

A THEORY OF (DE)CENTRALIZATION *

Peter Klibanoff[†] and Michel Poitevin[‡]

April 2013

Abstract

This paper compares the efficacy of a centralized and a decentralized rights structure in determining the size of an externality generating project. Consider a central authority and two localities. One locality can operate a variable-size project which produces an externality that affects the other locality. Each locality may have some private information concerning its own net benefit from the project. Under centralization, localities are vertically integrated with a benevolent central authority who effectively possesses all property rights. Under decentralization, localities are separate legal entities (endowed with property rights) who bargain to determine the project size. We examine the performance of these two regimes and show how one or the other may dominate depending on the distributions of private and external benefits from the project. The effect of the size and variation in the externality on this trade-off is of particular interest.

JEL: Organizational Behavior; Transaction Costs; Property Rights; Externalities; Asymmetric and Private Information; Structure, Scope, and Performance of Government (D23, D62, D82, H11).

*We thank Robin Boadway, Faruk Gul, Lu Hu, Eric Maskin, Dilip Mookherjee, Stefan Reichelstein, Mike Peters, Patrick Rey, Jacques Robert, Lars Stole and François Vaillancourt for comments and thank a number of seminar and conference audiences. The first version of this paper was written while the second author was visiting the MEDS Department at Northwestern University. He would like to thank MEDS for its hospitality and support during his visit, and also C.I.R.A.N.O., C.R.S.H. and F.C.A.R. for their financial support.

[†]MEDS, Northwestern University

[‡]Sciences économiques, Université de Montréal

1 Introduction

There are at least two classic approaches to the problem of externalities in the economic literature. One, Pigouvian taxation, solves the problem through centrally imposed taxes/subsidies on production. The Pigouvian solution specifies that a central authority imposes a tax, a subsidy, a quota, or a standard that must be obeyed by the agents. In a frictionless world with a benevolent central authority, this regulation leads to an efficient outcome.

Another, Coasian bargaining [Coase, 1960] offers a decentralized solution. The Coase Theorem states that, in the absence of transaction costs, the central authority only has to assign and/or enforce property rights of the concerned agents, and bargaining between the agents will generate an efficient outcome.

Both these approaches lead to efficiency when there are no market imperfections of any sort. If there are imperfections, however, the comparison between these two modes of regulation becomes more complicated. For example, if the central authority is imperfectly informed about the social costs and benefits of the project, it has to extract this information from the informed agents before putting in place its regulatory scheme. Similarly, if agents are asymmetrically informed, this will affect the outcome of bargaining between them. Since the two approaches may behave quite differently in the presence of asymmetric information, the problem of choosing the better one is not trivial.

The main difference between these two approaches is in the attribution or not of property rights to the regulated agents. Under a centralized scheme, property rights are retained by the center who imposes a solution on the agents. Under a decentralized scheme, agents are endowed with property rights which they can trade or bargain with.

An example of a centralized rights structure is the portrait of a centralized regime as depicted by De Long and Shleifer [1993] who study the impact of centralization of power on economic growth in European cities between 1050 and 1800. They define a centralized (absolutist) regime as one where:

“Subjects have no rights; they have privileges, which endure only as long as the prince wishes.”

In such a setting there are no enforceable agreements or bargains. The central authority can always break any promise. In a more modern example, consider two localities

that have been merged or integrated and are now under the authority of a central entity. The merged localities have no say in the central authority's decision making outside the communication channels and committees that may have been instituted. In this case, the central authority may consult local officials but it retains all decision making power.

In contrast, in a decentralized setting, regulated agents are attributed rights and therefore have some scope for independent action. In the face of an externality on one agent generated by the actions of another, it is natural to suppose that these agents will bargain to try to internalize the externality. Because agents have rights and these rights can be traded, enforceable agreements and transfers are possible in this decentralized structure. Again, our view of decentralization finds expression in De Long and Shleifer [1993]. They argue that, under decentralized (non-absolutist) regimes,

"...the legal framework was, not an instrument of the prince's rule, but more of a semifeudal contract between different powers establishing the framework of their interactions. (...) Taxes could be raised only with the consent of feudal estates."

Given a choice, would we expect a centralized or a decentralized regime to cope better with an externality in the presence of asymmetric information? Our model analyzes this problem by comparing a centralized/integrated Pigouvian setting, where no property rights exist and where centrally imposed quotas dictate the allocation of resources, as opposed to a decentralized Coasian environment, where local agents have property rights and bargain to determine the allocation of resources.

We now informally describe our model. The problem is to determine the proper size of a project affecting the welfare of two agents. We cast this problem in an environment where two neighboring localities are affected by a project. Each locality is privately informed of the benefit (harm) the project will provide to that locality. In a centralized setting, the two localities are merged and thus have no rights in bargaining with the central authority. A benevolent, but uninformed, central authority can impose any project size on the localities. The absence of property rights can be formalized by saying the two localities have no participation constraints that the central authority must respect. The information of the localities is incorporated in the decision process only through lobbying or informal communication. Such lobbying is akin to cheap talk and credibility becomes an issue which has to be solved by the central authority.

In a decentralized setting, the two localities are legal entities and thus possess rights concerning the size of the project and taxation. Specifically, one locality owns the right to carry out the project and to determine its size and each locality has the right to refuse taxation. The central authority therefore has no means of imposing a project size. The attribution of property rights can be formalized by the introduction of participation constraints for the localities. These constraints reflect their control over productive activity and right to refuse involuntary taxation. The decision about the project is by the localities through bargaining about project size and transfers subject to these participation constraints.

Without further assumptions, the fact that participation constraints (rights) are present under decentralization but not under centralization leads a welfare optimizing approach to always favor (at least weakly) centralization. Thus there would be no scope for a theory of decentralization versus centralization in dealing with externalities. To develop such a theory, we make one crucial assumption. We assume that there is a (small) social cost to taxation. This assumption is often found in the literature on public economics and may be based on inefficiencies resulting from the distortionary effects of taxes.

A critical difference between centralization and decentralization in our model is the existence and the allocation of property rights. Under centralization, localities have no rights in dealing with the central authority. Localities effectively become internal divisions of the central authority's organization. They are vertically integrated and they are an integral part of the legal structure of the central authority. Within this organization or legal structure, divisions have no property or contracting rights. In effect, no legal contract can be enforced between the localities and the central authority.

We note that this view of centralization in which the central authority has all property rights has some empirical relevance. In the recent financial crisis, the U.S. government devised relief programs to help troubled financial firms. Partnerships were set up between the government and financial firms. Chris Low, chief economist at FTN Financial, said: *"You'd have to be crazy as a big investor to go into a partnership with the government right now, because as we've seen with TARP (Troubled Asset Relief Program) and AIG the government changes terms when they don't like how things work out."* Michael Feroli, economist at JP Morgan Chase in New York, added: *"There's some concern that the government, being the government, can change the rules that they want, for example impose executive compensation*

*restrictions ex-post.*¹

The alternative is to allocate property rights to localities. Property rights give autonomy and legal means for signing and enforcing contracts. Localities are outside the central authority's organization. Localities can thus bargain and trade those rights under an enforceable legal framework. We call this regime decentralization.

We thus look at two polar cases for the endogenous distribution of property rights. Under centralization, localities are vertically integrated and have no property rights. Under decentralization, localities are spun off and they have full property rights. The bargaining and trading opportunities differ significantly across the two regimes.

One application of our model is to evaluate some "folk wisdom" about externalities and government control. Three statements often made (see e.g., Oates [1972]) are: (1) large externalities justify central control or regulation; (2) heterogeneity in localities' characteristics favors decentralization; and (3) centralized policies tend to be insensitive to the preferences of localities or regions.

Complete results are reported below, but the flavor of our findings as they relate to the above statements is given here. First, increases in the (average) size of the externality do not always favor centralization. Whether this is true or not depends on how large the expected externality is to begin with compared to the possible variation in the private benefit level of the project. Second, our results confirm that ex ante heterogeneity (variance) in the size of the externality favors decentralization. However, ex ante heterogeneity in the size of the private benefit of the project may favor either centralization or decentralization. Finally, we do find that, for a wide range of cases, optimal centralized policies specify a uniform project size independent of localities' ex post realized benefits and costs. Note that this uniformity is not assumed, as in many other models, but is derived as a result. The basic intuition for these results comes from the fact that decentralization is good at incorporating (through bargaining) variation in the externality, while both centralization and decentralization imperfectly accommodate variation in the private benefit. Centralization does this imperfectly because of the inability of the central authority to sign contracts binding its future behavior that would facilitate the elicitation of private information. Decentralization is imperfect because individual rationality

¹These citations were taken from "AIG debacle chills investor interest in bailout plans," by Kristina Cooke and Jennifer Ablan, Reuters, Wednesday, March 18, 2009 (www.reuters.com/article/2009/03/18/us-financial-aig-investors-idUSTRE52H4FI20090318).

constraints create a trade-off between incentives and informational rents.

Before going to the formal analysis, we mention the following related literature. There are a number of papers that examine the problem of externalities in asymmetric information environments (see, for example, Farrell, 1987; Rob, 1989; Klibanoff and Morduch, 1995). This strand of literature adopts a mechanism design approach and emphasizes the crucial role of individual rationality (or participation) constraints in hindering efficient solutions as pointed out by Laffont and Maskin [1979] and Myerson and Satterthwaite [1983].² These papers are interested in characterizing allocations in a setting where agents have individual-rationality constraints, and therefore cannot have a project imposed upon them by the higher-authority principal. Thus, in our language, these papers all examine variations on decentralized environments. Their underlying theme can be characterized as inefficiencies caused by the trade-off between incentives and informational rents.

Less related to externalities, but quite related to our formal modeling of centralization, are models of communication such as Crawford and Sobel [1982] and Melumad and Shibano [1991]. Our model draws on elements from both these literatures to generate a comparison of centralized and decentralized structures in handling externalities.

An alternative, political-economy, approach to some of the questions in this paper has been developed by Lockwood [2002] and Besley and Coate [2003]. These papers present models where the central authority's decisions do not aim to be welfare maximizing, but rather are the outcome of an explicit voting or legislative decision-making process. Like this paper, and unlike much of the earlier literature on centralization versus decentralization, these papers do not assume that a central policy, must, by definition be a uniform one.

A model that shares some of the features of the one we develop here is the limited communication model of Melumad, Mookherjee, and Reichelstein [1995] which compares two-tier and three-tier hierarchies under limited communication and asymmetric information. The limited communication in their model yields screening problems similar to those that come from limited commitment in our model. More broadly, our model can be viewed as part of the literature on organizational design under asymmetric information

²A notable exception to this is Greenwood and McAfee [1991] who focus on inefficiencies generated by incentive constraints alone.

(see, for example, Milgrom, 1988; Poitevin, 2000). Dessein [2002], Alonso, Dessein, and Matouschek [2008] and Rantakari [2008] focus on the tradeoff between centralization and a form of decentralization when coordination is an issue. One major difference is that they do not allow for the allocation of rights and explicit contracts as in our model of decentralization.

In the next section, we formally describe the model. Section 3 examines the symmetric information benchmark for centralization and decentralization. Section 4 solves for the optimal centralized and decentralized outcomes under asymmetric information, gives a welfare comparison, and provides comparative statics. Section 5 concludes. Proofs are presented in an appendix.

2 The Model

There are two local governments, denoted localities 1 and 2 respectively, and a central government, denoted by C . In locality 1, there is a public project that can be undertaken with intensity $q \in [0, \infty)$. This project has some external effects on locality 2. We assume that the choice of q , once made, is extremely costly or impossible to change.

The public project might be the construction of an electric power plant. In this case q would represent the capacity of the plant. Locality 1 would benefit from the increased generating capacity, and locality 2 might suffer from increased pollution. Or, the project might be the development of a new vocational training program or other improvement in the educational system. Here q could be an indication of the size or quality of the program. This could provide direct benefits to the residents and businesses in locality 1 and may also result in benefits to businesses located in a neighboring locality through helping develop or attract a skilled workforce.

More formally, locality 1's utility function is given by

$$u_1(q, \theta, t) = \theta q - q^2/2 + t,$$

where q is project intensity and $\theta (\geq 0)$ is a parameter which measures the desirability of the project to locality 1. The expression θq represents the gross benefit to residents of locality 1. The expression $q^2/2$ represents the cost of the public project. We assume that

it has to be financed in locality 1. Transfers, which can be associated with equalization payments or regional subsidies and are denoted t , are, however, possible to shift some of the cost burden to locality 2.

Locality 2's utility is given by

$$u_2(q, \gamma, t) = -\gamma q - t,$$

where γ is a parameter which measures the degree to which the project hurts (or benefits) locality 2. Our formal analysis deals with the case of negative externalities ($\gamma \geq 0$). The case of positive externalities is entirely symmetric and thus our analysis (with appropriate absolute values inserted) applies to that case as well.

The central government maximizes equally-weighted social utility:

$$u_C(q, \theta, \gamma, t) = u_1 + u_2 = (\theta - \gamma)q - q^2/2.$$

For most of the paper, θ and γ will be assumed to be private information of localities 1 and 2 respectively. Since some parties may be uninformed, it is necessary to specify prior beliefs over these parameters. We assume that it is common knowledge that θ and γ are independently distributed according to distribution functions $F(\theta)$ and $G(\gamma)$ with strictly positive densities $f(\theta)$ on $\Theta = [\underline{\theta}, \bar{\theta}]$ and $g(\gamma)$ on $\Gamma = [\underline{\gamma}, \bar{\gamma}]$ respectively. Furthermore, we assume $\bar{\gamma} \leq \underline{\theta}$, so that a positive project intensity (or size) is always socially optimal. To keep things simple, and in order to derive explicit solutions, we assume that θ is uniformly distributed on Θ .

The externality problem we investigate can be clearly seen by noting that there is a difference between the project size that is optimal for locality 1 and the project size that is optimal for the central government (socially optimal). The level $q_p(\theta) = \theta$ is privately optimal for locality 1, while $q_s(\theta, \gamma) = \theta - \gamma$ is socially optimal. Note that the social optimum depends on both localities' information.

We assume that raising funds for transfers is socially costly (due to inefficiencies in taxation, for example). In particular, we assume that the central government and both localities have a lexicographic dislike of giving transfers. The lexicographic dislike can be viewed as the limit case where these social costs are infinitely small compared to the effects of the choice of project size, q . We focus on this limit case mainly for simplicity: assuming some small social cost $\epsilon > 0$, so that giving a transfer t would subtract

$(1 + \epsilon)t$ from utility, would complicate the algebra without qualitatively affecting our results. That transfers have some social cost is a common assumption in the regulation literature. This assumption is important for our results, as without it or an alternative reason for the central government to care about transfers, our model of centralization could always implement the first-best outcome (and thus do at least as well as our model of decentralization).

We consider two different constitutional environments within which the problem of determining the project level can be tackled. Our task will be to compare the expected outcomes in these two settings from the point of view of the central government, that is, according to expected social welfare.

In the first environment, called centralization, all property rights over the public project and transfers reside with the central government. Localities are vertically integrated; they become internal divisions of the central government's organization. Specifically, the central government can mandate a project level, q , and may also require transfers between the two localities. The center's difficulty lies in inferring the localities' information so as to pick an appropriate project level q . Because localities are not legal entities, they have no legal rights and there can be no enforceable contract between the central government and its divisions (localities) that constrains the center's choice of project level or transfers. Localities thus communicate with the central government knowing that it will then make a unilateral decision about the project level and transfers. Without property rights, localities have no legal grounds for appealing the central government's decisions.

Accordingly, we model the implementation of the public project under centralization by the following (centralization) game:

1. Locality 1 chooses an element in $\Theta = [\underline{\theta}, \bar{\theta}]$ and locality 2 chooses an element in $\Gamma = [\underline{\gamma}, \bar{\gamma}]$. These choices are communicated to the central government.
2. The central government chooses a project level, q , and transfer, t , to implement.

The central government's strategy is to choose a project level and a transfer contingent on the information reported by localities 1 and 2 respectively. The centralization game draws on the cheap talk communication game of Crawford and Sobel [1982].

As an alternative to centralization, we consider an environment in which locality 1 is endowed with property rights over the public project and both localities have property

rights over transfers. In this environment, localities are legal entities separate from the central government. We assume that the enforcement of property rights is achieved by a constitution which may only be modified with the consent of all the parties, and that this constitution establishes a law of contract by which the parties may voluntarily agree to give up these rights. For example, the parties may agree to allow the transfer to be chosen as a function of locality 1's choice of q . Such agreements are enforceable by a court that may rule only on whether an action deprived a party of its constitutional rights. For example, if localities 1 and 2 write a contract in which 2 promises certain transfers as a function of the project level and then locality 2 refuses to pay up, locality 1 may argue before the court that, had 1 known locality 2 would renege on its promised transfer, 1 would have exercised its right to choose q in a different way. The court would then rule in favor of locality 1 and force the transfer. (It is worth noting that such a court would be irrelevant under centralization because no rights belong to the localities there.) We refer to this environment, in which localities have property rights, as a decentralized environment.

For tractability, we assume that locality 2 has all the bargaining power when property rights are decentralized to the two localities. Locality 2 offers a contractual agreement to locality 1 that specifies a transfer from locality 2 to locality 1 as a function of the project level chosen by locality 1.³

Thus a decentralized setting for choice of the project level is modeled by the following (decentralization) game:

1. Locality 2 offers a contract specifying $t(q)$, a transfer from locality 2 to locality 1 as a function of the project level.
2. Locality 1 can accept or reject this offer. If locality 1 accepts locality 2's contract proposal, locality 1 chooses and implements a project level q and locality 2 pays locality 1 the transfer $t(q)$. If locality 1 rejects the contract proposal, it chooses and implements a project level q' , and locality 2 is not obligated to make any transfers.

³In a more general analysis, we could also consider the case where locality 1 makes the offer to locality 2. This would significantly complicate the analysis since locality 1 would be an informed principal. Multiplicity of equilibria would occur. We thus prefer to focus on the simpler case with the understanding that the payoffs from decentralization may be underestimated (depending on which equilibrium we would select in the alternative formulation).

We are interested in comparing social welfare in the two environments. Whether centralization or decentralization performs better will depend on the characteristics of the problem, specifically the distributions of the direct benefit and externality parameters, θ and γ . The special case of full information (i.e., degenerate distributions) is examined in the next section. The following section contains our main results exploring the trade-off between the two environments under non-degenerate distributions and the resulting asymmetric information.

In our analysis, we use the standard solution concept of Perfect Bayesian Equilibrium as defined in Fudenberg and Tirole [1991].

3 Full Information

Here we analyze project choice under centralization and decentralization assuming that the values of θ (the private benefit parameter) and γ (the externality parameter) are common knowledge.

3.1 Centralization

Consider the centralization game with θ and γ being commonly known to all players. It is easy to show that the central authority chooses the project level, q , and the transfer, t , which maximize the expectation of the center's utility, u_C , conditional on the true θ and γ .

$$\max_{q,t} u_C(q, \theta, \gamma, t) = (\theta - \gamma)q - q^2/2.$$

The solution is

$$\begin{aligned} q_C^{fi}(\theta, \gamma) &= \theta - \gamma \\ t_C^{fi}(\theta, \gamma) &= 0. \end{aligned}$$

The central authority chooses the socially optimal project level $q_s(\theta, \gamma)$ conditional on the true observed state $\{\theta, \gamma\}$ and sets transfers equal to zero because it has a mild (lexicographic) dislike for transfers.

First-best expected social welfare is achieved.

$$SW_C^{fi} = E_{\theta, \gamma} \left\{ (\theta - \gamma)^2 / 2 \right\}.$$

Under full information, there is no efficiency loss in centralizing all rights with the central authority.

3.2 Decentralization

Consider the decentralization game with the private benefit and externality parameters being commonly known to all players. It is easy to show that the equilibrium contract solves the following maximization problem.

$$\begin{aligned} \max_{q, t} \quad & u_2(q, \gamma, t) = -\gamma q - t \\ \text{s.t.} \quad & u_1(q, \theta, t) = \theta q - q^2 / 2 + t \geq \max_{q'} \left\{ \theta q' - q'^2 / 2 \right\} = \theta^2 / 2. \end{aligned}$$

The constraint is locality 1's participation constraint where the right-hand-side represents locality 1's welfare if it rejects the contract and produces at its privately optimal level. Solving this, we find

$$\begin{aligned} q_D^{fi}(\theta, \gamma) &= \theta - \gamma \\ t_D^{fi}(\theta, \gamma) &= \gamma^2 / 2. \end{aligned}$$

For each pair $\{\theta, \gamma\}$, locality 2 chooses the socially optimal project level and sets transfers such that locality 1 accepts producing below its private optimum.

The resulting expected social welfare (up to the lexicographic dislike for transfers) is

$$SW_D^{fi} = E_{\theta, \gamma} \left\{ (\theta - \gamma)^2 / 2 \right\}.$$

Thus, with known parameters, centralization always does at least as well as decentralization with the difference vanishing in the social cost of transfers. This will serve as a benchmark for the more interesting case explored below.

Our analysis also confirms Coase's [1960] intuition that, without transaction costs or bargaining imperfections, assigning property rights (to the project to locality 1 and to

refuse transfers to both localities) yields an efficient outcome. We now turn to the analysis of the case where information is asymmetrically distributed.

4 Asymmetric Information

Here we analyze centralization and decentralization assuming that the realizations of θ and γ , the benefit and externality parameters, are private information of localities 1 and 2 respectively. In each environment, we solve for the equilibrium of the associated game assuming that private information is realized before the first stage of the game.

4.1 Centralization

In the last stage of the centralization game, the central authority optimally chooses the project intensity and the transfer conditional on whatever information may have been revealed by the two localities in the preceding stage. As in the case with full information, in equilibrium, no transfers will ever be made at this stage because they are disliked by the center.⁴

In the first stage, each locality chooses an element from its set of possible parameters taking into account how this may affect the project intensity through the information this conveys to the center.

In equilibrium, two conditions —truth-telling for the two localities and conditional optimality of the central authority's choice of project level, q — restrict the amount of information that will be transmitted. A first implication is that, because transfers are constant in equilibrium, it is impossible to separate out the different possible externality values of locality 2. Locality 2's preferences are monotonically decreasing in the project size, q , and thus it will report whatever externality parameter will lead the center to lower the project level the most.

⁴In fact, the important feature is not that transfers are disliked, and that there are none, but rather that the central authority has some level of transfers that it strictly prefers for reasons external to and independent of the project and the externality. Without commitment, ex post the central authority would choose the transfer that maximizes its preferences. This implies that transfers will not provide incentives under centralization.

In contrast, it may be possible to elicit some information from locality 1 since its preferences are not monotonic in the project level q . However, even here, full revelation of the benefit parameter by locality 1 cannot be incentive compatible whenever the expected externality from the project is not zero (i.e., $E\gamma > 0$). To see this, suppose it were possible to achieve an incentive-compatible full separation of the different benefit parameters, θ . Conditional optimality then dictates implementing the project level that maximizes expected social welfare, $q_s(\theta, E\gamma)$. If $E\gamma > 0$, $q_s(\theta, E\gamma) < q_p(\theta)$ (locality 1's privately optimal project level) which implies that locality 1 would rather shade its announcement upward rather tell the truth.

Our analysis of centralization is closely related to Crawford and Sobel's [1982] seminal analysis of sender-receiver games. They study communication in a setting without transfers and are interested in the nature of communication for different parameterizations of preferences for the center and locality 1. In our model of centralization, the expected externality plays the role of the bias generating the conflict between the sender and receiver in Crawford and Sobel's.

As there, a coarser revelation of information may, however, be feasible. Consider an incentive compatible partition of Θ , $\mathcal{P}_\Theta = \{\Theta_1, \dots, \Theta_J\}$, where $\Theta_j = [\theta_{j-1}, \theta_j]$ for $j = 1, \dots, J-1$, $\Theta_J = [\theta_{J-1}, \theta_J]$ with $\theta_0 = \underline{\theta}$ and $\theta_J = \bar{\theta}$. Suppose that in the first stage of the game, locality 1 selects the interval that includes the true benefit parameter, θ . Then upon selection of Θ_j by locality 1, the central authority chooses the project level q_j that maximizes its expected social welfare conditional on θ belonging to Θ_j ,

$$q_j = \arg \max_q \left\{ \int_{\theta_{j-1}}^{\theta_j} [(\theta - E\gamma)q - q^2/2] f(\theta) d\theta \right\} \div (F(\theta_j) - F(\theta_{j-1})).$$

It is straightforward to show that q_j is the conditional expected benefit parameter minus the expected externality parameter. Under our assumption of a Uniform distribution for θ , this implies that $q_j = (\theta_j + \theta_{j-1})/2 - E\gamma$.

What is the optimal partition of a given size J (i.e., the optimal way to divide the set of possible benefit parameters into J categories for screening purposes)?⁵ It is the solution

⁵Note that this reasoning *assumes* that locality 1 self-selects according to a partition. That the restriction to partitions is without loss of generality follows from the results of Crawford and Sobel [1982]. See also Melumad and Shibano [1991].

to the following maximization problem.

$$\begin{aligned}
(1) \quad SW_C(J, E\gamma) := & \max_{\{q_j, \theta_j\}_{j=1}^J} \sum_{j=1}^J \int_{\theta_{j-1}}^{\theta_j} \left[(\theta - E\gamma)q_j - q_j^2/2 \right] f(\theta) d\theta \\
& \text{s.t.} \quad q_j = (\theta_j + \theta_{j-1})/2 - E\gamma \quad \forall j = 1, \dots, J \\
& \quad \theta_j \leq \theta_{j+1} \quad \forall j = 0, \dots, J-1 \\
& \quad \theta_0 = \underline{\theta}, \theta_J = \bar{\theta} \\
& \quad u_1(q_j, \theta_j, 0) \geq u_1(q_k, \theta_j, 0) \quad \forall \theta_j \in \Theta_j, \forall j, k = 1, \dots, J
\end{aligned}$$

In this problem, the optimal J -category partition maximizes expected social welfare subject to the constraint that project levels are conditionally optimal given the element of the partition announced by locality 1, that the cut-off types, θ_j do, indeed, define a J -category partition, and that the partition induces locality 1 to announce the interval containing the true benefit parameter. The next result characterizes the solution to problem (1).

Lemma 1 *Fix the partition size at J .*

(i) *A solution to the central authority's problem of optimally screening locality 1 using J categories (i.e., solving problem (1)) exists if and only if the expected externality is small enough. Specifically, it exists if and only if $E\gamma \leq (\bar{\theta} - \underline{\theta}) / (2J(J-1))$ holds.*

(ii) *The optimal cut-off types determining the partition are given by:*

$$\theta_j = \frac{(J-j)\underline{\theta} + j\bar{\theta}}{J} - 2E\gamma(J-j)j \quad \forall j = 0, \dots, J,$$

and the corresponding optimal project levels, by:

$$q_j = (\theta_j + \theta_{j-1})/2 - E\gamma \quad \forall j = 1, \dots, J.$$

(iii) *The expected social welfare loss under the optimal partition with J categories as compared to first-best expected social welfare is:*

$$SWL_C(J) = \frac{\sigma_\theta^2}{2J^2} + \frac{\sigma_\gamma^2}{2} + (E\gamma)^2 \frac{(J^2 - 1)}{6}$$

where σ_z^2 denotes the variance of z .

It is interesting to note that this characterization is obtained solely from incentive compatibility and conditional optimality. Imposing these two constraints eliminates all but at most one way (described in (ii)) of dividing $[\underline{\theta}, \bar{\theta}]$ into J subintervals.

How does the externality interact with incentive compatibility and conditional optimality? As the expected externality parameter becomes larger, optimality requires that project levels decrease (since the externality is negative). Incentive compatibility requires that each cut-off type, θ_j , be indifferent between project levels q_j and q_{j+1} . To maintain this indifference as project levels shift downward, the cut-off types must decrease as well (since production costs are convex). Therefore, the externality makes screening less efficient by distorting the partition downward, away from the ideal of J equal-sized intervals. This reduction in screening efficiency is why the expected externality parameter appears in the expression for the expected social welfare loss. The downward distortion by the externality also makes clear why, as stated in the first part of the lemma, high levels of screening (i.e., large J) are not possible when the expected externality parameter is large. For $E\gamma$ large enough to violate the inequality in (i) for some J , the distortion of cut-off types is so large as to imply that $\theta_1 \leq \underline{\theta}$, making that level of screening infeasible.

We show next that finer screening increases expected social welfare.

Lemma 2 *For all feasible partition sizes J , expected social welfare under the optimal partition is increasing in J .*

There are no social costs but some benefits to increased screening by the central authority as this allows finer tuning of project levels to locality 1's private information. We can now characterize the optimal partition for screening locality 1 and the central authority's expected social welfare loss under centralization. To do this, we follow the mechanism design literature in focusing on the unique Pareto-dominant equilibrium outcome.

Proposition 1 (i) *There is an equilibrium outcome which ex ante Pareto-dominates all other equilibrium outcomes.*

Under this equilibrium outcome,

(ii) *The optimal partition size is $J^* = \left\lfloor 1/2 + \sqrt{1/4 + (\bar{\theta} - \underline{\theta}) / (2E\gamma)} \right\rfloor$, where $\lfloor x \rfloor$ denotes the largest integer weakly smaller than x , and the partition described in Lemma 1(ii) for $J = J^*$ characterizes locality 1's equilibrium reports and the center's corresponding equilibrium project*

levels in the centralization game;

(iii) The equilibrium expected social welfare loss under the centralized regime is:

$$(2) \quad SWL_C(J^*) = \frac{\sigma_\theta^2}{2(J^*)^2} + \frac{\sigma_\gamma^2}{2} + (E\gamma)^2 \frac{((J^*)^2 - 1)}{6}.$$

The expression for J^* (the optimal number of screening categories) comes from inverting the feasibility condition of Lemma 1(i) to find the largest integer J that satisfies it. The proof works by showing that in all equilibria, expected output is identical. Since locality 2's utility is linear in output, it is indifferent among all these equilibria. From Lemma 2 we know that expected social welfare increases with the fineness of information revelation by locality 1 (because the expected output can be more efficiently allocated across θ types). Thus, it must be that locality 1 (as well as the center) prefers the most revealing equilibrium consistent with incentive compatibility and conditional optimality. The equilibrium path strategies and beliefs that support this equilibrium outcome are:

1. Locality 1 of type θ chooses its most preferred element of $\mathcal{P}_\Theta^* = \{\Theta_1^*, \dots, \Theta_{J^*}^*\}$ where $\{\theta_j^*\}_{j=0}^{J^*}$ solves problem (1), (i.e., chooses Θ_j^* such that $\theta \in \Theta_j^*$). Locality 2 selects the first element of Γ .
2. The central authority believes that locality 1's type is distributed uniformly on Θ_j^* and it implements $q_j^* = (\theta_{j-1}^* + \theta_j^*)/2 - E\gamma$ and $t^* = 0$.

The expected social welfare loss under centralization depends positively on the mean of the externality parameter, γ , and the variances of both the benefit and externality parameters. The variances affect the efficiency of the centralized regime through "screening effects." Suppose, for example, that the equilibrium involves no screening ($J^* = 1$), and consider a mean-preserving increase in the variance of one of the parameters. It is clear that the efficiency of the centralized solution is reduced because there is now more weight on types further away from the average type, which is what determines the chosen project level. This same argument can be applied to any fixed level of screening, J^* . However, the impact on social welfare loss of variation in θ is attenuated when J^* is larger since θ is then screened more closely.⁶ Since there is no effective screening of locality 2, the impact

⁶Note that, for a Uniform distribution, the variance of θ may also affect the optimal amount of screening since $\sigma_\theta^2 = (\bar{\theta} - \underline{\theta})^2 / 12$. As the support of θ increases, it may become feasible to increase the number of screening categories, J . Our discussion focuses on changes in the variance of θ that do not impact J^* .

of the variance of γ is independent of J^* .

Finally, there is an additional “screening effect” due to the mean value of the externality parameter. As the expected magnitude of the externality ($E\gamma$) increases, the social welfare loss compared to the full information optimum increases. There are two parts to this effect. First, there is a decrease in the efficiency of screening, holding the number of reporting categories, J^* , fixed. Second, a larger expected externality may make screening worse by reducing the feasible number of reporting categories. We now provide intuition for these two effects.

Fix the number of reporting categories, J^* , and consider an increase in the expected externality parameter such that J^* does not change. The optimal project levels q_j^* decrease because each unit of q is now socially less valuable. As argued earlier, the presence of truth-telling constraints for locality 1 forces a socially costly downward distortion in the cut-off types used for screening as the expected externality increases. This effect is captured by the term $(E\gamma)^2$ multiplying the last term of the expression (2).

There is an additional effect of the expected externality on J^* itself, because increasing $E\gamma$ may result in such a large distortion of cut-off types that using J^* categories to screen becomes impossible. When this happens, J^* must be reduced. All else equal, Lemma 2 shows that a coarser partition results in lower expected social welfare.

These arguments show that increasing the expected size of the externality ($E\gamma$) has an unambiguously negative effect on social welfare under centralization. In subsection 4.3 we present more comparative statics results and numerical examples incorporating the screening effects described here.

4.2 Decentralization

Under decentralization, locality 1 produces at its private optimum if no agreement has been reached, or it produces at its preferred level taking into account the promised transfer associated with this level if an agreement has been reached. This implies that locality 1 accepts all contract offers which are weakly preferred to producing at its privately optimal project level and receiving zero transfers. Locality 2 thus offers, in the first stage of the game, its preferred contract among those accepted by locality 1. This contract solves

the following maximization problem as a function of the externality parameter, γ .

$$(3) \quad \begin{aligned} & \max_{\{q(\theta), t(\theta)\}_{\theta=\underline{\theta}}^{\bar{\theta}}} \int_{\underline{\theta}}^{\bar{\theta}} u_2(q(\theta), \gamma, t(\theta)) f(\theta) d\theta \\ & \text{s.t.} \quad u_1(q(\theta), \theta, t(\theta)) \geq u_1(q_p(\theta), \theta, 0) \quad \forall \theta \\ & \quad \quad u_1(q(\theta), \theta, t(\theta)) \geq u_1(q(\theta'), \theta, t(\theta')) \quad \forall \theta, \theta' \end{aligned}$$

The first set of constraints are individual rationality constraints. Each type θ of locality 1 must get as much from accepting the contract as it can get by rejecting it and just producing at its privately optimal level, $q_p(\theta)$. The second set of constraints are standard incentive compatibility constraints that ensure that locality 1 chooses the project level designated for its value of θ . Locality 2 of type γ simply maximizes over all individually rational and incentive compatible contracts for locality 1.⁷

The following lemma characterizes the solution to locality 2's contracting problem.

Lemma 3 *For a given externality parameter $\gamma \in \Gamma$,*

(i) *The optimal project level is $q_\gamma(\theta) = \min \{2\theta - \gamma - \underline{\theta}, q_p(\theta)\}$.*

(ii) *The optimal transfer is*

$$t_\gamma(\theta) = \begin{cases} 0 & \text{if } \theta > \underline{\theta} + \gamma \\ (\theta - \gamma - \underline{\theta})^2 - (\underline{\theta} + \gamma)^2 / 2 + \min\{\bar{\theta}, \underline{\theta} + \gamma\}(\underline{\theta} + \gamma) - \min\{\bar{\theta}, \underline{\theta} + \gamma\}^2 / 2 & \text{if } \theta \leq \underline{\theta} + \gamma. \end{cases}$$

This lemma shows that the project level is at most as large as the privately optimal level q_p . Locality 2 trades off between decreasing the project level below q_p and giving rents for doing so to locality 1. When the socially optimal output, $\theta - \gamma$, is large, locality 2 prefers not to induce any decrease in q below $q_p(\theta)$, as $q_p(\theta)$ rises slower than the increase in q dictated by the incentive constraint. In standard screening problems, this effect is not present since the outside option is usually assumed to be type independent. For those

⁷If we had assumed that locality 1 made the offer, the corresponding contracting problem would need to include incentive constraints for both localities since locality 2's individual rationality constraints would depend on its beliefs about locality 1's type which influences locality 1's choice of project in the event of a rejection. Some low value θ 's may wish to mimic high θ 's to induce locality 2 into believing that θ is high and that locality 2 should expect a high project level if it rejects the contract offer, thus forcing locality 2 to accept an unfavorable contract offer. In this alternative formulation, there may exist multiple equilibria depending on the specification of out-of-equilibrium beliefs. This explains why, for simplicity and tractability, we assume that locality 2 makes all offers in the decentralized environment.

types producing at $q_p(\theta)$ the transfer is 0. When, however, the socially optimal output is small enough, locality 2 induces less production, and gives a compensating transfer to locality 1. In this case, the expression for this transfer varies depending on whether there are any types producing at the privately optimal level. If all θ types produce below their private optimum, then the transfer is chosen to give the highest type, $\bar{\theta}$, zero rents. If some θ types produce at their privately optimal level, then transfers provide zero rents to the lowest type in that set. This explains the two “min” terms in the expression for the optimal transfers.

This characterization can now be used to compute the central authority’s expected social welfare loss under decentralization as compared to the full information environment.

Proposition 2 (i) *The solution to locality 2’s contracting problem (3) is an equilibrium outcome of the decentralization game.*

(ii) *For a given externality parameter, γ , the social welfare loss of the central authority under decentralization is:*

$$SWL_D(\gamma) = \begin{cases} 2\sigma_\theta^2 & \text{if } \gamma > \bar{\theta} - \underline{\theta} \\ \gamma^2 (3(\bar{\theta} - \underline{\theta}) - 2\gamma) / 6(\bar{\theta} - \underline{\theta}) & \text{if } \gamma \leq \bar{\theta} - \underline{\theta} \end{cases}$$

(iii) *The expected social welfare loss under decentralization as compared to first-best expected social welfare is given by:*

$$SWL_D = G(\bar{\theta} - \underline{\theta}) \left\{ \frac{(\bar{\sigma}_\gamma^2 + (\bar{E}\gamma)^2)}{2} - \frac{(\bar{E}\gamma)^3 + 3\bar{\sigma}_\gamma^2 \bar{E}\gamma + \bar{\xi}_\gamma}{3(\bar{\theta} - \underline{\theta})} \right\} + (1 - G(\bar{\theta} - \underline{\theta})) 2\sigma_\theta^2$$

where $\bar{E}\gamma$ is the mean of γ conditional on $\gamma \leq \bar{\theta} - \underline{\theta}$, $\bar{\sigma}_\gamma^2$, the variance of γ conditional on the same event, and $\bar{\xi}_\gamma$, the third central moment of γ conditional on the same event.

The first part of the proposition establishes that the solution to problem (3) is an equilibrium outcome of the decentralization game. It is supported by the following equilibrium path strategies and beliefs:

1. Locality 2 of type γ offers the schedule $t_\gamma(q_\gamma^{-1}(q))$.
2. Locality 1 correctly infers the type γ of locality 2 and accepts the offered schedule. Locality 1 of type θ chooses the project level $q_\gamma(\theta)$ and is remunerated according to

the offered schedule. If locality 1 rejects the schedule, it chooses its privately optimal project level $q_p(\theta)$.

The second part of the proposition gives the social welfare loss for each possible value of the externality parameter. When the externality is large enough, that is, $\gamma > \bar{\theta} - \underline{\theta}$, all θ types are induced into producing strictly less than their private optimum, $q_p(\theta)$. Furthermore, as can be seen in the characterization of q_γ in Lemma 3, all marginal increases in the externality are fully internalized by bargaining between the two localities in this case. Thus the social welfare loss is independent of γ . It is proportional to the variance of θ , which reflects the effect of asymmetric information about θ on the efficiency of the separating allocation for each γ . The greater the variation in benefits for locality 1, the greater the social loss from the optimal contract's need to economize on informational rents.

When the externality parameter is relatively small, however, some values of θ will produce at their privately optimal level, and marginal increases in the externality will not be fully internalized by the schedule q_γ . As γ increases, there are two competing effects on social welfare loss. First, the number of θ types producing at their private optimum is decreasing. This reduces social welfare loss. Second, the distortion away from efficient project levels, for all types that remain at their private optimum, is increasing because this difference is equal to γ . The total effect on the social welfare loss of increasing the externality parameter is positive. The rate of increase, however, is not constant. When γ is low ($\gamma < (\bar{\theta} - \underline{\theta}) / 2$), the welfare loss is convex in γ . When γ is high (but below $\bar{\theta} - \underline{\theta}$), the welfare loss is concave in γ .

When the expectation of the social welfare loss is taken over γ , higher moments of its distribution matter. *Ceteris paribus*, the conditional variance affects the expected social welfare loss in the following way. The magnitude of the effect of variance is the difference between the expected loss and the loss evaluated at the expectation. By Jensen's inequality, if the conditional mean is in the convex region of $SWL_D(\gamma)$, that is, when $\overline{E\gamma} < (\bar{\theta} - \underline{\theta}) / 2$, increases in the variance increase the expected social welfare loss. If the conditional mean is in the concave region, then increases in the variance decrease the expected social welfare loss.

There are two effects of the conditional mean of γ on expected social welfare loss. The first effect is the exact analog of the effect of increasing γ described above. Namely, increases in $\overline{E\gamma}$ increase the expected social welfare loss, first at an increasing rate, then

at a decreasing rate. There is an additional effect due to the dispersion in γ . Since the convexity of the social welfare loss is everywhere decreasing in γ , the expected social welfare loss from variance, as described above, decreases as $\overline{E\gamma}$ increases.

Finally, the conditional third central moment,⁸ also affects the expected social welfare loss. As it increases, expected social welfare loss decreases. All else fixed, increasing “skewness” puts more weight in the lower end of the distribution. This decreases expected social welfare loss since social welfare loss is increasing in γ .

We are now in a position to compare the social welfare loss under centralization and decentralization and present some relevant comparative statics.

4.3 Comparison of the two regimes

Above we found the expected social welfare loss under centralization and decentralization. We now compare these two regimes and show how the distributions of the benefit and externality parameters, θ and γ , affect their relative ranking.

In Proposition 2, we saw that the expected social welfare loss under decentralization involves conditional moments of the distribution of γ . It is therefore difficult to make general comparisons of centralization and decentralization without making specific assumptions about the distribution of γ . We focus on two polar cases with the understanding that intermediate cases are combinations of the effects that we identify for these polar cases. The cases are chosen so that in the first, no matter what the value of the externality γ , all types of locality 1 produce below their private optimum under decentralization. In the other case, there is always at least some value of θ for which production is at the privately optimal level.

The first case is defined by the assumption that the externality is large compared to the support of possible private benefits, namely, $\underline{\gamma} > \bar{\theta} - \underline{\theta}$. Under centralization, this will imply that no screening can be done, since separating the possible types of locality 1 into more than one category ($J^* \geq 2$) is feasible if and only if $E\gamma \leq (\bar{\theta} - \underline{\theta}) / 4$ (Lemma 1). Therefore the optimal centralized policy will be a uniform one: the project level is set at $E\theta - E\gamma$ independent of the actual parameter values. From Proposition 1, the expected

⁸When normalized by $\bar{\sigma}^3$, this moment is referred to as the coefficient of skewness.

social welfare loss under centralization is then:

$$SWL_C(1) = \frac{\sigma_\theta^2}{2} + \frac{\sigma_\gamma^2}{2}.$$

Under decentralization, the expected social welfare loss is:

$$SWL_D = 2\sigma_\theta^2.$$

Only the variances of γ and θ influence the choice between centralization and decentralization in this case. The comparison is summarized in the following proposition.

Proposition 3 *Assume that $\underline{\gamma} > \bar{\theta} - \underline{\theta}$. Decentralization is preferred to centralization if and only if the variance of γ is sufficiently large compared to the variance of θ (specifically $\sigma_\gamma^2 > 3\sigma_\theta^2$).*

Under centralization, the central authority can implement only one project level. This implies that all terms in the expected social welfare loss involving the expected externality parameter disappear since they reflect the presence of incentive constraints, which are absent when no screening is done. The social welfare loss reflects only the loss due to the pooling of all types at a common project size. Under decentralization, for these parameter values, screening costs depend only on the variance of θ , and not on the externality, as all marginal increases in γ are internalized (Lemma 3).

The preferred regime is that with the lowest screening costs. Only when the variance of γ is large enough can decentralization be preferred. Suppose, for example, that the externality is known with certainty ($\sigma_\gamma^2 = 0$). Then, the centralized pooling allocation is preferred to the decentralized separating allocation. This implies that projects with potentially large and uncertain externalities should be decentralized, while projects with large but well-known externalities should be centralized. More generally, decentralization is the preferred regime when the externality component of the project is significantly more uncertain than the private benefit component. This is so because decentralized bargaining is better at tuning project levels to the size of the externality than is centralized decision making.

We can compare these results with those of Dessein [2002]. In a model of communication (centralization) versus delegation with no enforceable contracts (à la Crawford and Sobel [1982]), Dessein shows that decentralization dominates centralization if the variance of private information is large enough. This is similar to our result with regard to

the variance of the externality. However, our result goes in the opposite direction when considering the variance of the private benefit θ . This occurs because the variance of θ increases the welfare loss due to project size distortions under decentralization faster than it increases the loss due to lack of sensitivity to θ under centralization.

The second case we consider is defined by the assumption that the externality parameter is relatively small as compared to the support of the benefit parameter, namely, $\bar{\gamma} \leq \bar{\theta} - \underline{\theta}$. We can use the results of Propositions 1 and 2 to determine the difference in expected social welfare loss between centralization and decentralization, $SWL_C(J^*) - SWL_D$, to be equal to

$$(4) \quad \frac{\sigma_\theta^2}{2(J^*)^2} + \frac{(E\gamma)^2}{2} \left\{ -1 + \frac{(J^*)^2 - 1}{3} \right\} + \frac{(E\gamma)^3 + 3\sigma_\gamma^2 E\gamma + \xi_\gamma}{3(\bar{\theta} - \underline{\theta})}.$$

The following proposition summarizes the comparison between the two regimes in this case.

Proposition 4 *Assume that $\bar{\gamma} \leq \bar{\theta} - \underline{\theta}$. All else equal,*

- (i) *an increase in the variance of γ (σ_γ^2) increases the relative loss of centralization;*
- (ii) *an increase in the “skewness” of γ (ξ_γ) increases the relative loss of centralization;*
- (iii) *if centralization involves full pooling ($J^* = 1$), an increase in the expectation of γ ($E\gamma$) reduces the relative loss of centralization, while the opposite is true whenever centralization involves some separation ($J^* \geq 2$).*
- (iv) *if centralization is optimal, then it involves full pooling ($J^* = 1$).*

As in the previous case, variance in the externality is detrimental to centralization. This reinforces the observation that decentralization is good at incorporating locality 2’s information, while centralization is unable to do this.

A larger expected external effect favors decentralization unless it is already large enough to require full pooling (i.e., a uniform policy) under centralization. In the latter case, an increase in the expected externality favors centralization. When the average externality is small enough that $J^* \geq 2$, increasing $E\gamma$ worsens the screening problem under centralization as project sizes and cut-off types must adjust. When the optimal centralized policy is uniform, increases favor centralization since the screening problem cannot get any worse. Finally, as in the previous case, we establish that centralization can only be optimal when its policy is uniform.

It is interesting to note that Dessein [2002] derives a similar result concerning centralization only being optimal when communication is at its minimum (a uniform policy requires no communication). His comparison is in a setting where enforceable contracts are not feasible in either regime. In our model, the choice between regimes affects the feasibility of enforceable contracts. Both our results and Dessein's with regard to the uniformity of centralization may be read as emphasizing that screening under centralization without enforceable contracts is very costly.

5 Conclusion

There are at least two ways to interpret our results. First, we have shown that there are environments in which a central authority may desire to decentralize power (by conferring rights to localities) to provide incentives for greater incorporation of localities' ex post preferences in policies. Second, despite the limited contractual ability of our central authority, we have shown that centralization can still be optimal in some circumstances, even when the externality is not too large. This would imply that it may be optimal not to confer any rights to attain the second-best efficiency level. Going back to the Coase theorem, an attribution of rights may not be the optimal thing to do when there are bargaining imperfections.

In the Introduction we presented three pieces of "folk wisdom" about externalities and government structure. First, it is often suggested that large externalities justify central control. In our model, this statement must be qualified somewhat. It is not only the size of the externality per se that is relevant but also the degree of uncertainty about the externality relative to that about the private benefit parameter. For example, if two localities' parameters are distributed ex-ante identically, then centralization is preferred if the expected externality is large enough (Propositions 3 and 4(iii-iv)). If, however, the distributions are different, then decentralization is preferred if the variance of the externality is high enough (Propositions 3 and 4(i)). This implies that, for the case of a large expected externality, central control may be justified for projects known to have a narrow range of external effects. In that case, the (derived) uniformity of the optimal central policy is not too costly.

Second, it is often argued that decentralization is favored in situations of ex post local

heterogeneity. Again, our model shows that this argument should be qualified. If one interprets *ex post* local heterogeneity as the size of the variance of the parameters, then more heterogeneity in the externality favors decentralization (Propositions 3 and 4), while more heterogeneity in the private benefit favors centralization if the expected externality is high enough (Proposition 3), and has ambiguous effects otherwise.

Third, much of the prior literature takes for granted (or assumes) that central policies are insensitive to (*ex post*) local preferences. Our model shows that the contractual inability problem associated with central control endogenously generates limits on the center's ability to discriminate according to *ex post* realized preferences. In fact, whenever centralization is better than decentralization the optimal centralized policy is a uniform one (Propositions 3 and 4(iv)). Thus insensitivity to local preferences and the uniformity of optimal centralized policies is a result rather than an assumption.

Appendix

Proof of Lemma 1 (i) The first step of the proof is to examine the implications of the incentive compatibility constraints in problem (1). Consider a marginal type θ_j and project level q_{j+1} . Suppose that θ_j strictly prefers q_{j+1} to q_j . This would imply that there is a type $\theta_j - \epsilon$ that strictly prefers q_{j+1} to q_j . This would contradict the structure of the solution where type $\theta_j - \epsilon$ should have q_j . It must therefore be the case that type θ_j is indifferent between q_{j+1} and q_j . This implies that

$$u_1(q_j, \theta_j, 0) = u_1(q_{j+1}, \theta_j, 0).$$

Solving for this equation yields $\theta_j = (q_j + q_{j+1})/2$.

Furthermore, when these equalities are satisfied for all j , then all incentive constraints are satisfied.

Secondly, conditional optimality yields $q_j = (\theta_j + \theta_{j-1})/2 - E\gamma$.

These two sets of conditions give a system of $2J + 1$ linear equations with $2J + 1$ unknowns.

$$\begin{aligned} \theta_0 &= \underline{\theta} \\ \theta_j &= (q_j + q_{j+1})/2 \quad \forall j = 1, \dots, J-1 \\ \theta_J &= \bar{\theta} \\ q_j &= (\theta_{j-1} + \theta_j)/2 - E\gamma \quad \forall j = 1, \dots, J \end{aligned}$$

The solution to this system entails

$$\theta_j = \frac{(J-j)\underline{\theta} + j\bar{\theta}}{J} - 2E\gamma(J-j)j \quad \forall j = 0, \dots, J.$$

This solution is feasible if the intervals constructed from the θ_j 's form a partition of Θ and all the project levels, q_j , are non-negative. A necessary condition for this is that $\underline{\theta} = \theta_0 \leq \theta_1$. This condition amounts to:

$$\frac{(J-1)\underline{\theta} + \bar{\theta}}{J} - 2E\gamma(J-1) \geq \underline{\theta}.$$

It reduces to:

$$E\gamma \leq \frac{\bar{\theta} - \underline{\theta}}{2J(J-1)},$$

which is the feasibility condition in the lemma.

We have just shown that this condition is necessary for feasibility of the solution. We now show that it is also sufficient by showing that, under this condition, $\theta_{j-1} \leq \theta_j$ and $q_j \geq 0, \forall j = 1, \dots, J$. It is easy to show that

$$\theta_j - \theta_{j-1} = \frac{(\bar{\theta} - \underline{\theta})}{J} - 2E\gamma(J - 2j + 1).$$

Observe that $\theta_j - \theta_{j-1}$ is increasing in j . Since the necessary condition yields $\theta_1 - \theta_0 \geq 0$, it must be that $\theta_j - \theta_{j-1} \geq 0$ for all $j = 2, \dots, J$ as well. Furthermore, $q_j = \frac{\theta_{j-1} + \theta_j}{2} - E\gamma \geq \underline{\theta} - E\gamma \geq 0$, where the first inequality follows from θ_j non-decreasing and the second inequality follows from our assumption that $\underline{\theta} \geq \bar{\gamma}$. The condition is thus also sufficient for feasibility.

(ii) The proof of this part follows directly from that of part (i).

(iii) The expected social welfare loss is:

$$E_{\theta, \gamma} \left\{ \frac{(\theta - \gamma)^2}{2} \right\} - \sum_{j=1}^J \int_{\theta_{j-1}}^{\theta_j} \left[(\theta - E\gamma) \left(\frac{(\theta_j + \theta_{j-1})}{2} - E\gamma \right) - \frac{((\theta_j + \theta_{j-1})/2 - E\gamma)^2}{2} \right] \frac{1}{\bar{\theta} - \underline{\theta}} d\theta$$

where the θ_j 's are defined in (ii). This expression was simplified to the expression in the lemma using *Mathematica*⁹ and induction. We do not reproduce these steps here. *Q.E.D.*

Proof of Lemma 2 Using the expression in Lemma 1, we have that $SWL_C(J) - SWL_C(J-1)$ is equal to:

$$\frac{\sigma_\theta^2}{2} \left\{ \frac{1}{J^2} - \frac{1}{(J-1)^2} \right\} + \frac{(E\gamma)^2}{6} \left\{ (J^2 - 1) - ((J-1)^2 - 1) \right\}.$$

The sign of this expression can be shown to be the same as that of

$$(E\gamma)^2(J-1)^2J^2 - 3\sigma_\theta^2.$$

⁹*Mathematica* is a registered trademark of Wolfram Research, Inc.

We want to show that this expression is negative. The feasibility condition for a partition of size J as stated in Lemma 1 is

$$E\gamma \leq (\bar{\theta} - \underline{\theta}) / (2J(J - 1)),$$

which is equivalent to

$$(E\gamma)^2 \leq \frac{3\sigma_\theta^2}{J^2(J - 1)^2},$$

using the fact that the distribution of θ is Uniform. This expression is equivalent to

$$(E\gamma)^2(J - 1)^2J^2 - 3\sigma_\theta^2 \leq 0.$$

Q.E.D.

Proof of Proposition 1 (i) First we show that in any partition of Θ that may be revealed in an equilibrium outcome, expected output is the same. Specifically, from the proof of Lemma 1, $q_j = (\theta_{j-1} + \theta_j)/2 - E\gamma$. Therefore, $E_{\theta,\gamma}q = \sum_j q_j(\theta_j - \theta_{j-1})/(\bar{\theta} - \underline{\theta}) = E\theta - E\gamma$. Next, note that locality 2's utility is linear in output and independent of θ , thus locality 2 is ex-ante indifferent between all equilibrium outcomes. By Lemma 2, we know that the sum of the two localities' ex-ante utility is increasing in the fineness of the equilibrium revealed partition of Θ . Thus, it must be that locality 1 and the central authority both prefer the equilibrium outcome with the finest revelation of information about θ .

(ii) By Lemma 2 we know that the largest feasible partition is optimal. By Lemma 1 a J partition is feasible if and only if $E\gamma \leq (\bar{\theta} - \underline{\theta}) / (2J(J - 1))$. Inverting this expression and taking into account that J is an integer yields that $J^* = \left\lceil 1/2 + \sqrt{1/4 + (\bar{\theta} - \underline{\theta}) / (2E\gamma)} \right\rceil$.

The solution to problem (1) yields partitions $\mathcal{P}_\Theta^* = \{\Theta_1^*, \dots, \Theta_{J^*}^*\}$ and $\mathcal{P}_\Gamma^* = \{\Gamma\}$, where $\{\theta_j^*\}_{j=0}^{J^*}$ solves problem (1). The optimal project levels are $q_j^* = (\theta_{j-1}^* + \theta_j^*)/2 - E\gamma$. The following strategies and beliefs support this allocation as an equilibrium outcome of the centralization game.

1. Locality 1 of type θ chooses its most preferred element of \mathcal{P}_Θ^* , that is, the element Θ_j^* such that $\theta \in \Theta_j^*$.

Locality 2 always selects the first element of Γ .

2. The central authority believes that locality 1's type is distributed uniformly on Θ_j^* and it implements $q_j^* = (\theta_{j-1}^* + \theta_j^*)/2 - E\gamma$.

The central authority sets the transfer equal to zero.

It is easy to show that these strategies and beliefs form a PBE of the centralization game. In the last stage of the game, given the two localities' strategies and its own beliefs, the central authority maximizes expected social welfare when choosing the project levels and transfers. The central authority updates its prior beliefs and chooses the conditional optimal project level as given by the solution to problem (1). Transfers are set to zero.

In the first stage, locality 1 chooses the element of the partition \mathcal{P}_Θ^* that it prefers given the expected project level. Locality 2 selects the first element of Γ .

(iii) The expression (2) is the expression in Lemma 1 with J^* substituted in for J . *Q.E.D.*

Proof of Lemma 3 Define $U(\theta) \equiv u_1(q(\theta), \theta, t(\theta))$. As is standard in solving a problem like (3), consider the equivalent problem in terms of $q(\theta)$ and $U(\theta)$ and replace the incentive constraints with a constraint on the derivative of $U(\theta)$ and a monotonicity condition on $q(\theta)$. (See, e.g., Fudenberg and Tirole [1991], Chapter 7.)

$$(5) \quad \begin{aligned} & \max_{\{q(\theta), U(\theta)\}_{\theta=\underline{\theta}}^{\bar{\theta}}} \int_{\underline{\theta}}^{\bar{\theta}} \left((\theta - \gamma)q(\theta) - q(\theta)^2/2 - U(\theta) \right) f(\theta) d\theta \\ & \text{s.t.} \quad U(\theta) \geq \theta^2/2 \quad \forall \theta \\ & \quad \quad \frac{dU(\theta)}{d\theta} = q(\theta) \quad \forall \theta \\ & \quad \quad \frac{dq(\theta)}{d\theta} \geq 0 \quad \forall \theta \end{aligned}$$

Proceed by applying optimal control methods to problem (5), ignoring the monotonicity constraint on $q(\theta)$. The Hamiltonian is

$$\mathcal{H} = \left((\theta - \gamma)q(\theta) - q(\theta)^2/2 - U(\theta) \right) f(\theta) + \mu(\theta)q(\theta),$$

where $\mu(\theta)$ is the Pontryagin multiplier on the incentive constraint. The Lagrangian is

$$\mathcal{L} = \mathcal{H} + \tau(\theta)(U(\theta) - \theta^2/2),$$

where $\tau(\theta)$ is the Lagrange multiplier on the individual rationality constraint. Since the Hamiltonian is (weakly) concave and differentiable in q and U and the individual rationality constraint is quasi-concave in U , the following are sufficient conditions for a

solution to (5) (ignoring monotonicity):

$$(6) \quad \frac{\partial \mathcal{H}}{\partial q} = 0,$$

$$(7) \quad \frac{\partial \mathcal{L}}{\partial U} = -\mu'(\theta),$$

$$(8) \quad \frac{dU(\theta)}{d\theta} = q(\theta),$$

$$(9) \quad \tau(\theta)(U(\theta) - \theta^2/2) = 0, U(\theta) - \theta^2/2 \geq 0, \tau(\theta) \geq 0,$$

$$(10) \quad \mu(\underline{\theta})(U(\underline{\theta}) - \underline{\theta}^2/2) = 0, \mu(\bar{\theta})(U(\bar{\theta}) - \bar{\theta}^2/2) = 0, \mu(\underline{\theta}) \leq 0, \mu(\bar{\theta}) \geq 0.$$

This follows from modifying Seierstad and Sydsaeter [1987, Theorem 5.1] to incorporate an initial inequality on the state variable ($U(\underline{\theta}) \geq \underline{\theta}^2/2$) and to allow only continuous multipliers, $\mu(\theta)$. Note that for our problem, $f(\theta) = 1/(\bar{\theta} - \underline{\theta})$. To find a solution, guess that it will have (at most) two pieces: one in which the individual rationality constraint binds and $q(\theta)$ is at the privately optimal level θ , and another in which the IR does not bind and $q(\theta)$ is between the privately optimal and socially optimal level. In particular, consider the following values for the multipliers:

$$\mu^*(\theta) = \begin{cases} \gamma/(\bar{\theta} - \underline{\theta}) & \text{if } \theta > \underline{\theta} + \gamma \\ (\theta - \underline{\theta})/(\bar{\theta} - \underline{\theta}) & \text{if } \theta \leq \underline{\theta} + \gamma \end{cases}$$

and

$$\tau^*(\theta) = \begin{cases} 1/(\bar{\theta} - \underline{\theta}) & \text{if } \theta > \underline{\theta} + \gamma \\ 0 & \text{if } \theta \leq \underline{\theta} + \gamma. \end{cases}$$

Observe that the sign constraints on the multipliers in (9) and (10), as well as the relationship between them defined in (7) are satisfied. Now $q(\theta)$ is determined through (6), yielding

$$\begin{aligned} q^*(\theta) &= \begin{cases} \theta & \text{if } \theta > \underline{\theta} + \gamma \\ 2\theta - \gamma - \underline{\theta} & \text{if } \theta \leq \underline{\theta} + \gamma \end{cases} \\ &= \min \{2\theta - \gamma - \underline{\theta}, \theta\}. \end{aligned}$$

Next, $U(\theta)$ is determined from (8) and an initial condition determined by (10). This gives

$$U^*(\theta) = \begin{cases} \theta^2/2 & \text{if } \theta > \underline{\theta} + \gamma \\ \theta^2 - \theta(\underline{\theta} + \gamma) + \min \{\bar{\theta}, \underline{\theta} + \gamma\}(\underline{\theta} + \gamma) - \min \{\bar{\theta}, \underline{\theta} + \gamma\}^2/2 & \text{if } \theta \leq \underline{\theta} + \gamma. \end{cases}$$

It is readily verified that (6)–(10) are satisfied. Furthermore, as $q^*(\theta)$ is non-decreasing in θ , the monotonicity constraint (heretofore ignored) is satisfied as well. This proves (i) of Lemma 3. To show (ii), note that

$$t(\theta) = U(\theta) - \theta q(\theta) + q(\theta)^2/2.$$

Q.E.D.

Proof of Proposition 2 (i) The solution to problem (3) is characterized in Lemma 3. Define the function $t_\gamma^*(q) \equiv t_\gamma(q_\gamma^{-1}(q))$. The following strategies and beliefs support this solution as an equilibrium outcome of the decentralization game.

1. Locality 2 of type γ offers the schedule t_γ^* .
2. If t_γ^* has been offered, locality 1 correctly infers the type γ of locality 2 and accepts the offered schedule. If any other schedule has been offered, locality 1 accepts it if and only if it is weakly preferred to its private optimum.

In any case, locality 1 of type θ chooses the project level $q_\theta^* = \max_q u_1(q, \theta, t'(q))$, where t' is the offered and accepted schedule.

If locality 1 rejects the offered schedule, it produces at its private optimum $q_p(\theta)$.

It is easy to show that these strategies and beliefs form a PBE of the centralization game. In the second stage of the game, locality 1 accepts the schedule t_γ^* and believes that locality 2 has type γ with probability one.

If any other schedule has been offered, locality 1 keeps its prior beliefs and accepts it if it is weakly preferred to its private optimum.

In any case, locality 1 selects q_θ^* .

If the schedule is rejected, locality 1 produces at its private optimum $q_p(\theta)$.

In the first stage of the game, given the acceptance rule of locality 1, locality 2 offers the schedule t_γ^* .

(ii) Suppose first that $\gamma > \bar{\theta} - \underline{\theta}$. By Lemma 3, this implies that $q_\gamma(\theta) = 2\theta - \gamma - \underline{\theta}$ for all $\theta \in \Theta$. The social welfare loss is then:

$$\int_{\underline{\theta}}^{\bar{\theta}} \left\{ \frac{(\theta - \gamma)^2}{2(\bar{\theta} - \underline{\theta})} - \left(\frac{(\theta - \gamma)(2\theta - \gamma - \underline{\theta}) - (2\theta - \gamma - \underline{\theta})^2/2}{(\bar{\theta} - \underline{\theta})} \right) \right\} d\theta.$$

The first term represents the first best social welfare, while the second term is the social welfare under decentralization. This expression reduces to:

$$\int_{\underline{\theta}}^{\bar{\theta}} \left\{ \frac{(\theta - \gamma - (2\theta - \gamma - \underline{\theta}))^2}{2(\bar{\theta} - \underline{\theta})} \right\} d\theta = \int_{\underline{\theta}}^{\bar{\theta}} \frac{(\underline{\theta} - \theta)^2}{2(\bar{\theta} - \underline{\theta})} d\theta = \frac{(\bar{\theta} - \underline{\theta})^2}{6} = 2\sigma_{\theta}^2.$$

Suppose now that $\gamma \leq \bar{\theta} - \underline{\theta}$. By Lemma 3, this implies that $q_{\gamma}(\theta) = \theta$ for all $\theta > \underline{\theta} + \gamma$ and $q_{\gamma}(\theta) = 2\theta - \gamma - \underline{\theta}$ for all $\theta \leq \underline{\theta} + \gamma$. The social welfare loss is then:

$$\begin{aligned} & \int_{\underline{\theta} + \gamma}^{\bar{\theta}} \left\{ \frac{(\theta - \gamma)^2}{2(\bar{\theta} - \underline{\theta})} - \frac{(\theta - \gamma)(\theta) - \theta^2/2}{(\bar{\theta} - \underline{\theta})} \right\} d\theta \\ & + \int_{\underline{\theta}}^{\underline{\theta} + \gamma} \left\{ \frac{(\theta - \gamma)^2}{2(\bar{\theta} - \underline{\theta})} - \frac{(\theta - \gamma)(2\theta - \gamma - \underline{\theta}) - (2\theta - \gamma - \underline{\theta})^2/2}{(\bar{\theta} - \underline{\theta})} \right\} d\theta \\ & = \int_{\underline{\theta} + \gamma}^{\bar{\theta}} \frac{\gamma^2}{2(\bar{\theta} - \underline{\theta})} d\theta + \int_{\underline{\theta}}^{\underline{\theta} + \gamma} \frac{(\underline{\theta} - \theta)^2}{2(\bar{\theta} - \underline{\theta})} d\theta = \frac{3\gamma^2(\bar{\theta} - \underline{\theta}) - 2\gamma^3}{6(\bar{\theta} - \underline{\theta})}. \end{aligned}$$

(iii) The expected social welfare loss is computed by taking the expectation of $SWL_D(\gamma)$ over γ .

$$\begin{aligned} SWL_D &= G(\bar{\theta} - \underline{\theta}) \int_{\underline{\gamma}}^{\bar{\theta} - \underline{\theta}} \frac{3\gamma^2(\bar{\theta} - \underline{\theta}) - 2\gamma^3}{6(\bar{\theta} - \underline{\theta})} \frac{g(\gamma)}{G(\bar{\theta} - \underline{\theta})} d\gamma + (1 - G(\bar{\theta} - \underline{\theta})) 2\sigma_{\theta}^2 \\ &= G(\bar{\theta} - \underline{\theta}) \int_{\underline{\gamma}}^{\bar{\theta} - \underline{\theta}} \left\{ \frac{\gamma^2}{2} - \frac{\gamma^3}{3(\bar{\theta} - \underline{\theta})} \right\} \frac{g(\gamma)}{G(\bar{\theta} - \underline{\theta})} d\gamma + (1 - G(\bar{\theta} - \underline{\theta})) 2\sigma_{\theta}^2 \\ &= G(\bar{\theta} - \underline{\theta}) \left\{ \frac{(\bar{\sigma}_{\gamma}^2 + (\overline{E\gamma})^2)}{2} - \frac{(\overline{E\gamma})^3 + 3\bar{\sigma}_{\gamma}^2 \overline{E\gamma} + \bar{\xi}_{\gamma}}{3(\bar{\theta} - \underline{\theta})} \right\} + (1 - G(\bar{\theta} - \underline{\theta})) 2\sigma_{\theta}^2 \end{aligned}$$

where $\overline{E\gamma}$ is the mean of γ conditional on $\gamma \leq (\bar{\theta} - \underline{\theta})$, $\bar{\sigma}_{\gamma}^2$, the variance of γ conditional on the same event, and $\bar{\xi}_{\gamma}$, the third central moment of γ conditional on the same event. Q.E.D.

Proof of Proposition 3 This is a straightforward manipulation of $SWL_C(1)$ and SWL_D . Q.E.D.

Proof of Proposition 4 Parts (i) and (ii) are straightforward.

(iii) For $J^* = 1$, we have:

$$SWL_C(1) - SWL_D = \frac{\sigma_{\bar{\theta}}^2}{2} - \frac{(E\gamma)^2}{2} + \frac{(E\gamma)^3 + 3\sigma_{\gamma}^2 E\gamma + \xi_{\gamma}}{3(\bar{\theta} - \underline{\theta})}.$$

The derivative with respect to $E\gamma$ is then,

$$\frac{d}{d(E\gamma)}(SWL_C(1) - SWL_D) = -E\gamma + \frac{(E\gamma)^2}{\bar{\theta} - \underline{\theta}} + \frac{\sigma_{\gamma}^2}{\bar{\theta} - \underline{\theta}}.$$

Observe that this derivative is largest if, given $E\gamma$, σ_{γ}^2 is as large as possible. However, given the constraints on γ ($0 \leq \underline{\gamma} \leq \bar{\gamma} \leq \bar{\theta} - \underline{\theta}$) the maximum value of σ_{γ}^2 given $E\gamma$ is attained by the two-point distribution that puts probability $E\gamma / (\bar{\theta} - \underline{\theta})$ on $\gamma = \bar{\theta} - \underline{\theta}$ and probability $1 - (E\gamma / (\bar{\theta} - \underline{\theta}))$ on $\gamma = 0$. The resulting variance is $(\bar{\theta} - \underline{\theta})(E\gamma) - (E\gamma)^2$. Substituting this value into the expression for the derivative yields,

$$\frac{d}{d(E\gamma)}(SWL_C(1) - SWL_D) = 0.$$

This shows that the derivative must be non-positive and the first part of (iii) follows directly. To see the second part of (iii), observe that if $J^* \geq 2$ then all terms in (4) involving $E\gamma$ are positive and increasing in $E\gamma$.

(iv) The difference in expected social welfare loss between centralization and decentralization is

$$SWL_C(J^*) - SWL_D = \frac{\sigma_{\bar{\theta}}^2}{2(J^*)^2} + \frac{(E\gamma)^2}{2} \left\{ -1 + \frac{(J^*)^2 - 1}{3} \right\} + \frac{(E\gamma)^3 + 3\sigma_{\gamma}^2 E\gamma + \xi_{\gamma}}{3(\bar{\theta} - \underline{\theta})}.$$

After substituting for the various moments of γ , this expression reduces to

$$\frac{(\bar{\theta} - \underline{\theta})^3}{(J^*)^2} + 4(\bar{\theta} - \underline{\theta})((J^*)^2 - 4)(E\gamma)^2 + (E\gamma^3).$$

When $J^* \geq 2$, the first and last terms are positive, while the middle term is nonnegative, which implies that $SWL_C(J^* \geq 2) - SWL_D > 0$.

Q.E.D.

References

- [1] ALONSO, R., W. DESSEIN, AND N. MATOUSCHEK (2008) "When Does Coordination Require Centralization?," *American Economic Review* **98**, 145–179.
- [2] BESLEY, T. AND S. COATE (2003) "Centralized versus Decentralized Provision of Local Public Goods: a Political Economy Approach," *Journal of Public Economics* **87**, 2611–2637.
- [3] COASE, R. (1960) "The Problem of Social Cost," *Journal of Law and Economics* **3**, 1–44.
- [4] CRAWFORD, V.P. AND J. SOBEL (1982) "Strategic Information Transmission," *Econometrica* **50**, 1431–51.
- [5] DE LONG, J.B. AND A. SHLEIFER (1993) "Princes and Merchants: European City Growth Before the Industrial Revolution," *Journal of Law and Economics* **36**, 671–702.
- [6] DESSEIN, W. (2002) "Authority and Communication in Organizations," *Review of Economic Studies* **69**, 811–838.
- [7] FARRELL, J. (1987) "Information and the Coase Theorem," *Journal of Economic Perspectives* **1**, 113–129.
- [8] FUDENBERG, D. AND J. TIROLE (1991) *Game Theory*. M.I.T. Press, Cambridge.
- [9] GREENWOOD, J. AND R.P. MCAFEE (1991) "Externalities and Asymmetric Information," *Quarterly Journal of Economics* **106**, 103–121.
- [10] KLIBANOFF, P. AND J. MORDUCH (1995) "Decentralization, Externalities, and Efficiency," *Review of Economic Studies* **62**, 223–247.
- [11] LAFFONT, J.-J. AND E. MASKIN (1979) "A Differential Approach to Expected Utility Maximizing Mechanisms," in Laffont, J.-J., ed., *Aggregation and the Revelation of Preferences*. North Holland Publishing Co.
- [12] LOCKWOOD, B. (2002) "Distributive Politics and the Benefits of Decentralization," *Review of Economic Studies* **69**, 313–337.
- [13] MELUMAD, N., D. MOOKHERJEE, AND S. REICHELSTEIN (1995) "Hierarchical Decentralization of Incentive Contracts," *Rand Journal of Economics* **26**, 654–672.
- [14] MELUMAD, N. AND T. SHIBANO (1991) "Communication in Settings with no Transfers," *Rand Journal of Economics* **22**, 173–198.

- [15] MILGROM, P.R. (1988) "Employment Contracts, Influence Activities, and Efficient Organization Design," *Journal of Political Economy* **96**, 42–60.
- [16] MYERSON, R.B. AND M.A. SATTERTHWAIT (1983) "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory* **29**, 265–281.
- [17] OATES, W. (1972) *Fiscal Federalism*. Harcourt, Brace, Jovanovich, New York.
- [18] POITEVIN, M. (2000) "Can the Theory of Incentives Explain Decentralization?," *Canadian Journal of Economics* **33**, 878–906.
- [19] RANTAKARI, H. (2008) "Governing Adaptation," *Review of Economic Studies* **75**, 1257–1285.
- [20] ROB, R. (1989) "Pollution Claim Settlements Under Private Information," *Journal of Economic Theory* **47**, 307–333.
- [21] SEIERSTAD, A. AND K. SYDSAETER (1987) *Optimal Control Theory with Economic Applications*. North Holland, Amsterdam.