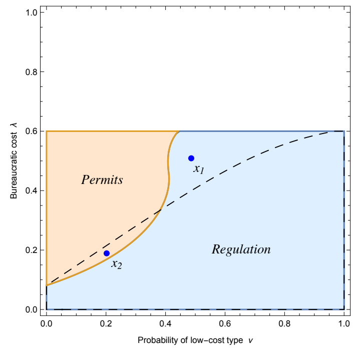

Figure 13: Regulation vs Permits in the no-delegation case

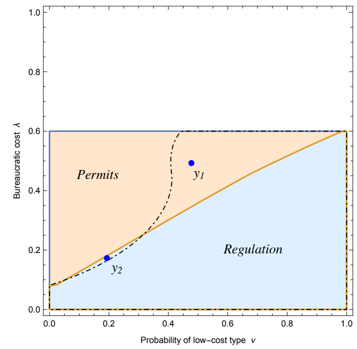

Figure 14: Regulation vs Permits with optimal partial delegation

bureaucratic drift creates.

While a full analysis is complicated, we can show, through an example, that information asymmetry and partial delegation may lead to a reversal of preference over the two regulatory environments. We return to the parametric specification ($G = 50$, $K = 10$, $\overline{\theta} = 4$, and $\underline{\theta} = 2$) of Example 1. The feasible range of λ, satisfying Assumption 2, is $[0, 0.6]$; while this Assumption does not apply to the permits setting, we need to invoke it for the purpose of this comparison. Figure 12 plots government's preference over the two frameworks in (λ, ν) space in the full-information case. In Figure 13, we compare government's payoff in the two frameworks in the asymmetric-information case with no delegation (i.e., when an uninformed government decides the regulation policy). The dashed curve presents the parameter values at which government is indifferent between regulation and permits in the full-information case. Observe, by comparing the two Figures, that lack of information can lead to preference reversal. Government's preference is reversed at points x_1 and x_2, depending on whether or not government has information about the firm's production technology. We next put $\beta = 0.4$ and compare government's value in the two frameworks in the case when it is uninformed but has an option to partially delegate to an informed bureaucrat. Figure 14 shows government's preference over the two frameworks in the asymmetric-information case with optimal partial delegation. The dot-dashed curve represents the parameter values of (λ, ν) at which government is indifferent when it regulates the firm with no delegation. The Figure also shows that the possibility of partial delegation can lead to preference reversal in the asymmetric-information case − government's preference will be reversed at points y_1 and y_2, depending on whether or not it is allowed to delegate.

6.3 Bureaucratic motivation

The analysis of the partial-delegation problem is related to the direction of policy bias rather than to a specific form of the bureaucrat's objective function. In the basic model (referred to as BM hereafter in this subsection), the bureaucrat has a private interest in transfers, which results in a downward policy bias. Below we consider an alternative form of bureaucratic objective − a case of an environment-friendly bureaucrat. We show that it leads to a similar downward policy bias and generates qualitatively similar result in terms of partial delegation rule.

We model an environment-friendly bureaucrat's objective function as a weighted average of the pollution cost and government's payoff.

$$\begin{aligned} U_B(\theta, t, d) &= \delta\left(-\frac{1}{2}d^2\right) + (1-\delta)\, U_C(t,d) \\ &= \delta\left(-\frac{1}{2}d^2\right) + (1-\delta)\left[\left(G - \frac{1}{2}d^2\right) - \frac{t}{1-\lambda}\right] \\ &= (1-\delta)\left[\left(G - \frac{d^2}{2(1-\delta)}\right) - \frac{t}{1-\lambda}\right], \end{aligned} \tag{15}$$

where $\delta \in (0,1)$ is the extent of the bureaucrat's aversion to pollution and is comparable to the parameter β considered in BM. The bureaucrat has the same cost of pollution as government at $\delta = 0$ and its cost increases with δ. Since the bureaucrat's payoff is decreasing in t, the firm's participation constraint will be binding. We can, therefore, replace t by $\theta(K-d)$ and solve for the bureaucrat's choice of contract for type θ, which is given by

$$\begin{aligned} d_{BI}(\theta) &= (1-\delta)\frac{\theta}{1-\lambda}, \\ t_{BI}(\theta) &= \theta(K - d_{BI}(\theta)). \end{aligned}$$

The bureaucratic aversion to pollution results in a downward policy bias. Compared to BM, the policy bias will be higher if and only if δ is higher than $\frac{\lambda\beta}{1-\lambda}$. The analysis of BM can, however, give us clear insight into the partial delegation problem in this case. The pollution-aversion parameter δ affects government's value only in case of weak delegation. Therefore, compared to BM, weak delegation will be more (less) prevalent for δ higher (lower) than $\frac{\lambda\beta}{1-\lambda}$, whereas the trade-off between strict delegation and no delegation will remain unchanged.

7 Related literature

Our paper relates to several strands of literature. Firstly, there is a large literature on when and how to delegate decision power to bureaucrats; see Bendor et al. [4], Huber and Shipan [24], Gailmard and Patty [18], and Moe [36] for summaries of this literature. This literature points to several factors responsible for delegating decision power to bureaucrats.[17] One is the effect of political uncertainty, which has (at least) two facets. First, politicians currently in position who are worried that they will lose next election may want to delegate decision power to bureaucrats. Secondly, government faces a credibility problem in that industry may hold back investments when there is uncertainty as to whether current regulatory policy will be continued in the

[17] The list can easily be made longer. Gailmard and Patty [17], for example, argue that the extent of discretion given to bureaucrats affects their motivation to acquire information. In our analysis, we take the view that bureaucrats are fully informed.

future; see, e.g., Spulber and Besanko [38], Levine, et al. [32], and Evans, et al. [12] on how delegating the task of regulation to bureaucrats may alleviate this hold-up problem. As argued by de Figueiredo [8], however, such political uncertainty calls for giving the bureaucrat limited discretion, or what we here call strict delegation, in order to make sure that current government's policy is carried out also in the future. While Gilardi [20] argues for the increased importance of the credibility problem to explain the expansion of independent regulatory agencies in many Western European countries in recent decades, it is not clear how much discretion for bureaucrats that this argument can explain. In the current analysis, we disregard any future elections and thus sidestep political uncertainty altogether.

The second is the misalignment of interests between government and bureaucrat, what we here call bureaucratic drift. The so-called *ally principle* states that there is more delegation, the more aligned the two are, or the lower the bureaucratic drift is (see, e.g., Epstein and O'Halloran [10], Bendor et al. [4], and Huber and Shipan [24]). In our analysis, we model the bureaucratic drift in a way particularly suited to the regulatory setting we discuss: the bureaucrat puts weights in her objectives on both the consumer surplus and the transfers diverted into the bureaucracy. Still, the ally principle shows up clearly also in our context: the more weight the bureaucrat puts on the diverted transfers, the less delegation there will be in equilibrium.

The literature also points to the importance of government's policy uncertainty (e.g., Bendor et al. [4], and Huber and Shipan [24]): the more uncertain government is about the effect of the decisions to be made, the more willing it is to delegate those decisions to an informed bureaucrat; this is oftentimes called the *uncertainty principle*. In our setting, government is incompletely informed about the regulated firm's production technology. The uncertainty is the highest when, in our two-type case, it is equally likely that the firm has low and high costs, respectively. What we find is slightly in contrast to the uncertainty principle. The bureaucratic drift gives rise to a downward bias in pollution levels: the bureaucrat prefers lower pollution level than government does. Weak delegation means putting a cap on the bias. When this is not enough, the government may want to resort to strict delegation. However, such strict delegation is based on an ex-ante expectation of firm's production costs and does not work well when the firm is likely to have low costs. The upshot is that, in discussions of delegation of regulatory tasks, one cannot expect the uncertainty principle to hold unmodified.

Other influencing factors include various institutional aspects of government and bureaucracy. For example, Huber and Shipan [24] point to government's ability to monitor the bureaucrat *ex post*. When the scope for such *ex-post* monitoring is large, government is more interested in delegating *ex ante*. This can clearly be relevant in a regulatory setting. For example, government may be able to overcome some of the costs of delegation by instituting a way for regulated firms to appeal to the ministry. Also, it may in some cases be possible to write contracts with the bureaucrat in order to incentivize her to regulate more in line with government's interest. We have still chosen not to include such aspects of delegated regulation in the present analysis, and there is no form of *ex-post* monitoring in our model. Thus, when we find that government imposes limits on bureaucrats' monitoring, this does not show up as a substitute for *ex-post* monitoring.[18]

Gailmard [16] analyzes the relative merits of a menu law and a partial delegation rule (referred to as an action restriction) as an instrument to control the choices of an informed bureaucrat. A menu law allows the legislature to regulate a bureaucrat with a state-contingent menu of contracts, specifying combinations of policy choice and transfer payments for each state. Gailmard finds that an action restriction can outperform the best available menu law when the legislature cannot necessarily commit to a state-contingent menu law. In our study, the legislature's objective is to regulate a firm with private information. Unlike Gailmard, we study the relative merit of partial delegation to an informed bureaucrat against an alternative scenario, in which the legislature can regulate the firm with an incentive-compatible menu-based regulation contract. Thus, the government's decision to delegate with limits on bureaucrats' actions arises from the relative inefficiency in regulating firms with menu-based contracts.

There is a small but growing literature addressing the strategic reasons behind delegation – why and when the legislature may benefit from delegating to a biased bureaucrat or limiting discretion of an unbiased bureaucrat, defying the ally principle (see, e.g., Boehmke, et al. [5], Gailmard and Patty [17], McCarty [35], and Montagnes and Wolton [37]). Some of these studies look at the strategic impact of delegation on the actions of the regulated agencies, such as firms. Delegation affects the regulated agent's behavior and her interaction with the bureaucracy in various ways, including transmission of useful policy-relevant information (Boehmke, et al. [5]), effort in self-regulation (McCarty [35]), and productive investment (Montagnes and Wolton [37]). This gives reasons for the legislature's strategic delegation and granting of discretionary authority. Gailmard and Patty [17, 19] study the impact of delegation on the bureaucrat's incentive to acquire expertise. We disregard such strategic aspects of delegation in our model by considering only informed bureaucrats and limiting the action of the regulated firm.

Secondly, there is a literature discussing regulation of firms by a bureaucrat where the focus is not on whether or not to delegate but on how to avoid regulatory capture; see, e.g., Laffont and Tirole [29, 30]. In these models, regulation is modeled as a three-tiered principal-agent problem with the bureaucrat in the middle tier, observing the firm's true type with a certain probability. Regulatory capture is modeled as collusion between the bureaucrat and the firm and the focus is on how to formulate contracts with the bureaucrat and the firm that are collusion-proof, thus avoiding regulatory capture. While we certainly believe that regulatory capture is a problem

[18] Other related studies examine the merit of delegation in accommodating policy drift in unified versus separated governments (Volden [40]), when legislative decision-making features supermajoritarianism (Callander and Krehbiel [7]), when bureaucracy may have limited capacity in implementing policy (Huber and McCarty [23]), and when the legislature is allowed to carry out *ex-post* sanctions towards delegatees within the executive branch (Lowande [33]). Unlike these studies, we focus away from the institutional features of the government and bureaucracy and restrict our attention in examining the regulatory implications of delegating the task of regulation to bureaucracy.

that should be taken seriously, we distract from it here in order to focus on government's use of various forms of partial delegation in order to make delegation less harmful and therefore more useful. In this literature on regulatory capture, delegation is taken as a given, and there is little discussion of how one can limit bureaucrats' discretion in order to avoid regulatory capture.[19] Moreover, this literature assumes that incentive contracts between government and bureaucrats are feasible, whereas we let bureaucrats be hired at an unmodeled fixed salary.

Thirdly, there exist models of bureaucrats regulating firms. One example is Khalil, et al. [25]. They model a bureaucrat who procures a good from a privately informed firm and who is given a fixed budget. The bureaucrat benefits in part from funds kept in the bureaucracy and not payed out to the firm. Although this is not a model of delegation and the bureaucrat is not informed, as ours is, there are some similarities in result. In their model, the government will keep the bureaucrat's budget low, to which the bureaucrat may choose to respond by offering the firm a pooling contract. This resembles our strict delegation, where the bureaucrat is tied up and, while not offering a pooling contract, at least is restricted to offer both firm types the same level of pollution. The work of Hiriart and Martimort [21] is quite complementary to ours. Like us, they discuss how much discretion government should give when delegating a regulatory task to a bureaucracy. But that regulatory task is quite different from ours, since the incentive problem involved is one of moral hazard rather than of asymmetric information. Moreover, the bureaucrat's bias in their model is pro-firm rather than pro-transfers. Also, they do not discuss whether or not government should delegate at all, as we do here.[20]

Fourthly, we contribute to the literature on the political economy of environmental policy. Our starting point is a model by Boyer and Laffont [6]. In their analysis, there are no bureaucrats. Instead there are two political parties, one that favors the regulated firm and one that favors others. They discuss whether politicians should be restricted to a non-discriminatory regulation policy, which is essentially what we here call strict delegation, with both firm types being offered the same pollution level. But although the outcome they discuss has similarities with ours, the issues involved are different. In particular, in our analysis, it is the politicians who formulate the delegation policy and decide whether to have strict or weak delegation, and the bureaucrat, when delegated the decision power, has full information about the firm technology. In the model of Strausz [39], there are bureaucrats but, in contrast to our setting, they are not informed and therefore resort to regulation through menus of contracts. The focus of Strausz is on how pre-election bargaining among the political parties affects the performance of the bureaucrats.

Finally, we are related to the literature on the delegation problem, which started out with the seminal work of Holmström [22]. There, a relationship between a principal and an agent is modeled, where incentive contracts are not feasible, and the agent is biased and privately

[19] One exception is the work of Laffont and Martimort [28], who discuss how government can institute multiple regulators of the same firm, in order to reduce bureaucrats' discretion and this way make regulatory capture more difficult.

[20] An interesting study of regulatory discretion is by Duflo, et al. [9], who show how giving discretion to regulators inspecting polluting plants in Gujarat, India, improves the regulatory outcome.

informed; all these features are shared with our model, where the two are called government and bureaucrat. But in contrast to this literature, the task in our setting is not to pick an action from the real line but to pick a contract in order to regulate the actual agent: the firm. Recent work in this literature includes Alonso and Matouschek [1] and Amador and Bagwell [2]. In Koessler and Martimort [26], Frankel [13–15], and Letina, et al. [31], the action space is multidimensional, just as we have in our model, with our two-dimensional regulatory contracts. Consider, in particular, Frankel [14]. One difference from our model is that Frankel assumes a state-independent bias, meaning the principal knows the bias even though he does not know the state, whereas in our regulatory set-up, the strength and direction of the bureaucratic drift vary with firm type.[21] In Melumad and Shibano [34], the action space is one-dimensional, but the agent's bias is state-dependent, in such a way that both its strength and direction are private information. Alonso and Matouschek [1] extend the analysis of Melumad and Shibano to more general distributions of the state variable and more general utility functions. Melumad and Shibano [34] also distinguish between decision rules that are what they call communication dependent and those that are communication independent, a distinction that closely resembles the one we make here between weak and strict delegation.

A common theme between these papers and ours is the need to cap the bias. In our case, this means putting a lower bound on pollution levels because the bureaucrat is less interested in pollution (and more interested in transfers) than is government. However, the focus is still quite different. The papers cited are interested in finding out whether interval delegation is optimal, meaning that the set of actions that the principal optimally admits is an interval (or the equivalent in the multi-dimensional version). Our model differs from those discussed in this literature in that, without delegation, there is a regulation problem with asymmetric information. Our primary concern is to discuss how delegation of regulation to an informed but biased bureaucrat is best done. In order to do this in a transparent way, we introduce a model with two types of firms, so that government's first-best choice is one of two single points in the action (transfer-pollution) space. Moreover, we limit government to put constraints in one dimension only, which is pollution in the main treatment, and we impose on government a requirement that the constraint is an interval. Of particular interest, relative to the literature on the delegation problem, is our finding of a scope for strict delegation: it may be optimal to delegate stricter than merely capping the bias: the government's lack of information about the regulated firm, a feature which is novel to our delegation problem, may cause it to limit the bureaucrat to perform a uniform regulatory policy. Finally, in our analysis, the principal has a natural outside-option: to go about regulating the firm himself by way of a menu of self-selection contracts. This way, we are able not only to discuss what is the best way to delegate, but also whether this optimum delegation is better than no delegation at all.

[21] In an extension, Frankel [14] allows for biases that vary with both the state and the action. However, unlike our model, the state and the action affect biases in an additively separable way.

8 Concluding remarks

In this paper, we develop a simple model to study a government's incentives to delegate to a bureaucrat the regulation of an industry. While the bureaucrat has more industry-specific knowledge, her interest may not align completely with that of government. Delegation thus creates an agency problem. Our analysis shows how partial delegation, i.e., delegation followed by laws and rules to restrict bureaucratic discretion, improves government's benefit from delegation.

The key result of the paper is the characterization of the optimal partial-delegation rule. In particular, we point to the occurrence of strict delegation, in which government caps the bureaucrat's actions based on expected costs. Such strict delegation has interesting implications. It gives rise to a regulatory contract featuring distortion at the top; and it leads to an interesting modification of the uncertainty principle. We describe how various factors, including bureaucratic drift, bureaucratic cost, and government's lack of information, affect the delegation rule and, subsequently, the equilibrium regulation policy. We show that, while bureaucratic discretion typically reduces with bureaucratic cost and bureaucratic drift, the equilibrium regulation policy changes in a non-trivial way. This is because there are substantive differences in the type of regulation policy implemented in various delegation regimes. In a related sense, uncertainty plays a critical role in determining the optimal rule of delegation. Bureaucratic discretion is low in situations with little uncertainty. However, the form of delegation rule depends on the exact nature of uncertainty and how the government interacts with the industry. For example, when government regulates pollution, it prefers no delegation over strict delegation only if the firm is more likely to be low-cost; while this feature of the delegation policy is more or less reversed when government sells pollution permits.

While our analyses provide some normative suggestions for the design of delegation rules, many interesting questions remain unanswered. First, the firm plays no strategic role in our model. In many contexts, it is more realistic to assume that the producer may influence the regulation policy by colluding with the bureaucrat. Allowing a richer action space to the producer can provide new insights for designing effective delegation policies. Secondly, the delegation framework assumes no contractual relationship between the principal and the delegates. While the assumption properly reflects the relationship between a politician and a bureaucrat, there are other situations where the assumption may not be appropriate. Thirdly, we do not address the bureaucrat's incentive for acquiring information. Again, this assumption seems appropriate in situations in which bureaucrats can possibly be hired based on their industry-specific knowledge. In other situations, we might expect that the delegation rule could have a direct effect on her incentive to acquire information. For example, low bureaucratic discretion can demotivate a bureaucrat from a detailed investigation of the firm. In such a situation, a planner must take the issue of information acquisition into consideration for designing the optimal delegation rule. Finally, our search for the optimal delegation rule has been limited in that we have imposed on the partial delegation that it be an interval and on top of that have allowed partial delegation in only one dimension, pollution level or transfer. A richer discussion of whether the optimal delegation rule is an interval or, rather, a closed set in our setting of two-dimensional regulation contracts, would require a continuous type space for firms' private information on technology. We leave all these questions for future research.

Appendix A

The Appendix contains the proofs omitted in the text.

Proof of Lemma 1

Proof. Government's expected payoff is weakly decreasing in transfers, and the firm's payoff is increasing in transfers. Government would therefore prefer to reduce transfer as much as possible, subject to the firm's participation constraint. It can be shown that (IRH) and (ICL) will be binding at the optimum. The low-cost firm can always pretend to be the high-cost firm and gets a payoff of $\bar{t} - \underline{\theta}\left(K - \bar{d}\right)$. In order to make it choose $(\underline{t}, \underline{d})$, government therefore shares an information rent of $IR\left(\bar{d}\right) := \triangle\theta\left(K - \bar{d}\right)$. Thus, $\bar{t} = \bar{\theta}\left(K - \bar{d}\right)$, and $\underline{t} = \underline{\theta}\left(K - \underline{d}\right) + IR\left(\bar{d}\right)$. Replacing $\bar{t}$ and $\underline{t}$ in (6) and using the fact that (IRH) and (ICL) together imply (IRL), we can rewrite the optimization problem as:

$$\begin{aligned}\max_{\underline{\alpha},\bar{\alpha}} \ & \nu\left[V\left(\underline{\theta}, \underline{d}\right) - \frac{\lambda}{1-\lambda}\underline{\theta}\left(K - \underline{d}\right)\right] \\ & + (1-\nu)\left[V\left(\bar{\theta}, \bar{d}\right) - \frac{\lambda}{1-\lambda}\bar{\theta}\left(K - \bar{d}\right)\right] - \frac{\nu}{1-\lambda} IR\left(\bar{d}\right), \\ \text{subject to } & (ICH).\end{aligned} \tag{A1}$$

From the first-order conditions of the unconstrained problem, we see that the pollution levels are given by $d_{CN}\left(\underline{\theta}\right) = \frac{\underline{\theta}}{1-\lambda}$ and $d_{CN}\left(\bar{\theta}\right) = \min\left\{\left(\frac{1}{1-\lambda}\right)\left(\bar{\theta} + \frac{\nu}{1-\nu}\Delta\theta\right), K\right\}$.[22] The transfers given in the Lemma then follow. □

Proof of Proposition 1

Proof. Define $f(\theta, d) := U_C\left(\theta\left(K - d\right), d\right)$. Then,

$$f(\theta, d) = G - \frac{\theta K}{1-\lambda} + \left(\frac{\theta d}{1-\lambda} - \frac{1}{2}d^2\right). \tag{A2}$$

Observe that

$$f(\theta, d_1) - f(\theta, d_2) = (d_1 - d_2)\left(\frac{\theta}{1-\lambda} - \frac{d_1 + d_2}{2}\right). \tag{A3}$$

[22] Note that $d_{CN}\left(\bar{\theta}\right) \geq d_{CN}\left(\underline{\theta}\right)$, which ensures that ICH is satisfied at the unconstrained solution.

Government's expected payoff in the no-delegation regime is given by

$$
\begin{aligned}
& E_\theta U_C\left(\alpha_{CN}(\theta)\right) \\
= \ & \nu\left[G-\frac{1}{2} d_{CN}(\underline{\theta})^2-\frac{t_{CN}(\underline{\theta})}{1-\lambda}\right]+(1-\nu)\left[G-\frac{1}{2} d_{CN}(\overline{\theta})^2-\frac{t_{CN}(\overline{\theta})}{1-\lambda}\right] \\
= \ & \nu f\left(\underline{\theta}, d_{CN}(\underline{\theta})\right)+(1-\nu) f\left(\overline{\theta}, d_{CN}(\overline{\theta})\right)-\frac{\nu \triangle \theta\left(K-d_{CN}(\overline{\theta})\right)}{1-\lambda},
\end{aligned}
$$

where the second equality follows from using (A2) and replacing d and t with expressions derived in (7). Similarly, government's expected payoff in the full-delegation regime is given by

$$
\begin{aligned}
& E_\theta U_C\left(\alpha_{BI}(\theta)\right) \\
= \ & \nu\left[G-\frac{1}{2} d_{BI}(\underline{\theta})^2-\frac{t_{BI}(\underline{\theta})}{1-\lambda}\right]+(1-\nu)\left[G-\frac{1}{2} d_{BI}(\overline{\theta})^2-\frac{t_{BI}(\overline{\theta})}{1-\lambda}\right] \\
= \ & \nu f\left(\underline{\theta}, d_{BI}(\underline{\theta})\right)+(1-\nu) f\left(\overline{\theta}, d_{BI}(\overline{\theta})\right),
\end{aligned}
$$

where the second equality follows from again using (A2) and replacing d and t with expressions derived in (10). We can thus compute the payoff difference between the two regimes, ΔD defined in (11), as

$$
\begin{aligned}
\triangle D= & \ \nu\left[f\left(\underline{\theta}, d_{CN}(\underline{\theta})\right)-f\left(\underline{\theta}, d_{BI}(\underline{\theta})\right)\right] \\
& +(1-\nu)\left[f\left(\overline{\theta}, d_{CN}(\overline{\theta})\right)-f\left(\overline{\theta}, d_{BI}(\overline{\theta})\right)\right]-\frac{\nu \triangle \theta\left(K-d_{CN}(\overline{\theta})\right)}{1-\lambda} .
\end{aligned} \tag{A4}
$$

We first use (A3) to expand the first term in (A4), inserting for $d_{CN}(\underline{\theta})$ and $d_{BI}(\underline{\theta})$ from (7) and (10), respectively, to get

$$
\begin{aligned}
& \nu\left[f\left(\underline{\theta}, d_{CN}(\underline{\theta})\right)-f\left(\underline{\theta}, d_{BI}(\underline{\theta})\right)\right] \\
= \ & \nu\left(d_{CN}(\underline{\theta})-d_{BI}(\underline{\theta})\right)\left(\frac{\underline{\theta}}{1-\lambda}-\frac{d_{CN}(\underline{\theta})+d_{BI}(\underline{\theta})}{2}\right) \\
= \ & \frac{\nu\left(d_{CN}(\underline{\theta})-d_{BI}(\underline{\theta})\right)^2}{2} \\
= \ & \frac{\nu \lambda^2 \beta^2 \underline{\theta}^2}{2(1-\lambda)^2(1-\beta)^2} .
\end{aligned}
$$

Next, we similarly use (A3) to expand the second term in (A4), inserting for $d_{CN}(\overline{\theta})$ and $d_{BI}(\overline{\theta})$

from (7) and (10), respectively, to get

$$
\begin{aligned}
& (1-\nu)\left[f\left(\overline{\theta}, d_{CN}\left(\overline{\theta}\right)\right)-f\left(\overline{\theta}, d_{BI}\left(\overline{\theta}\right)\right)\right] \\
= \; & (1-\nu)\left[\left(d_{CN}\left(\overline{\theta}\right)-d_{BI}\left(\overline{\theta}\right)\right)\left(\frac{\overline{\theta}}{1-\lambda}-\frac{d_{CN}\left(\overline{\theta}\right)+d_{BI}\left(\overline{\theta}\right)}{2}\right)\right] \\
= \; & \frac{1-\nu}{2}\left[\left(d_{CN}\left(\overline{\theta}\right)-\frac{\overline{\theta}}{1-\lambda}+\frac{\lambda\beta\overline{\theta}}{(1-\lambda)(1-\beta)}\right)\left(\frac{\overline{\theta}}{1-\lambda}-d_{CN}\left(\overline{\theta}\right)+\frac{\lambda\beta\overline{\theta}}{(1-\lambda)(1-\beta)}\right)\right] \\
= \; & \frac{1-\nu}{2}\left[\left(\frac{\lambda\beta\overline{\theta}}{(1-\lambda)(1-\beta)}\right)^2-\left(d_{CN}\left(\overline{\theta}\right)-\frac{\overline{\theta}}{1-\lambda}\right)^2\right] \\
= \; & \begin{cases} \frac{1-\nu}{2(1-\lambda)^2}\left[\frac{\lambda^2\beta^2\overline{\theta}^2}{(1-\beta)^2}-\frac{\nu^2(\triangle\theta)^2}{(1-\nu)^2}\right], & \text{if } d_{CN}\left(\overline{\theta}\right)<K; \\ \\ \frac{1-\nu}{2(1-\lambda)^2}\left[\frac{\lambda^2\beta^2\overline{\theta}^2}{(1-\beta)^2}-\left((1-\lambda)K-\overline{\theta}\right)^2\right], & \text{if } d_{CN}\left(\overline{\theta}\right)=K. \end{cases}
\end{aligned}
$$

Finally, we use the expression for $d_{CN}\left(\overline{\theta}\right)$ in (7) to write the third term in (A4) as

$$
\frac{\nu\triangle\theta\left(K-d_{CN}\left(\overline{\theta}\right)\right)}{1-\lambda} = \begin{cases} \frac{\nu\triangle\theta}{(1-\lambda)^2}\left((1-\lambda)K-\overline{\theta}-\frac{\nu\triangle\theta}{(1-\nu)}\right), & \text{if } d_{CN}\left(\overline{\theta}\right)<K; \\ 0, & \text{if } d_{CN}\left(\overline{\theta}\right)=K. \end{cases}.
$$

Combining these expressions, we can write ΔD in (A4) as

$$
\triangle D = \begin{cases} \frac{\nu\lambda^2\beta^2\underline{\theta}^2}{2(1-\lambda)^2(1-\beta)^2}+\frac{(1-\nu)\lambda^2\beta^2\overline{\theta}^2}{2(1-\lambda)^2(1-\beta)^2} \\ +\frac{(1-\nu)}{2(1-\lambda)^2}\left(\frac{\nu\triangle\theta}{1-\nu}\right)^2-\frac{\nu\triangle\theta}{(1-\lambda)^2}\left((1-\lambda)K-\overline{\theta}\right), & \text{if } d_{CN}\left(\overline{\theta}\right)<K; \\ \\ \frac{\nu\lambda^2\beta^2\underline{\theta}^2}{2(1-\lambda)^2(1-\beta)^2}+\frac{(1-\nu)\lambda^2\beta^2\overline{\theta}^2}{2(1-\lambda)^2(1-\beta)^2} \\ -\frac{(1-\nu)}{2(1-\lambda)^2}\left((1-\lambda)K-\overline{\theta}\right)^2, & \text{if } d_{CN}\left(\overline{\theta}\right)=K. \end{cases}. \tag{A5}
$$

Note that $\frac{d(\Delta D)}{d\beta}>0$. When $\beta=0$, $\triangle D$ can take a positive value (for example, for low values of ν) or a negative value (for example, for high values of ν) depending on other parameter values. When $\beta=1$, $\triangle D>0$ always. However, Assumption 1 sets an implicit upper bound on β. At this upper bound, $\triangle D$ can be positive or negative depending on other parameter values. Since $\frac{d(\Delta D)}{d\beta}>0$, we can determine a threshold $\overline{\beta}$ such that $\triangle D>0$, and no delegation is preferred to full delegation, if and only if $\beta>\overline{\beta}$. Note that this threshold can be zero or one, such that one regime is preferred for all values of β, depending on other parameter values.

Recall the definition of ν^*, in equation (8), and that $\frac{d\nu^*}{d\lambda}<0$. We can equivalently define $\lambda^*:=1-\frac{\overline{\theta}+(\nu\triangle\theta/(1-\nu))}{K}$ as that λ which, for a given ν, makes $\nu^*=\nu$, so that we have $d_{CN}\left(\overline{\theta}\right)<K$ if $\lambda<\lambda^*$, and $d_{CN}\left(\overline{\theta}\right)=K$ if $\lambda^*\leq\lambda\leq 1-\frac{\overline{\theta}}{K}$, where the upper bound comes from Assumption

2. ΔD is not necessarily monotone in λ. However, the sign of $\triangle D$ is determined by the sign of an expression that is increasing in λ. For fixed values of β and ν, we can, from (A5), write

$$\triangle D = \frac{G(\lambda)}{(1-\lambda)^2},$$

where

$$G(\lambda) = \begin{cases} \lambda^2 \left\{ \frac{\nu\beta^2\underline{\theta}^2}{2(1-\beta)^2} + \frac{(1-\nu)\beta^2\overline{\theta}^2}{2(1-\beta)^2} \right\} + \frac{1-\nu}{2}\left(\frac{\nu\triangle\theta}{1-\nu}\right)^2 & \\ -\nu\triangle\theta\left((1-\lambda)K - \overline{\theta}\right), & \text{if } \lambda < \lambda^*; \\ & \\ \lambda^2 \left\{ \frac{\nu\beta^2\underline{\theta}^2}{2(1-\beta)^2} + \frac{(1-\nu)\beta^2\overline{\theta}^2}{2(1-\beta)^2} \right\} & \\ -\frac{1-\nu}{2}\left((1-\lambda)K - \overline{\theta}\right)^2, & \text{if } \lambda \geq \lambda^*. \end{cases}.$$

Observe that $G'(\lambda) > 0$, with $G(0) < 0$. If λ is large, then $d_{CN}\left(\overline{\theta}\right) = K$, and $G(\lambda)$ can be positive or negative depending on other parameter values. Furthermore, as $(1-\lambda)^2 > 0$, we have that $\triangle D \geq 0$ if and only if $G(\lambda) \geq 0$. We can therefore determine a threshold $\overline{\lambda}$ such that $\triangle D > 0$, and no delegation is preferred to full delegation, if and only if $\lambda > \overline{\lambda}$. Note that the threshold can be as high as the implicit upper bound of λ, given by $\min\left\{\frac{1-\beta}{\beta}, 1 - \frac{\overline{\theta}}{K}\right\}$ according to Assumptions 1 and 2, such that full delegation is preferred for all values of λ.

Next, consider the effect of ν. For $\nu < \nu^*$, where ν^* is defined in (8), we have $d_{CN}\left(\overline{\theta}\right) < K$, and we can write $\triangle D = \mu_1 + \mu_2\nu + \mu_3\frac{\nu^2}{1-\nu}$, where μ_1, μ_2, and μ_3 are constants, and $\mu_3 = \frac{(\triangle\theta)^2}{2(1-\lambda)^2} > 0$. As $\frac{\nu^2}{1-\nu}$ is convex in ν, $\triangle D$ is also convex in ν for $\nu < \nu^*$. For $\nu \geq \nu^*$, we have $d_{CN}\left(\overline{\theta}\right) = K$, and $\triangle D$ is linear in ν. Note, moreover, that $\triangle D > 0$ when $\nu \in \{0,1\}$. For $\nu \in (0,1)$, $\triangle D$ can be negative depending on other parameter values. In combination, these observations imply that $\triangle D$ is convex for $\nu < \nu^*$, is linear for $\nu \geq \nu^*$, is continuous in ν, and takes positive values at $\nu = 0$ and $\nu = 1$. These features can be satisfied simultaneously only if $\triangle D$ either is always positive or has two zeros, at $\underline{\nu}^{FD}$ and $\overline{\nu}^{FD}$ with $0 < \underline{\nu}^{FD} \leq \overline{\nu}^{FD} < 1$, such that $\triangle D < 0$ if and only if $\underline{\nu}^{FD} < \nu < \overline{\nu}^{FD}$. The interval $\left[\underline{\nu}^{FD}, \overline{\nu}^{FD}\right]$ can, however, be a null set, so that no delegation is the preferred choice for all values of ν. □

Proof of Lemma 5

Proof. In this proof, we make use of Lemma 3 repeatedly while determining the bureaucrat's optimal response for a given d_1. Consider first $d_{BI}\left(\overline{\theta}\right) \leq d_{CI}\left(\underline{\theta}\right)$. For $d_1 < d_{BI}\left(\underline{\theta}\right)$, the bureaucrat sets $d_{BI}^P\left(\underline{\theta}, d_1, d_2\right) = d_{BI}\left(\underline{\theta}\right)$ and $d_{BI}^P\left(\overline{\theta}, d_1, d_2\right) = d_{BI}\left(\overline{\theta}\right)$ and government's payoff is independent of d_1. For $d_{BI}\left(\underline{\theta}\right) \leq d_1 \leq d_{BI}\left(\overline{\theta}\right)$, the bureaucrat sets $d_{BI}^P\left(\underline{\theta}, d_1, d_2\right) = d_1$ and $d_{BI}^P\left(\overline{\theta}, d_1, d_2\right) = d_{BI}\left(\overline{\theta}\right)$, and government's payoff is increasing in d_1, since $d_{BI}\left(\underline{\theta}\right) \leq d_{CI}\left(\underline{\theta}\right)$. For $d_{BI}\left(\overline{\theta}\right) \leq d_1 \leq d_2$, the bureaucrat sets $d_{BI}^P\left(\underline{\theta}, d_1, d_2\right) = d_{BI}^P\left(\overline{\theta}, d_1, d_2\right) = d_1$, resulting in a uniform pollution level for both types of firms. In such a case, government's payoff increases with d_1 for $d_1 \leq \frac{E_\theta \theta}{1-\lambda}$ and decreases thereafter. Note that $\frac{E_\theta \theta}{1-\lambda} = d_{CI}(E_\theta \theta) < K$, by Assumption 2.

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

Proof of Proposition 3

Proof. Recalling the expression for $f(\theta, d)$ in (A2) and using (7), we can write government's expected payoff in the no-delegation regime as

$$\begin{aligned} E_\theta U_C^{ND} &= \nu \left[G - \frac{1}{2} d_{CN}(\underline{\theta})^2 - \frac{t_{CN}(\underline{\theta})}{1-\lambda} \right] + (1-\nu) \left[G - \frac{1}{2} d_{CN}(\overline{\theta})^2 - \frac{t_{CN}(\overline{\theta})}{1-\lambda} \right] \\ &= \nu f(\underline{\theta}, d_{CN}(\underline{\theta})) + (1-\nu) f(\overline{\theta}, d_{CN}(\overline{\theta})) - \frac{\nu \triangle\theta \left(K - d_{CN}(\overline{\theta})\right)}{1-\lambda}. \end{aligned} \tag{A6}$$

Similarly, using (5) and (10), we write government's expected payoff in the weak-delegation regime as

$$\begin{aligned} E_\theta U_C^{WD} &= \nu \left[G - \frac{1}{2} d_{CI}(\underline{\theta})^2 - \frac{t_{CI}(\underline{\theta})}{1-\lambda} \right] + (1-\nu) \left[G - \frac{1}{2} d_{BI}(\overline{\theta})^2 - \frac{t_{BI}(\overline{\theta})}{1-\lambda} \right] \\ &= \nu f(\underline{\theta}, d_{CI}(\underline{\theta})) + (1-\nu) f(\overline{\theta}, d_{BI}(\overline{\theta})), \end{aligned} \tag{A7}$$

and government's expected payoff in the strict-delegation regime as

$$\begin{aligned} E_\theta U_C^{SD} &= \nu \left[G - \frac{1}{2} d_{CI}(E_\theta \theta)^2 - \frac{\underline{\theta}(K - d_{CI}(E_\theta \theta))}{1-\lambda} \right] \\ &\quad + (1-\nu) \left[G - \frac{1}{2} d_{CI}(E_\theta \theta)^2 - \frac{\overline{\theta}(K - d_{CI}(E_\theta \theta))}{1-\lambda} \right] \\ &= \nu f(\underline{\theta}, d_{CI}(E_\theta \theta)) + (1-\nu) f(\overline{\theta}, d_{CI}(E_\theta \theta)). \end{aligned} \tag{A8}$$

We first compare the payoffs between the two regimes of weak delegation and strict delegation. We denote the payoff difference by $\triangle D^{SD-WD}$.

$$\triangle D^{SD-WD} = \nu \left[f(\underline{\theta}, d_{CI}(E_\theta \theta)) - f(\underline{\theta}, d_{CI}(\underline{\theta})) \right] + (1-\nu) \left[f(\overline{\theta}, d_{CI}(E_\theta \theta)) - f(\overline{\theta}, d_{BI}(\overline{\theta})) \right].$$

Using (A3), (A7), and (A8), as well as expressions in (5) and (10), we have

$$
\begin{aligned}
f\left(\underline{\theta}, d_{CI}\left(E_\theta\theta\right)\right)-f\left(\underline{\theta}, d_{CI}\left(\underline{\theta}\right)\right) &= \left(d_{CI}\left(E_\theta\theta\right)-d_{CI}\left(\underline{\theta}\right)\right)\left(\frac{\underline{\theta}}{1-\lambda}-\frac{d_{CI}\left(E_\theta\theta\right)+d_{CI}\left(\underline{\theta}\right)}{2}\right) \\
&= -\frac{(1-\nu)^2\left(\triangle\theta\right)^2}{2\left(1-\lambda\right)^2}, \\
f\left(\overline{\theta}, d_{CI}\left(E_\theta\theta\right)\right)-f\left(\overline{\theta}, d_{BI}\left(\overline{\theta}\right)\right) &= \left(d_{CI}\left(E_\theta\theta\right)-d_{BI}\left(\overline{\theta}\right)\right)\left(\frac{\overline{\theta}}{1-\lambda}-\frac{d_{CI}\left(E_\theta\theta\right)+d_{BI}\left(\overline{\theta}\right)}{2}\right) \\
&= \frac{1}{2\left(1-\lambda\right)^2}\left[\frac{\lambda^2\beta^2\overline{\theta}^2}{(1-\beta)^2}-\nu^2\left(\triangle\theta\right)^2\right],
\end{aligned}
$$

and

$$
\begin{aligned}
\triangle D^{SD-WD} &= -\frac{\nu\left(1-\nu\right)^2\left(\triangle\theta\right)^2}{2\left(1-\lambda\right)^2}+\frac{\left(1-\nu\right)}{2\left(1-\lambda\right)^2}\left[\frac{\lambda^2\beta^2\overline{\theta}^2}{(1-\beta)^2}-\nu^2\left(\triangle\theta\right)^2\right] \\
&= \frac{\left(1-\nu\right)}{2\left(1-\lambda\right)^2}\left[\frac{\lambda^2\beta^2\overline{\theta}^2}{(1-\beta)^2}-\nu\left(\triangle\theta\right)^2\right]. \qquad \text{(A9)}
\end{aligned}
$$

Next, we compare the payoffs between the two regimes of no delegation and weak delegation. We denote this payoff difference by $\triangle D^{ND-WD}$. Making use of (A3), (A6), and (A7), as well as expressions in (7) and (10), we have

$$
\begin{aligned}
&\triangle D^{ND-WD} \\
&= \nu\left[f\left(\underline{\theta}, d_{CN}\left(\underline{\theta}\right)\right)-f\left(\underline{\theta}, d_{CI}\left(\underline{\theta}\right)\right)\right] \\
&+\left(1-\nu\right)\left[f\left(\overline{\theta}, d_{CN}\left(\overline{\theta}\right)\right)-f\left(\overline{\theta}, d_{BI}\left(\overline{\theta}\right)\right)\right]-\frac{\nu\triangle\theta\left(K-d_{CN}\left(\overline{\theta}\right)\right)}{1-\lambda} \\
&= \left(1-\nu\right)\left[\left(d_{CN}\left(\overline{\theta}\right)-d_{BI}\left(\overline{\theta}\right)\right)\left(\frac{\overline{\theta}}{1-\lambda}-\frac{d_{CN}\left(\overline{\theta}\right)+d_{BI}\left(\overline{\theta}\right)}{2}\right)\right]-\frac{\nu\triangle\theta\left(K-d_{CN}\left(\overline{\theta}\right)\right)}{1-\lambda} \\
&= \frac{1-\nu}{2}\left[\left(\frac{\lambda\beta\overline{\theta}}{\left(1-\lambda\right)\left(1-\beta\right)}\right)^2-\left(d_{CN}\left(\overline{\theta}\right)-\frac{\overline{\theta}}{1-\lambda}\right)^2\right]-\frac{\nu\triangle\theta\left(K-d_{CN}\left(\overline{\theta}\right)\right)}{1-\lambda} \\
&= \begin{cases}
\frac{(1-\nu)}{2(1-\lambda)^2}\left[\frac{\lambda^2\beta^2\overline{\theta}^2}{(1-\beta)^2}+\frac{\nu^2(\triangle\theta)^2}{(1-\nu)^2}-\frac{2\nu\triangle\theta}{(1-\nu)}\left(\left(1-\lambda\right)K-\overline{\theta}\right)\right] & \text{if } d_{CN}\left(\overline{\theta}\right)<K \\
\\
\frac{(1-\nu)}{2(1-\lambda)^2}\left[\frac{\lambda^2\beta^2\overline{\theta}^2}{(1-\beta)^2}-\left(\left(1-\lambda\right)K-\overline{\theta}\right)^2\right] & \text{if } d_{CN}\left(\overline{\theta}\right)=K
\end{cases}. \qquad \text{(A10)}
\end{aligned}
$$

Finally, the difference in payoffs between the two regimes of no delegation and strict delegation is denoted by $\triangle D^{ND-SD}$. Using (5), (7), (A3), (A6), and (A8), we have

$$\begin{aligned}
&\triangle D^{ND-SD} \\
&= \nu \left[f\left(\underline{\theta}, d_{CN}\left(\underline{\theta}\right)\right) - f\left(\underline{\theta}, d_{CI}\left(E_{\theta}\theta\right)\right)\right] \\
&+ (1-\nu)\left[f\left(\overline{\theta}, d_{CN}\left(\overline{\theta}\right)\right) - f\left(\overline{\theta}, d_{CI}\left(E_{\theta}\theta\right)\right)\right] - \frac{\nu\triangle\theta\left(K - d_{CN}\left(\overline{\theta}\right)\right)}{1-\lambda} \\
&= \nu\left[\left(d_{CN}\left(\underline{\theta}\right) - d_{CI}\left(E_{\theta}\theta\right)\right)\left(\frac{\underline{\theta}}{1-\lambda} - \frac{d_{CN}\left(\underline{\theta}\right) + d_{CI}\left(E_{\theta}\theta\right)}{2}\right)\right] \\
&+ (1-\nu)\left[\left(d_{CN}\left(\overline{\theta}\right) - d_{CI}\left(E_{\theta}\theta\right)\right)\left(\frac{\overline{\theta}}{1-\lambda} - \frac{d_{CN}\left(\overline{\theta}\right) + d_{CI}\left(E_{\theta}\theta\right)}{2}\right)\right] - \frac{\nu\triangle\theta\left(K - d_{CN}\left(\overline{\theta}\right)\right)}{1-\lambda} \\
&= \frac{\nu(1-\nu)^2(\triangle\theta)^2}{2(1-\lambda)^2} - \frac{\nu\triangle\theta\left(K - d_{CN}\left(\overline{\theta}\right)\right)}{1-\lambda} \\
&+ \frac{(1-\nu)}{2(1-\lambda)^2}\left[\left((1-\lambda)d_{CN}\left(\overline{\theta}\right) - \overline{\theta} + \nu\triangle\theta\right)\left(\overline{\theta} + \nu\triangle\theta - (1-\lambda)d_{CN}\left(\overline{\theta}\right)\right)\right] \\
&= \frac{\nu(1-\nu)^2(\triangle\theta)^2}{2(1-\lambda)^2} - \frac{\nu\triangle\theta\left(K - d_{CN}\left(\overline{\theta}\right)\right)}{1-\lambda} + \frac{(1-\nu)}{2(1-\lambda)^2}\left[(\nu\triangle\theta)^2 - \left((1-\lambda)d_{CN}\left(\overline{\theta}\right) - \overline{\theta}\right)^2\right] \\
&= \left[\frac{\nu(1-\nu)^2(\triangle\theta)^2}{2(1-\lambda)^2} + \frac{(1-\nu)\nu^2(\triangle\theta)^2}{2(1-\lambda)^2}\right] \\
&- \frac{(1-\nu)}{2(1-\lambda)^2}\left[\frac{2(1-\lambda)\nu\triangle\theta}{(1-\nu)}\left(K - d_{CN}\left(\overline{\theta}\right)\right) + \left((1-\lambda)d_{CN}\left(\overline{\theta}\right) - \overline{\theta}\right)^2\right]
\end{aligned}$$

$$= \frac{\nu(1-\nu)(\triangle\theta)^2}{2(1-\lambda)^2} - \frac{(1-\nu)}{2(1-\lambda)^2}\left[\frac{2(1-\lambda)\nu\triangle\theta}{(1-\nu)}\left(K - d_{CN}\left(\overline{\theta}\right)\right) + \left((1-\lambda)d_{CN}\left(\overline{\theta}\right) - \overline{\theta}\right)^2\right].$$

Suppose first $d_{CN}\left(\overline{\theta}\right) < K$. Then,

$$\begin{aligned}
&\triangle D^{ND-SD} \\
&= \frac{\nu(1-\nu)(\triangle\theta)^2}{2(1-\lambda)^2} - \frac{(1-\nu)}{2(1-\lambda)^2}\left[\frac{2(1-\lambda)\nu\triangle\theta}{(1-\nu)}\left(K - d_{CN}\left(\overline{\theta}\right)\right) + \left((1-\lambda)d_{CN}\left(\overline{\theta}\right) - \overline{\theta}\right)^2\right] \\
&= \frac{\nu(1-\nu)(\triangle\theta)^2}{2(1-\lambda)^2} - \frac{(1-\nu)}{2(1-\lambda)^2}\left[\left(\frac{2\nu\triangle\theta}{1-\nu}\right)\left((1-\lambda)K - \overline{\theta} - \frac{\nu\triangle\theta}{1-\nu}\right) + \left(\frac{\nu\triangle\theta}{1-\nu}\right)^2\right] \\
&= \frac{\nu(1-\nu)(\triangle\theta)^2}{2(1-\lambda)^2} + \frac{\nu^2(\triangle\theta)^2}{2(1-\lambda)^2(1-\nu)} - \frac{\nu\triangle\theta}{(1-\lambda)^2}\left((1-\lambda)K - \overline{\theta}\right) \\
&= \frac{\nu\left(1-\nu+\nu^2\right)(\triangle\theta)^2}{2(1-\lambda)^2(1-\nu)} - \frac{\nu\triangle\theta}{(1-\lambda)^2}\left((1-\lambda)K - \overline{\theta}\right).
\end{aligned}$$

Suppose next $d_{CN}\left(\overline{\theta}\right)=K$. Then,

$$\begin{aligned}
&\triangle D^{ND-SD} \\
&= \frac{\nu\left(1-\nu\right)\left(\triangle\theta\right)^2}{2\left(1-\lambda\right)^2} - \frac{\left(1-\nu\right)}{2\left(1-\lambda\right)^2}\left[\frac{2\left(1-\lambda\right)\nu\triangle\theta}{\left(1-\nu\right)}\left(K-d_{CN}\left(\overline{\theta}\right)\right)+\left(\left(1-\lambda\right)d_{CN}\left(\overline{\theta}\right)-\overline{\theta}\right)^2\right] \\
&= \frac{\left(1-\nu\right)}{2\left(1-\lambda\right)^2}\left[\nu\left(\triangle\theta\right)^2-\left(\left(1-\lambda\right)K-\overline{\theta}\right)^2\right].
\end{aligned}$$

Together, this means

$$\triangle D^{ND-SD}=\begin{cases}\frac{\nu\left(1-\nu+\nu^2\right)\left(\triangle\theta\right)^2}{2\left(1-\lambda\right)^2\left(1-\nu\right)}-\frac{\nu\triangle\theta}{\left(1-\lambda\right)^2}\left(\left(1-\lambda\right)K-\overline{\theta}\right) & \text{if } d_{CN}\left(\overline{\theta}\right)<K \\ \\ \frac{\left(1-\nu\right)}{2\left(1-\lambda\right)^2}\left[\nu\left(\triangle\theta\right)^2-\left(\left(1-\lambda\right)K-\overline{\theta}\right)^2\right] & \text{if } d_{CN}\left(\overline{\theta}\right)=K\end{cases}. \tag{A11}$$

Consider the effects of λ and β. For a fixed ν, we can write $\triangle D^{ND-SD}$ as $\frac{A(\lambda)}{(1-\lambda)^2}$ where $A\left(\lambda\right)$ is a continuous and $A'\left(\lambda\right)>0$. Thus, we can determine a threshold value λ^{ND} such that $\triangle D^{ND-SD}>0$ if and only if $\lambda>\lambda^{ND}$. Note that λ^{ND} does not vary with β, as $\triangle D^{ND-SD}$ is independent of β. Similarly, we can write $\triangle D^{ND-WD}$ as $\frac{B(\lambda,\beta)}{(1-\lambda)^2}$, and $\triangle D^{SD-WD}$ as $\frac{C(\lambda,\beta)}{(1-\lambda)^2}$, where $B\left(\lambda,\beta\right)$ and $C\left(\lambda,\beta\right)$ are continuous functions that are increasing in λ and β. We can therefore determine threshold values $\lambda^1\left(\beta\right)$ and $\lambda^2\left(\beta\right)$ such that $\triangle D^{ND-WD}>0$ is positive if and only if $\lambda>\lambda^1\left(\beta\right)$ and $\triangle D^{SD-WD}>0$ if and only if $\lambda>\lambda^2\left(\beta\right)$. Furthermore, $\lambda^1\left(\beta\right)$ and $\lambda^2\left(\beta\right)$ are decreasing in β, since $B\left(\lambda,\beta\right)$ and $C\left(\lambda,\beta\right)$ increase with β.

Define $\lambda^D\left(\beta\right):=\min\left\{\lambda^1\left(\beta\right),\lambda^2\left(\beta\right)\right\}$. If $\lambda<\lambda^D\left(\beta\right)$, then government receives the highest value in the weak-delegation regime and chooses weak delegation in equilibrium. In contrast, if $\lambda\geq\lambda^D\left(\beta\right)$, then government prefers either no delegation or strict delegation, depending on whether λ is above or below λ^{ND}.

To find the effect of ν, fix λ and β. From (A9), note that $\triangle D^{SD-WD}>0$ if and only $\nu<\left(\frac{\lambda\beta\overline{\theta}}{(1-\beta)\triangle\theta}\right)^2$. Thus, between the two partial-delegation rules, government prefers strict delegation to weak delegation if and only if $\nu<\hat{\nu}$, where $\hat{\nu}$ is defined in the Proposition. Consequently, for $\nu\leq\hat{\nu}$, strict delegation or no delegation can occur in equilibrium, while for $\nu\geq\hat{\nu}$, only weak delegation or no delegation can occur in equilibrium. We will consider these two cases separately. To find out the equilibrium outcome in these two cases, we need to study the effect of ν on $\triangle D^{ND-SD}$ and $\triangle D^{ND-WD}$. Recall the definition of ν^*, in (8), and that $d_{CN}\left(\overline{\theta}\right)<K$ if and only if $\nu<\nu^*$. Note that $\triangle D^{ND-SD}$ and $\triangle D^{ND-WD}$ are continuous in ν but not necessarily differentiable at $d_{CN}\left(\overline{\theta}\right)=K$, or, equivalently, at $\nu=\nu^*$, and that ν^* can lie below or above $\hat{\nu}$, depending on other parameter values.

First, consider $\nu\leq\hat{\nu}$. If now $\triangle D^{ND-SD}>0$, then no delegation occurs in equilibrium; otherwise, strict delegation occurs in equilibrium. We first study the expression of $\triangle D^{ND-SD}$

in (A11) for all ν in $[0,1]$. Then, we will consider the constraint that $\nu \leq \hat{\nu}$.

The following two observations are useful in studying the effect of ν on $\triangle D^{ND-SD}$. First, $\triangle D^{ND-SD} = 0$ at both $\nu = 0$ and $\nu = 1$. Secondly, $\frac{\partial^2}{\partial \nu^2}\left(\triangle D^{ND-SD}\right) = \frac{\nu\left(3-3\nu+\nu^2\right)(\triangle\theta)^2}{(1-\nu)^3(1-\lambda)^2} > 0$ for $\nu < \nu^*$, and $\frac{\partial^2}{\partial \nu^2}\left(\triangle D^{ND-SD}\right) = -\frac{(\triangle\theta)^2}{(1-\lambda)^2} < 0$ for $\nu > \nu^*$. Therefore, $\triangle D^{ND-SD}$ is convex in ν for $\nu < \nu^*$, and concave in ν for $\nu > \nu^*$. If $\triangle D^{ND-SD} \leq 0$ at $\nu = \nu^*$, then the above two observations together imply that, either $\triangle D^{ND-SD} \leq 0$ for every $0 < \nu < 1$, or $\triangle D^{ND-SD}$ has exactly one zero at some $\underline{\nu}^k \in (\nu^*, 1)$, such that $\triangle D^{ND-SD} < 0$ for every $0 < \nu < \underline{\nu}^k$, and $\triangle D^{ND-SD} > 0$ for every $\underline{\nu}^k < \nu < 1$. In contrast, if $\triangle D^{ND-SD} > 0$ at $\nu = \nu^*$, then the two observations together imply that, either $\triangle D^{ND-SD} > 0$ for every $0 < \nu < 1$, or $\triangle D^{ND-SD}$ has exactly one zero at some $\underline{\nu}^\ell \in (0, \nu^*)$, such that $\triangle D^{ND-SD} < 0$ for every $0 < \nu < \underline{\nu}^\ell$, and $\triangle D^{ND-SD} > 0$ for every $\underline{\nu}^\ell < \nu < 1$. While $\triangle D^{ND-SD}$ can take either sign at $\nu = \nu^*$, we see that the following three possibilities are the only ones that can arise: (a) $\triangle D^{ND-SD} > 0$ for $\nu \in (0,1)$; (b) $\triangle D^{ND-SD} < 0$ for $\nu \in (0,1)$; and (c) $\triangle D^{ND-SD} < 0$ for $\nu \in (0, \underline{\nu}^m)$ and $\triangle D^{ND-SD} > 0$ for $\nu \in (\underline{\nu}^m, 1)$, where $0 < \underline{\nu}^m < 1$, and $\underline{\nu}^m \in \left\{\underline{\nu}^k, \underline{\nu}^\ell\right\}$. Now we adjust to the constraint $\nu \leq \hat{\nu}$, and define the threshold $\underline{\nu}^{SD}$ as 0 in scenario (a), as $\hat{\nu}$ in scenario (b), and as $\min\{\hat{\nu}, \underline{\nu}^m\}$ in scenario (c). Therefore, $\underline{\nu}^{SD} \in [0, \hat{\nu}]$. Furthermore, strict delegation occurs in equilibrium if $0 < \nu \leq \underline{\nu}^{SD}$; and no delegation occurs in equilibrium if $\underline{\nu}^{SD} \leq \nu \leq \hat{\nu}$.

Next, consider $\nu \geq \hat{\nu}$. If now $\triangle D^{ND-WD} > 0$, then no delegation occurs in equilibrium; otherwise, weak delegation occurs in equilibrium. As before, we will first study the expression of $\triangle D^{ND-WD}$ in (A10) for all ν in $[0,1]$ and then adjust for the constraint that $\nu \geq \hat{\nu}$.

The following two observations are useful in this case. First, $\triangle D^{ND-WD} > 0$ at $\nu = 0$, and $\triangle D^{ND-WD} = 0$ at $\nu = 1$. Secondly, $\frac{\partial^2}{\partial \nu^2}\left(\triangle D^{ND-WD}\right) = \frac{(\triangle\theta)^2}{(1-\nu)^3(1-\lambda)^2} > 0$ for $\nu < \nu^*$, and $\frac{\partial^2}{\partial \nu^2}\left(\triangle D^{ND-WD}\right) = 0$ for $\nu > \nu^*$. Therefore, $\triangle D^{ND-WD}$ is convex in ν for $\nu < \nu^*$ and linear in ν for $\nu \geq \nu^*$. If $\triangle D^{ND-WD} < 0$ at $\nu = \nu^*$, then the above two observations together imply that $\triangle D^{ND-WD}$ has exactly one zero at some $\underline{\nu}^r \in (0, \nu^*)$, such that $\triangle D^{ND-WD} > 0$ for every $0 < \nu < \underline{\nu}^r$, and $\triangle D^{ND-WD} < 0$ for every $\underline{\nu}^r < \nu < 1$. In contrast, if $\triangle D^{ND-WD} \geq 0$ at $\nu = \nu^*$, then the two observations together imply that, either $\triangle D^{ND-WD} \geq 0$ for every $0 < \nu < 1$, or $\triangle D^{ND-WD}$ can have exactly two zeros at $\underline{\nu}^s$ and $\overline{\nu}^s$, with $0 < \underline{\nu}^s \leq \overline{\nu}^s < \nu^*$, such that $\triangle D^{ND-WD} < 0$ if and only if $\underline{\nu}^s < \nu < \overline{\nu}^s$. While $\triangle D^{ND-WD}$ can take either sign at $\nu = \nu^*$, the following three possibilities are the only ones that can arise: (a) $\triangle D^{ND-WD} > 0$ for $\nu \in \left(0, \underline{\nu}^t\right)$ and $\triangle D^{ND-WD} < 0$ for $\nu \in \left(\underline{\nu}^t, 1\right)$, where $0 < \underline{\nu}^t < 1$, and $\underline{\nu}^t \in \{\underline{\nu}^r, \underline{\nu}^s\}$; (b) $\triangle D^{ND-WD} \geq 0$ for all $\nu \in (0,1)$; and (c) $\triangle D^{ND-WD} > 0$ for $\nu \in \left(0, \underline{\nu}^t\right) \cup (\overline{\nu}^s, 1)$ and $\triangle D^{ND-WD} < 0$ for all $\nu \in \left(\underline{\nu}^t, \overline{\nu}^s\right)$ and $0 < \underline{\nu}^t \leq \overline{\nu}^s < 1$. Adjusting for the constraint $\nu \geq \hat{\nu}$, we define the thresholds $\underline{\nu}^{WD}$ and $\overline{\nu}^{WD}$ as $\max\left\{\underline{\nu}^t, \hat{\nu}\right\}$ and 1, respectively, in scenario (a); as 1 and 1, respectively, in scenario (b); and as $\max\left\{\underline{\nu}^t, \hat{\nu}\right\}$ and $\max\{\overline{\nu}^s, \hat{\nu}\}$, respectively, in scenario (c). Clearly, $\hat{\nu} \leq \underline{\nu}^{WD} \leq \overline{\nu}^{WD} \leq 1$. Furthermore, weak delegation occurs in equilibrium if $\underline{\nu}^{WD} \leq \nu \leq \overline{\nu}^{WD}$, while no delegation occurs in equilibrium if $\hat{\nu} \leq \nu \leq \underline{\nu}^{WD}$, or if $\overline{\nu}^{WD} \leq \nu < 1$. □

Appendix B

In this Appendix, we describe and analyze an alternate framework, which we call Permits. Unlike in the regulation framework, here the firm produces the good privately and sells it in the market. Government regulates production by issuing pollution-contingent permits to compensate for consumer's disutility from pollution. The firm makes a transfer to government to purchase these permits, and the transfer is affected by bureaucratic cost. A contract $\alpha \in A$ now determines a transfer t from the firm to government and a pollution level d. Letting $p(\theta)$ denote the market price set by the firm of type θ, we will simply put $p(\theta) = G$, so that all the gross value of the firm's production accrues to the firm. For a given permit contract $\alpha = (t, d)$, the firm's profit thus is

$$U_P^T(\theta, \alpha) = G - \theta(K - d) - t; \tag{B1}$$

note that we use the superscript T to distinguish this case from the one in the main text. We consider pairs of contracts $(\underline{\alpha}, \overline{\alpha}) = \left((\underline{t}, \underline{d}), (\overline{t}, \overline{d})\right) \in A^2$ satisfying *incentive-compatibility* constraints

$$-\overline{\theta}\left(K - \overline{d}\right) - \overline{t} \geq -\overline{\theta}(K - \underline{d}) - \underline{t}, \tag{ICH-T}$$

$$-\underline{\theta}(K - \underline{d}) - \underline{t} \geq -\underline{\theta}\left(K - \overline{d}\right) - \overline{t}. \tag{ICL-T}$$

Like in the main text, a contract (t, d) satisfies the *individual-rationality* constraint if

$$G - \theta(K - d) - t \geq 0, \tag{IR-T}$$

and a pair of contracts$\left((\underline{t}, \underline{d}), (\overline{t}, \overline{d})\right)$ satisfy individual-rationality constraints if

$$G - \overline{\theta}\left(K - \overline{d}\right) - \overline{t} \geq 0, \tag{IRH-T}$$

$$G - \underline{\theta}(K - \underline{d}) - \underline{t} \geq 0. \tag{IRL-T}$$

Permit fees have a bureaucratic cost. For each unit of fee transferred, government receives a fraction $1-\lambda$ of it, and the remaining fraction λ is consumed by the bureaucracy. Government's payoff is therefore

$$U_C^T(\alpha) = G + (1-\lambda)t - \frac{1}{2}d^2 - p = (1-\lambda)t - \frac{1}{2}d^2. \tag{B2}$$

An informed bureaucrat can implement a type-contingent transfer policy. The bureaucrat has a vested interest in the transfer, and her payoff is

$$U_B^T(\theta, \alpha) = \beta\lambda t + (1-\beta)U_C^T(\alpha) = (1-\beta)\left[\left(1 + \lambda\frac{2\beta - 1}{1-\beta}\right)t - \frac{1}{2}d^2\right]. \tag{B3}$$

Comparison between full delegation and no delegation

The contract that government would have chosen for type θ if it had perfect information solves the following problem:

$$\max_{\alpha} G + (1-\lambda)\, t - \frac{1}{2}d^2 - G, \text{ subject to (IR-T).} \tag{B4}$$

The fact that (IR-T) is binding, together with the first-order condition of (B4), give us the optimal contract $\alpha_{CI}^{T}(\theta) = \left(t_{CI}^{T}(\theta), d_{CI}^{T}(\theta)\right)$, where

$$d_{CI}^{T}(\theta) = (1-\lambda)\,\theta,\; t_{CI}^{T}(\theta) = G - \theta\left(K - d_{CI}^{T}(\theta)\right). \tag{B5}$$

When an uninformed government offers an incentive-compatible pair of contracts $(\underline{\alpha}, \overline{\alpha})$ to the firm, the optimal contract pair solves:

$$\max_{\underline{\alpha},\overline{\alpha}} \nu\left[(1-\lambda)\,\underline{t} - \frac{1}{2}\underline{d}^2\right] + (1-\nu)\left[(1-\lambda)\,\overline{t} - \frac{1}{2}\overline{d}^2\right], \tag{B6}$$
$$\text{subject to (IRH-T), (IRL-T), (ICH-T), and (ICL-T).}$$

As (ICL-T) and (IRH-T) are binding, the first-order condition of (B6), give us the optimal contract pair $\left(\alpha_{CN}^{T}(\underline{\theta}), \alpha_{CN}^{T}(\overline{\theta})\right) = \left(\left(t_{CN}^{T}(\underline{\theta}), d_{CN}^{T}(\underline{\theta})\right), \left(t_{CN}^{T}(\overline{\theta}), d_{CN}^{T}(\overline{\theta})\right)\right)$, where

$$\begin{aligned}
d_{CN}^{T}(\underline{\theta}) &= (1-\lambda)\,\underline{\theta}, \\
d_{CN}^{T}(\overline{\theta}) &= \min\left\{(1-\lambda)\left(\overline{\theta} + \frac{\nu}{1-\nu}\triangle\theta\right), K\right\}, \\
t_{CN}^{T}(\underline{\theta}) &= G - \overline{\theta}\left(K - d_{CN}^{T}(\underline{\theta})\right) - \triangle\theta\left(K - d_{CN}^{T}(\overline{\theta})\right), \\
t_{CN}^{T}(\overline{\theta}) &= G - \overline{\theta}\left(K - d_{CN}^{T}(\overline{\theta})\right).
\end{aligned} \tag{B7}$$

When an informed bureaucrat, if delegated, offers a type-contingent pair of contracts to the firm, the optimal contract for type θ solves:

$$\max_{(t,d)} (1-\beta)\left[\left(1 + \lambda\frac{2\beta-1}{1-\beta}\right)t - \frac{1}{2}d^2\right], \tag{B8}$$
$$\text{subject to (IR-T).}$$

The fact that (IR-T) is binding, together with the first-order condition of (B8), give us the optimal contract $\alpha_{BI}^{T}(\theta) = \left(t_{BI}^{T}(\theta), d_{BI}^{T}(\theta)\right)$, where

$$d_{BI}^{T}(\theta) = \min\left\{\left(1+\lambda\frac{2\beta-1}{1-\beta}\right)\theta, K\right\},$$
$$t_{BI}^{T}(\theta) = G-\theta\left(K-d_{BI}^{T}(\theta)\right). \tag{B9}$$

Unlike the regulation framework, the bureaucrat's preferred pollution level is now always above that of government as the transfer flows from the firm to government in the permits framework. In order to increase the transfer, the bureaucrat allows the firm to over-pollute.[23]

We compare government's payoff in the two cases to derive the condition under which government prefers no delegation to full delegation:

$$\triangle D^{T} := E_{\theta}U_{C}^{T}\left(\alpha_{CN}^{T}(\theta)\right) - E_{\theta}U_{C}^{T}\left(\alpha_{BI}^{T}(\theta)\right) \geq 0. \tag{B10}$$

The effects of β and υ are qualitatively similar to their effects on $\triangle D$, derived for the regulation framework in the main text; see Proposition 1, parts (i) and (iii). We describe them in the following proposition. The effect of λ is, however, different in this permits setting. In particular, it is not the case here that full delegation being optimum for some $\lambda' \in (0,1)$ implies that it will also be optimum for all $\lambda < \lambda'$. In fact, cases exist where, in the permits framework, no delegation is preferred to full delegation for both the highest and the lowest values of λ.[24]

Proposition B.1. Consider the game in which government chooses between full delegation and no delegation. The equilibrium is characterized as follows:

(i) For given λ and ν, there exists a threshold $\overline{\beta}^{T}$ such that government prefers no delegation to full delegation if and only if $\beta \geq \overline{\beta}^{T}$.

(ii) For given λ and β, there exist $0 < \underline{\nu}^{FDT} \leq \overline{\nu}^{FDT} < 1$ such that government prefers full delegation to no delegation if and only if $\underline{\nu}^{FDT} \leq \nu \leq \overline{\nu}^{FDT}$.

Proof. Define $h(\theta,d) := U_{C}^{T}(G-\theta(K-d), d)$. Then,

$$h(\theta,d) = (1-\lambda)(G-\theta K) + \left((1-\lambda)\theta d - \frac{1}{2}d^{2}\right). \tag{B11}$$

Observe that

$$h(\theta,d_1) - h(\theta,d_2) = (d_1-d_2)\left((1-\lambda)\theta - \frac{d_1+d_2}{2}\right). \tag{B12}$$

[23] As we shall see shortly, this complicates the comparison of no delegation and full delegation in Proposition B.1, since we there have to take into account the possibility that also the bureaucrat, with her interest for high pollution levels, may choose a no-regulation contract for the high-cost type, with $d_{BI}^{T}\left(\overline{\theta}\right) = K$. This complication does not show up in the subsequent analysis of partial delegation in Proposition B.2, since government, when capping the bias, restricts the bureaucrat from choosing such a high pollution level.

[24] We do not know whether the effect of λ on the decision to choose full delegation over no delegation can be even more complicated than that.

Government's expected payoff in the no-delegation regime is given by

$$\begin{aligned}
& E_\theta U_C^T\left(\alpha_{CN}^T(\theta)\right) \\
= \; & \nu\left[(1-\lambda)\, t_{CN}^T(\underline{\theta}) - \frac{1}{2} d_{CN}^T(\underline{\theta})^2\right] + (1-\nu)\left[(1-\lambda)\, t_{CN}^T(\overline{\theta}) - \frac{1}{2} d_{CN}^T(\overline{\theta})^2\right] \\
= \; & \nu h\left(\underline{\theta}, d_{CN}^T(\underline{\theta})\right) + (1-\nu)\, h\left(\overline{\theta}, d_{CN}^T(\overline{\theta})\right) - \nu(1-\lambda)\triangle\theta\left(K - d_{CN}^T(\overline{\theta})\right),
\end{aligned}$$

where the second equality follows from using (B11) and inserting from (B7). Similarly, government's expected payoff in the full-delegation regime is given by

$$\begin{aligned}
& E_\theta U_C^T\left(\alpha_{BI}^T(\theta)\right) \\
= \; & \nu\left[(1-\lambda)\, t_{BI}^T(\underline{\theta}) - \frac{1}{2} d_{BI}^T(\underline{\theta})^2\right] + (1-\nu)\left[(1-\lambda)\, t_{BI}^T(\overline{\theta}) - \frac{1}{2} d_{BI}^T(\overline{\theta})^2\right] \\
= \; & \nu h\left(\underline{\theta}, d_{BI}^T(\underline{\theta})\right) + (1-\nu)\, h\left(\overline{\theta}, d_{BI}^T(\overline{\theta})\right),
\end{aligned}$$

where the second equality follows from again using (B11) and inserting from (B9). We can thus compute the payoff difference between the two regimes, ΔD^T defined in (B10), as

$$\begin{aligned}
\triangle D^T = & \; \nu\left[h\left(\underline{\theta}, d_{CN}^T(\underline{\theta})\right) - h\left(\underline{\theta}, d_{BI}^T(\underline{\theta})\right)\right] \\
& + (1-\nu)\left[h\left(\overline{\theta}, d_{CN}^T(\overline{\theta})\right) - h\left(\overline{\theta}, d_{BI}^T(\overline{\theta})\right)\right] - \nu(1-\lambda)\triangle\theta\left(K - d_{CN}^T(\overline{\theta})\right).
\end{aligned} \tag{B13}$$

We first use (B12) to expand the first term in (B13), inserting for $d_{CN}^T(\underline{\theta})$ and $d_{BI}^T(\underline{\theta})$ from (B7) and (B9), respectively, to get

$$\begin{aligned}
& \nu\left[h\left(\underline{\theta}, d_{CN}^T(\underline{\theta})\right) - h\left(\underline{\theta}, d_{BI}^T(\underline{\theta})\right)\right] \\
& = \nu\left(d_{CN}^T(\underline{\theta}) - d_{BI}^T(\underline{\theta})\right)\left((1-\lambda)\,\underline{\theta} - \frac{d_{CN}(\underline{\theta}) + d_{BI}(\underline{\theta})}{2}\right) \\
& = \frac{\nu\left(d_{CN}^T(\underline{\theta}) - d_{BI}^T(\underline{\theta})\right)^2}{2} \\
& = \begin{cases} \frac{\nu\lambda^2\beta^2\underline{\theta}^2}{2(1-\beta)^2}, & \text{if } d_{BI}^T(\underline{\theta}) < K; \\ \frac{\nu}{2}\left(K - (1-\lambda)\,\underline{\theta}\right)^2, & \text{if } d_{BI}^T(\underline{\theta}) = K. \end{cases}
\end{aligned} \tag{B14}$$

Next, we similarly use (B12) to expand the second term in (B13), inserting for $d_{CN}^T(\overline{\theta})$ and

$d_{BI}^{T}\left(\overline{\theta}\right)$ from (B7) and (B9), respectively, to get

$$
\begin{aligned}
&(1-\nu)\left[h\left(\overline{\theta},d_{CN}^{T}\left(\overline{\theta}\right)\right)-h\left(\overline{\theta},d_{BI}^{T}\left(\overline{\theta}\right)\right)\right] \\
&=(1-\nu)\left[\left(d_{CN}^{T}\left(\overline{\theta}\right)-d_{BI}^{T}\left(\overline{\theta}\right)\right)\left((1-\lambda)\overline{\theta}-\frac{d_{CN}^{T}\left(\overline{\theta}\right)+d_{BI}^{T}\left(\overline{\theta}\right)}{2}\right)\right] \\
&=\begin{cases}
\frac{(1-\nu)}{2}\left((1-\lambda)\overline{\theta}+(1-\lambda)\frac{\nu}{1-\nu}\triangle\theta-d_{BI}^{T}\left(\overline{\theta}\right)\right)\times \\
\left((1-\lambda)\overline{\theta}-(1-\lambda)\left(\frac{\nu}{1-\nu}\triangle\theta\right)-d_{BI}^{T}\left(\overline{\theta}\right)\right), & \text{if } d_{CN}^{T}\left(\overline{\theta}\right)<K; \\
(1-\nu)\left[\left(K-d_{BI}^{T}\left(\overline{\theta}\right)\right)\left((1-\lambda)\overline{\theta}-\frac{K+d_{BI}^{T}\left(\overline{\theta}\right)}{2}\right)\right], & \text{if } d_{CN}^{T}\left(\overline{\theta}\right)=K.
\end{cases} \\
&=\begin{cases}
\frac{(1-\nu)}{2}\left[\frac{\lambda^{2}\beta^{2}\overline{\theta}^{2}}{(1-\beta)^{2}}-\frac{(1-\lambda)^{2}\nu^{2}(\triangle\theta)^{2}}{(1-\nu)^{2}}\right], & \text{if } d_{CN}^{T}\left(\overline{\theta}\right)<K \text{ and } d_{BI}^{T}\left(\overline{\theta}\right)<K; \\
\frac{(1-\nu)}{2}\left[\left(K-(1-\lambda)\overline{\theta}\right)^{2}-\frac{(1-\lambda)^{2}\nu^{2}(\triangle\theta)^{2}}{(1-\nu)^{2}}\right], & \text{if } d_{CN}^{T}\left(\overline{\theta}\right)<d_{BI}^{T}\left(\overline{\theta}\right)=K; \\
-\frac{(1-\nu)}{2}\left[\left(K-\left(1+\lambda\frac{2\beta-1}{1-\beta}\right)\overline{\theta}\right)\left(K-\left(1-\frac{\lambda}{1-\beta}\right)\overline{\theta}\right)\right], & \text{if } d_{BI}^{T}\left(\overline{\theta}\right)<d_{CN}^{T}\left(\overline{\theta}\right)=K; \\
0, & \text{if } d_{CN}^{T}\left(\overline{\theta}\right)=d_{BI}^{T}\left(\overline{\theta}\right)=K.
\end{cases}
\end{aligned}
\tag{B15}
$$

We use the expression for $d_{CN}^{T}\left(\overline{\theta}\right)$ in (B7) to write the third term in (B13) as

$$
\nu(1-\lambda)\triangle\theta\left(K-d_{CN}^{T}\left(\overline{\theta}\right)\right)=\begin{cases}
\nu(1-\lambda)\triangle\theta\left(K-(1-\lambda)\left(\overline{\theta}+\frac{\nu}{1-\nu}\triangle\theta\right)\right), & \text{if } d_{CN}^{T}\left(\overline{\theta}\right)<K; \\
0, & \text{if } d_{CN}^{T}\left(\overline{\theta}\right)=K.
\end{cases}
\tag{B16}
$$

Combining these expressions, we can write ΔD^{T} in (B13) as

$$\triangle D^T = \begin{cases} \frac{\nu\lambda^2\beta^2\underline{\theta}^2}{2(1-\beta)^2} + \frac{(1-\nu)}{2}\left[\frac{\lambda^2\beta^2\overline{\theta}^2}{(1-\beta)^2} - \frac{(1-\lambda)^2\nu^2(\triangle\theta)^2}{(1-\nu)^2}\right] & \\ -\nu(1-\lambda)\triangle\theta\left(K-(1-\lambda)\left(\overline{\theta}+\frac{\nu}{1-\nu}\triangle\theta\right)\right), & \text{if } d^T_{CN}\left(\overline{\theta}\right)<K,\, d^T_{BI}\left(\underline{\theta}\right)<d^T_{BI}\left(\overline{\theta}\right)<K; \\ \frac{\nu\lambda^2\beta^2\underline{\theta}^2}{2(1-\beta)^2} + \frac{(1-\nu)}{2}\left[\left(K-(1-\lambda)\overline{\theta}\right)^2 - \frac{(1-\lambda)^2\nu^2(\triangle\theta)^2}{(1-\nu)^2}\right] & \\ -\nu(1-\lambda)\triangle\theta\left(K-(1-\lambda)\left(\overline{\theta}+\frac{\nu}{1-\nu}\triangle\theta\right)\right), & \text{if } d^T_{CN}\left(\overline{\theta}\right)<K,\, d^T_{BI}\left(\underline{\theta}\right)<d^T_{BI}\left(\overline{\theta}\right)=K; \\ \frac{\nu}{2}\left(K-(1-\lambda)\underline{\theta}\right)^2 + \frac{(1-\nu)}{2}\left[\left(K-(1-\lambda)\overline{\theta}\right)^2\right. & \\ \left.-\frac{(1-\lambda)^2\nu^2(\triangle\theta)^2}{(1-\nu)^2}\right] & \\ -\nu(1-\lambda)\triangle\theta\left(K-(1-\lambda)\left(\overline{\theta}+\frac{\nu}{1-\nu}\triangle\theta\right)\right), & \text{if } d^T_{CN}\left(\overline{\theta}\right)<d^T_{BI}\left(\underline{\theta}\right)=d^T_{BI}\left(\overline{\theta}\right)=K; \\ \frac{\nu\lambda^2\beta^2\underline{\theta}^2}{2(1-\beta)^2} - \frac{(1-\nu)}{2}\left(K-\left(1+\lambda\frac{2\beta-1}{1-\beta}\right)\overline{\theta}\right)\times & \\ \left(K-\left(1-\frac{\lambda}{1-\beta}\right)\overline{\theta}\right), & \text{if } d^T_{BI}\left(\underline{\theta}\right)<d^T_{BI}\left(\overline{\theta}\right)<d^T_{CN}\left(\overline{\theta}\right)=K; \\ \frac{\nu\lambda^2\beta^2\underline{\theta}^2}{2(1-\beta)^2}, & \text{if } d^T_{BI}\left(\underline{\theta}\right)<d^T_{BI}\left(\overline{\theta}\right)=d^T_{CN}\left(\overline{\theta}\right)=K; \\ \frac{\nu}{2}\left(K-(1-\lambda)\overline{\theta}\right)^2, & \text{if } d^T_{BI}\left(\underline{\theta}\right)=d^T_{BI}\left(\overline{\theta}\right)=d^T_{CN}\left(\overline{\theta}\right)=K. \end{cases} \tag{B17}$$

Consider first the effect of β. Define $\beta^T(\theta) := \frac{K+\lambda\theta}{K+2\lambda\theta} > \frac{1}{2}$. Note that $d^T_{BI}(\theta) < K$, for $\beta < \beta^T(\theta)$, and $d^T_{BI}(\theta) = K$, otherwise. From (B17), we see that ΔD^T is increasing in β for $\beta < \beta^T(\underline{\theta})$, and independent of β otherwise.[25] Thus, we can determine a threshold $\overline{\beta}^T$ such that $\triangle D^T > 0$, and no delegation is preferred to full delegation, if and only if $\beta > \overline{\beta}^T$. Note that this threshold can be zero or one, such that one regime is preferred for all values of β, depending on other parameter values.

Next, consider the effect of ν. Note that $d^T_{CN}\left(\overline{\theta}\right) < K$ if and only if $\nu < \nu^T$, where

$$\nu^T := \frac{K-(1-\lambda)\overline{\theta}}{K-(1-\lambda)\underline{\theta}} \in [0,1]. \tag{B18}$$

For $\nu < \nu^T$, we can write $\triangle D^T = \mu_1 + \mu_2\nu + \mu_3\frac{\nu^2}{1-\nu}$, where μ_1, μ_2, and μ_3 are constants, and $\mu_3 = \frac{(1-\lambda)^2(\triangle\theta)^2}{2} > 0$. As $\frac{\nu^2}{1-\nu}$ is convex in ν, $\triangle D$ is also convex in ν for $\nu < \nu^T$. For $\nu \geq \nu^T$, we

[25] Showing that $\triangle D^T$ is increasing in β in the case $d^T_{BI}\left(\underline{\theta}\right) < d^T_{BI}\left(\overline{\theta}\right) < d^T_{CN}\left(\overline{\theta}\right) = K$ is tedious but straightforward.

have $d_{CN}^T\left(\overline{\theta}\right) = K$, and $\triangle D^T$ is linear in ν. Note, moreover, that $\triangle D^T > 0$ when $\nu \in \{0,1\}$. For $\nu \in (0,1)$, $\triangle D^T$ can be negative, depending on other parameter values. In combination, these observations imply that $\triangle D^T$ is convex for $\nu < \nu^T$, is linear for $\nu \geq \nu^T$, is continuous in ν, and takes positive values at $\nu = 0$ and $\nu = 1$. These features can be satisfied simultaneously only if $\triangle D^T$ either is always positive or has two zeros, at $\underline{\nu}^{FDT}$ and $\overline{\nu}^{FDT}$ with $0 < \underline{\nu}^{FDT} \leq \overline{\nu}^{FDT} < 1$, such that $\triangle D^T < 0$ if and only if $\underline{\nu}^{FDT} < \nu < \overline{\nu}^{FDT}$. The interval $\left[\underline{\nu}^{FDT}, \overline{\nu}^{FDT}\right]$ can, however, be a null set, such that no delegation is the preferred choice for all values of ν. □

Partial delegation

Like in the regulation framework, government can improve its value by restricting the choice of the bureaucrat. We impose a restriction that the bureaucrat chooses regulatory contracts $\alpha(\theta), \theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$ with the constraint that $d(\theta) \in [d_1, d_2] \subseteq [0, K]$. The bureaucrat's optimal contract for type θ solves

$$\max_{(t,d)} (1-\beta)\left[\left(1+\lambda\frac{2\beta-1}{1-\beta}\right)t - \frac{1}{2}d^2\right], \tag{B19}$$
$$\text{subject to (IR-T), and } d \in [d_1, d_2].$$

We denote the solution with a superscript PT and a subscript BI. The following lemma describes the bureaucrat's choice of contracts. The proof is similar to that of Lemma 3.

Lemma B.1. Assume that government delegates the decision-making authority with the restriction that $d \in [d_1, d_2] \subseteq [0, K]$. The bureaucrat's regulation contract for type $\theta \in \left\{\underline{\theta}, \overline{\theta}\right\}$ is given by $\alpha_{BI}^{PT}(\theta, d_1, d_2) = \left(t_{BI}^{PT}(\theta, d_1, d_2), d_{BI}^{PT}(\theta, d_1, d_2)\right)$, where

$$\begin{aligned} d_{BI}^{PT}(\theta, d_1, d_2) &= \begin{cases} d_1 & \text{if } d_1 \geq \left(1+\lambda\frac{2\beta-1}{1-\beta}\right)\theta \\ \left(1+\lambda\frac{2\beta-1}{1-\beta}\right)\theta & \text{if } d_1 < \left(1+\lambda\frac{2\beta-1}{1-\beta}\right)\theta < d_2 \\ d_2 & \text{if } \left(1+\lambda\frac{2\beta-1}{1-\beta}\right)\theta \geq d_2 \end{cases} \\ t_{BI}^{PT}(\theta, d_1, d_2) &= G - \theta\left(K - d_{BI}^{PT}(\theta, d_1, d_2)\right). \end{aligned}$$

Unlike the regulation framework, the lower bound d_1 has a limited effect in this case. This is because the bureaucrat's preferred pollution level is always above that of government, and therefore, government can increase its value by restricting the bureaucrat's choice using an upper bound. The following lemmata describe the optimal choices of the bounds; they are parallel to Lemmata 4 and 5, respectively.

Lemma B.2. Fix $d_2 \in [0, K]$. Suppose government partially delegates with a restriction that $d(\underline{\theta}), d\left(\overline{\theta}\right) \in [d_1, d_2]$, for some $d_1 \in [0, d_2]$. Government's payoff is maximized at any $d_1 \leq \min\left\{d_2, d_{BI}^T(\underline{\theta})\right\}$.

Disregarding government's indifference, we assume that it puts $d_1 = \min\left\{d_2, d^T_{CI}\left(\underline{\theta}\right)\right\} \leq \min\left\{d_2, d^T_{BI}\left(\underline{\theta}\right)\right\}$.

Lemma B.3. Fix $d_1 = d^T_{CI}\left(\underline{\theta}\right)$. Suppose government partially delegates with a restriction that $d\left(\underline{\theta}\right), d\left(\overline{\theta}\right) \in [d_1, d_2]$, for some $d_2 \in [d_1, K]$. If $d^T_{BI}\left(\underline{\theta}\right) \geq d^T_{CI}\left(\overline{\theta}\right)$, then, among all $d_2 \in [d_1, K]$, government's value is maximized at $d_2 = (1-\lambda) E_\theta \theta = d^T_{CI}\left(E_\theta \theta\right)$. If $d^T_{BI}\left(\underline{\theta}\right) < d^T_{CI}\left(\overline{\theta}\right)$, then, among all $d_2 \geq d^T_{BI}\left(\underline{\theta}\right)$, C's payoff is maximized at $d_2 = d^T_{CI}\left(\overline{\theta}\right)$, while among all $d_2 < d^T_{BI}\left(\underline{\theta}\right)$, government's value is maximized at $d_2 = (1-\lambda) E_\theta \theta = d^T_{CI}\left(E_\theta \theta\right)$.

The above lemmata show that we have two possible types of regime if government partially delegates in equilibrium. With weak delegation, government chooses $d_2 = d^T_{CI}\left(\overline{\theta}\right)$. In response, the bureaucrat sets $d^{PT}_{BI}\left(\underline{\theta}, d_1, d_2\right) = \min\left\{d^T_{BI}\left(\underline{\theta}\right), d^T_{CI}\left(\overline{\theta}\right)\right\}$, and $d^{PT}_{BI}\left(\overline{\theta}, d_1, d_2\right) = d^T_{CI}\left(\overline{\theta}\right)$. Thus, government implements the full-information regulation contract if the firm is high-cost. There is distortion at the contract offered to a low-cost firm, as $d^T_{BI}\left(\underline{\theta}\right) > d^T_{CI}\left(\underline{\theta}\right)$. With strict delegation, government chooses $d_1 = d^T_{CI}\left(E_\theta \theta\right)$. In response, the bureaucrat sets $d^P_{BI}\left(\underline{\theta}, d_1, d_2\right) = d^P_{BI}\left(\overline{\theta}, d_1, d_2\right) = d_1$, resulting in a uniform pollution level for both types of firms. Note that government's choice of upper bound d_2 is always strictly below K.

Comparing government's expected value in various cases, we can observe three possible regimes in equilibrium −weak, strict, or no delegation. The following proposition fully characterizes how different regimes can arise in equilibrium.

Proposition B.2. Consider the game in which government chooses between partial delegation and no delegation. The equilibrium regime is characterized as follows:

(i) Fix ν. There exists a threshold $\lambda^{NDT} \in [0, 1)$ such that government prefers strict delegation to no delegation if and only if $\lambda \geq \lambda^{NDT}$. Morever, $\lambda^{NDT} = 0$ if $\nu < \left(\frac{K-\overline{\theta}}{\triangle\theta}\right)^2$. In addition, there exists a threshold $\beta^{DT}(\lambda)$ such that weak delegation occurs if $\beta < \beta^{DT}(\lambda)$. For $\lambda > \lambda^{NDT}$, $\beta^{DT}(\lambda)$ is decreasing in λ.

(ii) Fix λ and β. Define $\hat{\nu}^T := \max\left\{0, 1 - \left(\frac{\lambda\beta\underline{\theta}}{(1-\lambda)(1-\beta)\triangle\theta}\right)^2\right\} \in [0, 1]$. Government prefers strict delegation to weak delegation if and only if $\nu \geq \hat{\nu}^T$. For $\nu \leq \hat{\nu}^T$, there exists a threshold $\underline{\nu}^{WDT} \in \left[0, \hat{\nu}^T\right]$ such that weak delegation occurs in equilibrium if $\nu \leq \underline{\nu}^{WDT}$; and no delegation occurs in equilibrium if $\nu \in \left[\underline{\nu}^{WDT}, \hat{\nu}^T\right]$. For $\nu \geq \hat{\nu}^T$, there exists a threshold $\overline{\nu}^{SDT} \in \left[\hat{\nu}^T, 1\right]$ such that strict delegation occurs in equilibrium if $\nu \in \left[\hat{\nu}^T, \overline{\nu}^{SDT}\right]$; and no delegation occurs in equilibrium if $\nu \geq \overline{\nu}^{SDT}$.

Proof. Recalling the expression for $h(\theta, d)$ in (B11), we write government's expected payoff in the no-delegation regime as

$$E_\theta U^{NDT}_C = \nu h\left(\underline{\theta}, d^T_{CN}\left(\underline{\theta}\right)\right) + (1-\nu) h\left(\overline{\theta}, d^T_{CN}\left(\overline{\theta}\right)\right) - \nu(1-\lambda)\triangle\theta\left(K - d^T_{CN}\left(\overline{\theta}\right)\right). \quad \text{(B20)}$$

Similarly, using (B5) and (B9), we write government's expected payoff in the weak-delegation regime as

$$
\begin{aligned}
E_\theta U_C^{WDT} &= \nu \left[(1-\lambda)\, t_{BI}^T(\underline{\theta}) - \frac{1}{2} d_{BI}^T(\underline{\theta})^2\right] + (1-\nu)\left[(1-\lambda)\, t_{CI}^T(\overline{\theta}) - \frac{1}{2} d_{CI}^T(\overline{\theta})^2\right] \\
&= \nu h\left(\underline{\theta}, d_{BI}^T(\underline{\theta})\right) + (1-\nu)\, h\left(\overline{\theta}, d_{CI}^T(\overline{\theta})\right). \qquad \text{(B21)}
\end{aligned}
$$

and government's expected payoff in the strict-delegation regime as

$$
\begin{aligned}
E_\theta U_C^{SDT} &= \nu \left[(1-\lambda)\left(G - \underline{\theta}\left(K - d_{CI}^T(E_\theta\theta)\right)\right) - \frac{1}{2} d_{CI}^T(E_\theta\theta)^2\right] \\
&\quad + (1-\nu)\left[(1-\lambda)\left(G - \overline{\theta}\left(K - d_{CI}^T(E_\theta\theta)\right)\right) - \frac{1}{2} d_{CI}^T(E_\theta\theta)^2\right] \\
&= \nu h\left(\underline{\theta}, d_{CI}^T(E_\theta\theta)\right) + (1-\nu)\, h\left(\overline{\theta}, d_{CI}^T(E_\theta\theta)\right). \qquad \text{(B22)}
\end{aligned}
$$

We first compare the two regimes of weak delegation and strict delegation. We denote the difference in government's expected payoff in the two regimes by $\triangle D^{SDT-WDT}$. As government's choice of upper bound is strictly less than K, we restrict ourselves to the possibility $d_{BI}^T(\underline{\theta}) < K$.

$$
\triangle D^{SDT-WDT} = \nu\left[h\left(\underline{\theta}, d_{CI}^T(E_\theta\theta)\right) - h\left(\underline{\theta}, d_{BI}^T(\underline{\theta})\right)\right] + (1-\nu)\left[h\left(\overline{\theta}, d_{CI}^T(E_\theta\theta)\right) - h\left(\overline{\theta}, d_{CI}^T(\overline{\theta})\right)\right].
$$

Using (B12), (B21), and (B22), as well as expressions in (B5) and (B9), we have

$$
\begin{aligned}
h\left(\underline{\theta}, d_{CI}^T(E_\theta\theta)\right) - h\left(\underline{\theta}, d_{BI}^T(\underline{\theta})\right) &= \left(d_{CI}^T(E_\theta\theta) - d_{BI}^T(\underline{\theta})\right)\left((1-\lambda)\,\underline{\theta} - \frac{d_{CI}^T(E_\theta\theta) + d_{BI}^T(\underline{\theta})}{2}\right) \\
&= \frac{1}{2}\left[\frac{\lambda^2\beta^2\underline{\theta}^2}{(1-\beta)^2} - (1-\nu)^2(1-\lambda)^2(\triangle\theta)^2\right],
\end{aligned}
$$

$$
\begin{aligned}
h\left(\overline{\theta}, d_{CI}^T(E_\theta\theta)\right) - h\left(\overline{\theta}, d_{CI}^T(\overline{\theta})\right) &= \left(d_{CI}^T(E_\theta\theta) - d_{CI}^T(\overline{\theta})\right)\left((1-\lambda)\,\overline{\theta} - \frac{d_{CI}^T(E_\theta\theta) + d_{CI}^T(\overline{\theta})}{2}\right) \\
&= -\frac{1}{2}\left[\nu^2(1-\lambda)^2(\triangle\theta)^2\right],
\end{aligned}
$$

and

$$
\begin{aligned}
\triangle D^{SDT-WDT} &= \frac{\nu}{2}\left[\frac{\lambda^2\beta^2\underline{\theta}^2}{(1-\beta)^2} - (1-\nu)^2(1-\lambda)^2(\triangle\theta)^2\right] - \frac{(1-\nu)}{2}\left[\nu^2(1-\lambda)^2(\triangle\theta)^2\right] \\
&= \frac{\nu}{2}\left[\frac{\lambda^2\beta^2\underline{\theta}^2}{(1-\beta)^2} - (1-\nu)(1-\lambda)^2(\triangle\theta)^2\right]. \qquad \text{(B23)}
\end{aligned}
$$

Next, we compare the two regimes of no delegation and weak delegation, and we denote this difference by $\triangle D^{NDT-WDT}$. Making use of (B12), (B20), and (B21), as well as expressions in

(B7) and (B9), we have

$$
\begin{aligned}
&\triangle D^{NDT-WDT}\\
&= \nu\left[h\left(\underline{\theta}, d_{CN}^{T}\left(\underline{\theta}\right)\right) - h\left(\underline{\theta}, d_{BI}^{T}\left(\underline{\theta}\right)\right)\right]\\
&+ (1-\nu)\left[h\left(\overline{\theta}, d_{CN}^{T}\left(\overline{\theta}\right)\right) - h\left(\overline{\theta}, d_{CI}^{T}\left(\overline{\theta}\right)\right)\right] - \nu\left(1-\lambda\right)\triangle\theta\left(K - d_{CN}^{T}\left(\overline{\theta}\right)\right)\\
&= \nu\left[\left(d_{CN}^{T}\left(\underline{\theta}\right) - d_{BI}^{T}\left(\underline{\theta}\right)\right)\left((1-\lambda)\underline{\theta} - \frac{d_{CN}^{T}\left(\underline{\theta}\right) + d_{BI}^{T}\left(\underline{\theta}\right)}{2}\right)\right]\\
&+ (1-\nu)\left[\left(d_{CN}^{T}\left(\overline{\theta}\right) - d_{CI}^{T}\left(\overline{\theta}\right)\right)\left((1-\lambda)\overline{\theta} - \frac{d_{CN}^{T}\left(\overline{\theta}\right) + d_{CI}^{T}\left(\overline{\theta}\right)}{2}\right)\right] - \nu\left(1-\lambda\right)\triangle\theta\left(K - d_{CN}^{T}\left(\overline{\theta}\right)\right)\\
&= \frac{\nu\lambda^2\beta^2\underline{\theta}^2}{2\left(1-\beta\right)^2} - \frac{1-\nu}{2}\left(d_{CN}^{T}\left(\overline{\theta}\right) - d_{CI}^{T}\left(\overline{\theta}\right)\right)^2 - \nu\left(1-\lambda\right)\triangle\theta\left(K - d_{CN}^{T}\left(\overline{\theta}\right)\right)\\
&= \begin{cases} \frac{\nu}{2}\left[\frac{\lambda^2\beta^2\underline{\theta}^2}{(1-\beta)^2} + \frac{\nu(1-\lambda)^2(\triangle\theta)^2}{(1-\nu)} - 2\left(1-\lambda\right)\triangle\theta\left(K - \left(1-\lambda\right)\overline{\theta}\right)\right] & \text{if } d_{CN}\left(\overline{\theta}\right) < K \\ \\ \frac{\nu\lambda^2\beta^2\underline{\theta}^2}{2(1-\beta)^2} - \frac{1-\nu}{2}\left(K - \left(1-\lambda\right)\overline{\theta}\right)^2 & \text{if } d_{CN}\left(\overline{\theta}\right) = K \end{cases}. \qquad \text{(B24)}
\end{aligned}
$$

Finally, the difference between the two regimes of no delegation and strict delegation is denoted by $\triangle D^{NDT-SDT}$. Making use of (B12), (B20), and (B22), as well as (B5) and (B7), we have

$$
\begin{aligned}
&\triangle D^{NDT-SDT}\\
&= \nu\left[h\left(\underline{\theta}, d_{CN}^{T}\left(\underline{\theta}\right)\right) - h\left(\underline{\theta}, d_{CI}^{T}\left(E_{\theta}\theta\right)\right)\right]\\
&+ (1-\nu)\left[h\left(\overline{\theta}, d_{CN}^{T}\left(\overline{\theta}\right)\right) - h\left(\overline{\theta}, d_{CI}^{T}\left(E_{\theta}\theta\right)\right)\right] - \nu\left(1-\lambda\right)\triangle\theta\left(K - d_{CN}^{T}\left(\overline{\theta}\right)\right)\\
&= \nu\left[\left(d_{CN}^{T}\left(\underline{\theta}\right) - d_{CI}^{T}\left(E_{\theta}\theta\right)\right)\left((1-\lambda)\underline{\theta} - \frac{d_{CN}^{T}\left(\underline{\theta}\right) + d_{CI}^{T}\left(E_{\theta}\theta\right)}{2}\right)\right]\\
&+ (1-\nu)\left[\left(d_{CN}^{T}\left(\overline{\theta}\right) - d_{CI}^{T}\left(E_{\theta}\theta\right)\right)\left((1-\lambda)\overline{\theta} - \frac{d_{CN}^{T}\left(\overline{\theta}\right) + d_{CI}^{T}\left(E_{\theta}\theta\right)}{2}\right)\right]\\
&- \nu\left(1-\lambda\right)\triangle\theta\left(K - d_{CN}^{T}\left(\overline{\theta}\right)\right)\\
&= \nu\left[\frac{\left(1-\lambda\right)^2\left(1-\nu\right)^2\left(\triangle\theta\right)^2}{2}\right] - \nu\left(1-\lambda\right)\triangle\theta\left(K - d_{CN}^{T}\left(\overline{\theta}\right)\right)\\
&+ \frac{(1-\nu)}{2}\left[\left(d_{CN}^{T}\left(\overline{\theta}\right) - (1-\lambda)\left(\overline{\theta} - \nu\triangle\theta\right)\right)\left(2\left(1-\lambda\right)\overline{\theta} - d_{CN}^{T}\left(\overline{\theta}\right) - (1-\lambda)\left(\overline{\theta} - \nu\triangle\theta\right)\right)\right]\\
&= \nu\left[\frac{\left(1-\lambda\right)^2\left(1-\nu\right)^2\left(\triangle\theta\right)^2}{2}\right] - \nu\left(1-\lambda\right)\triangle\theta\left(K - d_{CN}^{T}\left(\overline{\theta}\right)\right)\\
&+ \frac{(1-\nu)}{2}\left[\left(d_{CN}^{T}\left(\overline{\theta}\right) - (1-\lambda)\overline{\theta} + \nu\left(1-\lambda\right)\triangle\theta\right)\left((1-\lambda)\overline{\theta} - d_{CN}^{T}\left(\overline{\theta}\right) + \nu\left(1-\lambda\right)\triangle\theta\right)\right]
\end{aligned}
$$

$$= \nu \left[\frac{(1-\lambda)^2 (1-\nu)^2 (\triangle\theta)^2}{2} \right] - \nu (1-\lambda) \triangle\theta \left(K - d_{CN}^T \left(\overline{\theta} \right) \right)$$

$$+ \frac{(1-\nu)}{2} \left[\nu^2 (1-\lambda)^2 (\triangle\theta)^2 - \left(d_{CN}^T \left(\overline{\theta} \right) - (1-\lambda) \overline{\theta} \right)^2 \right]$$

$$= \frac{\nu (1-\nu) (1-\lambda)^2 (\triangle\theta)^2}{2} [(1-\nu) + \nu]$$

$$- \left[\nu (1-\lambda) \triangle\theta \left(K - d_{CN}^T \left(\overline{\theta} \right) \right) - \frac{(1-\nu)}{2} \left(d_{CN}^T \left(\overline{\theta} \right) - (1-\lambda) \overline{\theta} \right)^2 \right]$$

$$= \frac{\nu (1-\nu) (1-\lambda)^2 (\triangle\theta)^2}{2} - \frac{(1-\lambda)^2 (1-\nu)}{2} \left[\frac{2\nu\triangle\theta \left(K - d_{CN}^T \left(\overline{\theta} \right) \right)}{(1-\nu)(1-\lambda)} + \left(\frac{d_{CN}^T \left(\overline{\theta} \right)}{(1-\lambda)} - \overline{\theta} \right)^2 \right].$$

Suppose that $d_{CN} \left(\overline{\theta} \right) < K$. Then,

$$\triangle D^{NDT-SDT}$$

$$= \frac{\nu (1-\nu) (1-\lambda)^2 (\triangle\theta)^2}{2} - \frac{(1-\lambda)^2 (1-\nu)}{2} \left[\frac{2\nu\triangle\theta \left(K - d_{CN}^T \left(\overline{\theta} \right) \right)}{(1-\nu)(1-\lambda)} + \left(\frac{d_{CN}^T \left(\overline{\theta} \right)}{(1-\lambda)} - \overline{\theta} \right)^2 \right]$$

$$= \frac{\nu (1-\nu) (1-\lambda)^2 (\triangle\theta)^2}{2} - \frac{(1-\lambda)^2 (1-\nu)}{2} \left[\frac{2\nu\triangle\theta}{(1-\nu)} \left(\frac{K}{1-\lambda} - \overline{\theta} - \frac{\nu (\triangle\theta)}{(1-\nu)} \right) + \frac{\nu^2 (\triangle\theta)^2}{(1-\nu)^2} \right]$$

$$= \frac{\nu (1-\nu) (1-\lambda)^2 (\triangle\theta)^2}{2} - \frac{(1-\lambda)^2 (1-\nu)}{2} \left[\frac{2\nu\triangle\theta}{(1-\nu)} \left(\frac{K}{1-\lambda} - \overline{\theta} \right) - \frac{\nu^2 (\triangle\theta)^2}{(1-\nu)^2} \right]$$

$$= \frac{\nu (1-\nu) (1-\lambda)^2 (\triangle\theta)^2}{2} + \frac{\nu^2 (1-\lambda)^2 (\triangle\theta)^2}{2(1-\nu)} - (1-\lambda)^2 \nu\triangle\theta \left(\frac{K}{1-\lambda} - \overline{\theta} \right)$$

$$= \frac{\nu (1-\lambda)^2 (\triangle\theta)^2}{2(1-\nu)} \left[(1-\nu)^2 + \nu \right] - (1-\lambda)^2 \nu\triangle\theta \left(\frac{K}{1-\lambda} - \overline{\theta} \right)$$

$$= (1-\lambda)^2 \nu\triangle\theta \left[\frac{\left(1-\nu+\nu^2\right)(\triangle\theta)}{2(1-\nu)} - \left(\frac{K}{1-\lambda} - \overline{\theta} \right) \right].$$

Suppose that $d_{CN} \left(\overline{\theta} \right) = K$. Then,

$$\triangle D^{NDT-SDT}$$

$$= \frac{\nu (1-\nu) (1-\lambda)^2 (\triangle\theta)^2}{2} - \frac{(1-\lambda)^2 (1-\nu)}{2} \left[\frac{2\nu\triangle\theta \left(K - d_{CN}^T \left(\overline{\theta} \right) \right)}{(1-\nu)(1-\lambda)} + \left(\frac{d_{CN}^T \left(\overline{\theta} \right)}{(1-\lambda)} - \overline{\theta} \right)^2 \right]$$

$$= \frac{\nu (1-\nu) (1-\lambda)^2 (\triangle\theta)^2}{2} - \frac{(1-\lambda)^2 (1-\nu)}{2} \left(\frac{K}{1-\lambda} - \overline{\theta} \right)^2$$

$$= \frac{(1-\nu)(1-\lambda)^2}{2}\left[\nu(\triangle\theta)^2 - \left(\frac{K}{1-\lambda} - \overline{\theta}\right)^2\right].$$

Together, this means

$$\triangle D^{NDT-SDT} = \begin{cases} (1-\lambda)^2 \nu \triangle\theta \left[\frac{(1-\nu+\nu^2)(\triangle\theta)}{2(1-\nu)} - \left(\frac{K}{1-\lambda} - \overline{\theta}\right)\right] & \text{if } d_{CN}\left(\overline{\theta}\right) < K \\ \\ \frac{(1-\nu)(1-\lambda)^2}{2}\left[\nu(\triangle\theta)^2 - \left(\frac{K}{1-\lambda} - \overline{\theta}\right)^2\right] & \text{if } d_{CN}\left(\overline{\theta}\right) = K \end{cases}. \tag{B25}$$

Consider the effects of λ and β. For a fixed ν, we can write $\triangle D^{NDT-SDT}$ in (B25) as $(1-\lambda)^2 A^T(\lambda)$ where $A^T(\lambda)$ is continuous. We define $\lambda^{NDT} := \max\left\{0, \lambda_1^{NDT}\right\}$ where λ_1^{NDT} solves $A^T(\lambda) = 0$. Since $\frac{dA^T(\lambda)}{d\lambda} < 0$, government prefers strict delegation to no delegation if and only if $\lambda \geq \lambda^{NDT}$. It follows that $\lambda^{NDT} = 0$ if $\nu < \left(\frac{K-\overline{\theta}}{\triangle\theta}\right)^2$. λ^{NDT} does not vary with β, as $\triangle D^{NDT-SDT}$ is independent of β.

For $\lambda \geq \lambda^{NDT}$, either strict delegation or weak delegation occurs in equilibrium. We can write $\triangle D^{SDT-WDT}$ as $(1-\lambda)^2 B^T(\lambda,\beta)$, where $B^T(\lambda,\beta)$ is continuous and increasing in β. Therefore, we can find $\beta_1^{DT}(\lambda)$, which solves $B^T(\lambda,\beta) = 0$ for $\lambda \geq \lambda^{NDT}$, implying that $\triangle D^{SDT-WDT} \geq 0$ if $\beta \geq \beta_1^{DT}(\lambda)$. As $B^T(\lambda,\beta)$ is increasing in λ, $\beta_1^{DT}(\lambda)$ is decreasing in λ.

For $\lambda \leq \lambda^{NDT}$, either no delegation or weak delegation occurs in equilibrium. We can write $\triangle D^{NDT-WDT}$ as $(1-\lambda)^2 C^T(\lambda,\beta)$, where $C^T(\lambda,\beta)$ is continuous and increasing in β. Therefore, we can find $\beta_2^{DT}(\lambda)$, which solves $C^T(\lambda,\beta) = 0$ for $\lambda \leq \lambda^{NDT}$, implying that $\triangle D^{NDT-WDT} \geq 0$ if $\beta \geq \beta_2^{DT}(\lambda)$. Therefore, weak delegation occurs in equilibrium if

$$\beta < \beta^{DT}(\lambda) := \begin{cases} \beta_1^{DT}(\lambda) & \text{if } \lambda \geq \lambda^{NDT} \\ \beta_2^{DT}(\lambda) & \text{if } \lambda \leq \lambda^{NDT} \end{cases}. \tag{B26}$$

To find the effect of ν, fix λ and β. From (B23), note that $\triangle D^{SDT-WDT} > 0$ if and only $\nu > 1 - \left(\frac{\lambda\beta\underline{\theta}}{(1-\lambda)(1-\beta)\triangle\theta}\right)^2$. As ν is bounded below by 0, we define the threshold value $\hat{\nu}^T := \max\left\{0, 1 - \left(\frac{\lambda\beta\underline{\theta}}{(1-\lambda)(1-\beta)\triangle\theta}\right)^2\right\}$. Between the two partial-delegation rules, government prefers strict delegation to weak delegation if and only if $\nu > \hat{\nu}^T$. Consequently, for $\nu > \hat{\nu}^T$, strict delegation or no delegation can occur in equilibrium, while for $\nu \leq \hat{\nu}^T$, only weak delegation or no delegation can occur in equilibrium. We will consider these two cases separately. To find out the equilibrium outcome in these two cases, we need to study the effect of ν on $\triangle D^{NDT-SDT}$ and $\triangle D^{NDT-WDT}$. Recall the definition of ν^T, in (B18), and that $d_{CN}^T\left(\overline{\theta}\right) < K$ if and only if $\nu < \nu^T$. Note that $\triangle D^{NDT-SDT}$ and $\triangle D^{NDT-WDT}$ are continuous in ν but not necessarily differentiable at $d^T_{CN}\left(\overline{\theta}\right)=K$, or equivalently, at $\nu=\nu^T$, and that ν^T can lie below or above $\hat{\nu}^T$, depending on other parameter values.

First, consider $\nu\geq\hat{\nu}^T$. If now $\triangle D^{NDT-SDT}\geq 0$, then no delegation occurs in equilibrium; otherwise, strict delegation occurs in equilibrium. We first study the expression of $\triangle D^{NDT-SDT}$ in (B25) for all ν in $[0,1]$. Then, we will consider the constraint that $\nu\geq\hat{\nu}^T$.

The following two observations are useful in studying the effect of ν on $\triangle D^{NDT-SDT}$. First, $\triangle D^{NDT-SDT}=0$ at both $\nu=0$ and $\nu=1$. Secondly, $\frac{\partial^2}{\partial\nu^2}\left(\triangle D^{NDT-SDT}\right)=\frac{\nu\left(3-3\nu+\nu^2\right)(1-\lambda)^2(\triangle\theta)^2}{(1-\nu)^3}>0$ for $\nu<\nu^T$, and $\frac{\partial^2}{\partial\nu^2}\left(\triangle D^{NDT-SDT}\right)=-\left(1-\lambda\right)^2\left(\triangle\theta\right)^2<0$ for $\nu>\nu^T$. Therefore, $\triangle D^{NDT-SDT}$ is convex in ν for $\nu<\nu^T$, and concave in ν for $\nu>\nu^T$. If $\triangle D^{NDT-SDT}\leq 0$ at $\nu=\nu^T$, then the above two observations together imply that, either $\triangle D^{NDT-SDT}<0$ for every $0<\nu<1$, or $\triangle D^{NDT-SDT}$ has exactly one zero at some $\overline{\nu}^{kT}\in\left(\nu^T,1\right)$, such that $\triangle D^{NDT-SDT}<0$ for every $0<\nu<\overline{\nu}^{kT}$, and $\triangle D^{NDT-SDT}>0$ for every $\overline{\nu}^{kT}<\nu<1$. In contrast, if $\triangle D^{NDT-SDT}>0$ at $\nu=\nu^T$, then the two observations together imply that, either $\triangle D^{NDT-SDT}>0$ for every $0<\nu<1$, or $\triangle D^{NDT-SDT}$ has exactly one zero at some $\overline{\nu}^{\ell T}\in\left(0,\nu^T\right)$, such that $\triangle D^{NDT-SDT}<0$ for every $0<\nu<\overline{\nu}^{\ell T}$, and $\triangle D^{NDT-SDT}>0$ for every $\overline{\nu}^{\ell T}<\nu<1$. While $\triangle D^{NDT-SDT}$ can take either sign at $\nu=\nu^T$, we see that the following three possibilities are the only ones that can arise: (a) $\triangle D^{NDT-SDT}>0$ for $\nu\in(0,1)$; (b) $\triangle D^{NDT-SDT}<0$ for $\nu\in(0,1)$; and (c) $\triangle D^{NDT-SDT}<0$ for $\nu\in\left(0,\overline{\nu}^{mT}\right)$ and $\triangle D^{NDT-SDT}>0$ for $\nu\in\left(\overline{\nu}^{mT},1\right)$, where $0<\overline{\nu}^{mT}<1$, and $\overline{\nu}^{mT}\in\left\{\overline{\nu}^{kT},\overline{\nu}^{\ell T}\right\}$. Now we adjust to the constraint $\nu\geq\hat{\nu}^T$, and define the threshold $\overline{\nu}^{SDT}$ as $\hat{\nu}^T$ in scenario (a), as 1 in scenario (b), and as $\max\left\{\hat{\nu}^T,\overline{\nu}^{mT}\right\}$ in scenario (c). Therefore, $\overline{\nu}^{SDT}\in\left[\hat{\nu}^T,1\right]$. Furthermore, strict delegation occurs in equilibrium if $\hat{\nu}^T\leq\nu\leq\overline{\nu}^{SDT}$; and no delegation occurs in equilibrium if $\overline{\nu}^{SDT}\leq\nu<1$.

Next, consider $\nu\leq\hat{\nu}^T$. If now $\triangle D^{NDT-WDT}>0$, then no delegation occurs in equilibrium; otherwise, weak delegation occurs in equilibrium. As before, we will first study the expression of $\triangle D^{NDT-WDT}$ in (B24) for all ν in $[0,1]$ and then adjust for the constraint that $\nu\leq\hat{\nu}^T$.

The following two observations are useful in this case. First, $\triangle D^{NDT-WDT}=0$ at $\nu=0$, and $\triangle D^{NDT-WDT}>0$ at $\nu=1$. Secondly, $\frac{\partial^2}{\partial\nu^2}\left(\triangle D^{NDT-WDT}\right)=\frac{(1-\lambda)^2(\triangle\theta)^2}{(1-\nu)^3}>0$ for $\nu<\nu^T$, and $\frac{\partial^2}{\partial\nu^2}\left(\triangle D^{NDT-WDT}\right)=0$ for $\nu>\nu^T$. Therefore, $\triangle D^{NDT-WDT}$ is convex in ν for $\nu<\nu^T$ and linear in ν for $\nu\geq\nu^T$. If $\triangle D^{NDT-WDT}\leq 0$ at $\nu=\nu^T$, then the above two observations together imply that $\triangle D^{NDT-WDT}$ has exactly one zero at some $\underline{\nu}^{rT}\in\left[\nu^T,1\right)$, such that $\triangle D^{NDT-WDT}\leq 0$ for every $0<\nu\leq\underline{\nu}^{rT}$, and $\triangle D^{NDT-WDT}>0$ for every $\underline{\nu}^{rT}<\nu<1$. In contrast, if $\triangle D^{NDT-WDT}>0$ at $\nu=\nu^T$, then the two observations together imply that, either $\triangle D^{NDT-WDT}>0$ for every $0<\nu<1$, or $\triangle D^{NDT-WDT}$ can have exactly one zero at $\underline{\nu}^{sT}\in\left(0,\nu^T\right)$, such that $\triangle D^{NDT-WDT}<0$ if and only if $0<\nu<\underline{\nu}^{sT}$. Therefore, while $\triangle D^{NDT-WDT}$ can take either sign at $\nu=\nu^T$, the following two possibilities are the only ones that can arise: (a) $\triangle D^{NDT-WDT}>0$ for all $\nu\in(0,1)$; and (b) $\triangle D^{NDT-WDT}<0$ for $\nu\in\left(0,\underline{\nu}^{tT}\right)$ and $\triangle D^{NDT-WDT}>0$ for $\nu\in\left(\underline{\nu}^{tT},1\right)$, where $0<\underline{\nu}^{tT}<1$, and $\underline{\nu}^{tT}\in\left\{\underline{\nu}^{rT},\underline{\nu}^{sT}\right\}$.

Adjusting for the constraint $\nu \leq \hat{\nu}^T$, we define the threshold $\underline{\nu}^{WDT}$ as 0 in scenario (a); and as $\min\left\{\underline{\nu}^{tT}, \hat{\nu}^T\right\}$ in scenario (b). Clearly, $0 < \underline{\nu}^{WDT} \leq \hat{\nu}^T$. Furthermore, weak delegation occurs in equilibrium if $0 \leq \nu \leq \underline{\nu}^{WDT}$, while no delegation occurs in equilibrium if $\underline{\nu}^{WDT} \leq \nu \leq \hat{\nu}^T$. $\square$

References

[1] Alonso, R. and N. Matouschek (2009). "Optimal Delegation." *Review of Economic Studies* 75, 259-293.

[2] Amador, M. and K. Bagwell (2013). "The Theory of Optimal Delegation with an Application to Tariff Caps." *Econometrica* 81, 1541-1599.

[3] Baron, D.P. (1989). "Design of Regulatory Mechanisms and Institutions." In *Handbook of Industrial Organization*, vol 2 (R. Schmalensee and R.D. Willig, eds.), pp. 1347-1447.

[4] Bendor, J., A. Glazer, and T. Hammond (2001). "Theories of Delegation." *Annual Review of Political Science* 4(1), 235-269.

[5] Boehmke, F.J., S. Gailmard, and Patty, J.W. (2006). "Whose Ear to Bend? Information Sources and Venue Choice in Policy-Making." *Quarterly Journal of Political Science* 1, 139-169.

[6] Boyer, M. and J.-J. Laffont (1999). "Toward a Political Theory of the Emergence of Environmental Incentive Regulation." *RAND Journal of Economics* 30, 137-157.

[7] Callander, S. and K. Krehbiel (2014). "Gridlock and Delegation in a Changing World." *American Journal of Political Science* 58, 819-834.

[8] de Figueiredo, R.J. (2002). "Electoral Competition, Political Uncertainty, and Policy Insulation." *American Political Science Review* 96, 321-333.

[9] Duflo, E., M. Greenstone, R. Pande, and N. Ryan (2018). "The Value of Regulatory Discretion: Estimates from Environmental Inspections in India." *Econometrica*, forthcoming.

[10] Epstein, D. and S. O'Halloran (1994). "Administrative Procedures, Information, and Agency Discretion." *American Journal of Political Science* 38, 697-722.

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

[12] Evans, J., P. Levine, and F. Trillas (2008). "Lobbies, Delegation and the Under-Investment Problem in Regulation." *International Journal of Industrial Organization* 26, 17-40.

[13] Frankel, A. (2014). "Aligned Delegation." *American Economic Review* 104, 66-83.

[14] Frankel, A. (2016). "Delegating Multiple Decisions." *American Economic Journal: Microeconomics* 8, 16-53.

[15] Frankel, A. (2017). "Selecting Applicants." Unpublished manuscript, University of Chicago.

[16] Gailmard, S. (2009). "Discretion Rather Than Rules: Choice of Instruments to Control Bureaucratic Policy Making." *Political Analysis* 17, 25-44.

[17] Gailmard, S. and J.W. Patty (2007). "Slackers and Zealots: Civil Service, Policy Discretion, and Bureaucratic Expertise." *American Journal of Political Science* 51, 873-889.

[18] Gailmard, S. and J.W. Patty (2012a). "Formal Models of Bureaucracy." *Annual Review of Political Science* 15, 353-377.

[19] Gailmard, S., and J.W. Patty (2012b). *Learning While Governing: Expertise and Accountability in the Executive Branch*. University of Chicago Press.

[20] Gilardi, F. (1989). *Delegation in the Regulatory State: Independent Regulatory Agencies in Western Europe*. Edward Elgar Publishing.

[21] Hiriart, Y. and D. Martimort (2012). "How Much Discretion for Risk Regulators?" *RAND Journal of Economics* 43, 283-314.

[22] Holmström, B. (1984). "On the Theory of Delegation." In *Bayesian Models in Economic Theory* (M. Boyer and R.E. Kihlstrom, eds.), Elsevier, pp. 115-141.

[23] Huber, J.D. and N. McCarty (2004). "Bureaucratic Capacity, Delegation, and Political Reform." *American Political Science Review* 98, 481-494.

[24] Huber, J.D. and C.R. Shipan (2006). "Politics, Delegation, and Bureaucracy." In *Oxford Handbook of Political Economy* (D.A. Ritchie and B.R. Weingast, eds.), Oxford University Press, pp. 256-272.

[25] Khalil, F., D. Kim, and J. Lawarrée (2013). "Contracts Offered by Bureaucrats." *RAND Journal of Economics* 44, 686-711.

[26] Koessler, F. and D. Martimort (2012). "Optimal Delegation with Multi-Dimensional Decisions." *Journal of Economic Theory* 147, 1850-1881.

[27] Laffont, J.-J. (1994). "Regulation of Pollution with Asymmetric Information." In *Nonpoint Source Pollution Regulation: Issues and Analysis* (C. Dosi and T. Tomasi, eds.), Springer, pp. 39-66.

[28] Laffont, J.-J. and D. Martimort (1999). "Separation of Regulators against Collusive Behavior." *RAND Journal of Economics* 30, 232-262.

[29] Laffont, J.-J. and J. Tirole (1991). "The Politics of Government Decision-Making: A Theory of Regulatory Capture." *Quarterly Journal of Economics* 106, 1089-1127.

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

[31] Letina, I., S. Liu, and N. Netzer (2018). "Delegating Performance Evaluation." Unpublished manuscript, Universities of Bern and Zurich.

[32] Levine, P., J. Stern, and F. Trillas (2005). "Utility Price Regulation and Time Inconsistency: Comparisons with Monetary Policy." *Oxford Economic Papers* 57, 447-478.

[33] Lowande, K. (2018). "Delegation or Unilateral Action?" *Journal of Law, Economics, and Organization* 34, 54-78.

[34] Melumad, N.D. and T. Shibano (1991). "Communication in Settings with No Transfers." *RAND Journal of Economics* 22, 173-198.

[35] McCarty, N. (2017). "The Regulation and Self-Regulation of a Complex Industry." *The Journal of Politics* 79, 1220-1236.

[36] Moe, T.M. (2013). "Delegation, Control, and the Study of Public Bureaucracy." In *Handbook of Organizational Economics* (R. Gibbons and J. Roberts, eds.), Princeton University Press, pp. 1148-1181.

[37] Montagnes, B.P. and S. Wolton (2017). "Rule versus Discretion: Regulatory Uncertainty, Firm Investment, and Bureaucratic Organization." *The Journal of Politics* 79, 457-472.

[38] Spulber, D.F. and D. Besanko (1992). "Delegation, Commitment, and the Regulatory Mandate." *Journal of Law, Economics, and Organization* 8, 126-154.

[39] Strausz, R. (2017). "Politically Induced Regulatory Risk and Independent Regulatory Agencies." *International Journal of Industrial Organization* 54, 215-238.

[40] Volden, C. (2002). "A Formal Model of the Politics of Delegation in a Separation of Powers System." *American Journal of Political Science* 46, 111-133.