Title
Cleanup delays at hazardous waste sites: an incomplete information game

Authors
Rausser, Gordon C.
Simon, Leo K.
Zhao, Jinhua

Publication Date
1999-04-01

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS AND POLICY
DIVISION OF AGRICULTURE AND NATURAL RESOURCES
UNIVERSITY OF CALIFORNIA AT BERKELEY

WORKING PAPER NO. 839

CLEANUP DELAYS AT HAZARDOUS WASTE SITES: AN INCOMPLETE INFORMATION GAME

by

Gordon C. Rausser, Leo K. Simon, and Jinhua Zhao

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
April 1999

CLEANUP DELAYS AT HAZARDOUS WASTE SITES:
AN INCOMPLETE INFORMATION GAME

GORDON C. RAUSSER, LEO K. SIMON, AND JINHUA ZHAO

APRIL 6, 1999

ABSTRACT. This paper studies the incentives facing Potentially Responsible Parties at a hazardous waste site to promote excessive investigation of the site and thus postpone the beginning of the remediation phase of the cleanup. We model the problem as an incomplete information, simultaneous-move game between PRPs. We assume that PRP's liability shares are predetermined. Each PRP's type is its private information about the precision of its own records relating to the site. A strategy for a PRP is a function mapping its type into announced levels of precision. Once types have been realized, the regulator aggregates the realized precision announcements and imposes the investigation schedule according to a predetermined policy function. We show that a pure-strategy Nash equilibrium exists, in which each PRP's strategy is monotone increasing in its type. We prove that PRPs with higher liability shares have greater incentives to delay than those with lower shares. We also show that under certain conditions, when liability shares become more homogenous, delay becomes more likely. We demonstrate that when certain conditions are imposed on our model, it predicts that two widespread practices—*de minimis* buyouts and the formation of steering committtess—will tend to increase delay.

JEL classification: D82, Q28

Keywords: Environmental economics; Superfund; hazardous waste cleanups; Potentially responsible parties (PRPs); strategic information transmission; strategic delay; environmental remediation; cleanup.

The authors are, respectively: Dean, College of Natural Resources, Robert Gordon Sproul Distinguished Professor, Professor, Department of Agricultural and Resource Economics, and Member, Giannini Foundation of Agricultural Economics, University of California at Berkeley; Adjunct Professor, Department of Agricultural and Resource Economics, University of California at Berkeley; and Assistant Professor, Department of Economics, Iowa State University.

Address correspondence to: Jinhua Zhao, Department of Economics, Iowa State University, Heady Hall, Ames, IA 50011. Email: jzhao@iastate.edu.

The authors are grateful to participants at the Second Toulouse Conference on Environmental and Resource Economics and at the Economic Theory Seminar Series at Iowa State University. Susan Athey and Chris Shannon provided very helpful advice relating to the application of Athey's methodology to our problem.

1. Introduction

A major impediment to the rapid cleanup of hazardous waste sites is the information asymmetry between the regulatory authority and the corporations that have been identified as potentially liable for the damage, known as Potential Responsible Parties (PRPs). At any particular polluted site, each PRP will have private information about its own contribution to the site, including the nature and geographic distribution of the substances contributed. Typically, this private information will be imperfect, due to gaps in PRPs' records of the magnitude, transportation and diffusion of their contributions. However, PRPs' private information will generally be more precise than the information that is directly available to the regulatory authority. Because of this asymmetry, together with the enormously contentious issue of how to apportion liability shares among PRP's, the process of cleaning up hazardous waste sites has been characterized by prolonged negotiation and extensive litigation. As a result, the pace of cleanups has been extremely slow, generating considerable public concern.

Numerous experts have identified the litigation and negotiation processes as the main cause of cleanup delays and have called for ways of reducing the incentives for participating in these processes.[1] However, as Dixon (1994) observed, PRPs actually benefit from delay because it reduces their *discounted* cleanup costs. Cost savings due to discounting may be significant: as reported by Birdsall and Salah (1993), prejudgment interest is the single largest cost item at a site and accounts for nearly one-third of the total costs involved.[2] Thus discounting provides an incentive, in addition to disagreement about liability shares, for PRP's to litigate and negotiate, and thereby delay the cleanup process in legally acceptable ways.

[1] See, for example, Dixon (1994) and Church and Nakamura (1993).
[2] Prejudgment interest is accumulated when the government or a PRP sues other PRPs for past cleanup costs.

At most hazardous waste sites, the extent of contamination, and hence aggregate liability, is highly uncertain. It is, therefore, suboptimal to proceed very rapidly to the cleanup phase of the remediation process; rather, time is needed to conduct field investigations that will reduce uncertainty about the nature of the cleanup task. The issue of strategic information revelation naturally arises because of an inevitable information asymmetry between PRPs and the regulatory authority. The authority must determine how long a site should be investigated, based on the quality of its initial information about the degree of contamination. If the investigation period is too short, inappropriate remediation strategies may be adopted: if the extent of contamination is overestimated, then excessive resources may be allocated to remediation; if it is underestimated, then the remediation plan may be inadequate, resulting in exacerbated health risks and costly revisions to the original cleanup schedule. The greater the uncertainty, therefore, the longer is the optimal investigation period, and hence the longer is the optimal delay in cleanup. Apart from differences between their respective rates of discount, there are other factors that lead PRPs to prefer cleanup schedules that differ from the socially optimal schedule. In particular, they have different degrees of risk aversion and face liabilities that differ in both their nature and their extent. The EPA is less risk averse than the PRPs, due to risk pooling among the many sites an EPA office is overseeing. Moreover, PRPs only pay part of the off-site costs and the residual costs of a site.[3] For these reasons, PRPs are likely to overvalue the benefits of uncertainty reduction and to undervalue the costs of extending the investigation period beyond its socially optimal length. On the other hand, each individual

[3] Off-site costs are incurred when toxic substances migrate from the designated site and create health hazards, or when neighboring property values are diminished due to their proximity to the site. When health hazard is involved, PRP's typically do not bear the entire burden of off-site costs. This is because the causal link between any health damage and the actions of a particular PRP is typically difficult to establish, and because the burden of proof rests with the parties that have been harmed. Since transaction costs are so high and the probability of prevailing in litigation is so low, citizens are deterred from suing PRP's unless the hazard is significant. That is, PRP's expect to pay only a part of the off-site costs. Moreover, while health damage increases with exposure time, compensation for this damage can only be obtained through a private cost recovery action. Accordingly the full effect of time associated with health damages is unlikely to be reflected. Thus, PRP's have inadequate incentives to expedite the cleanup process to reduce health damages. Residual costs are incurred when the cleanup does not eliminate all the hazardous substances, so that further monitoring, maintenance or cleanup of the site is required and, possibly, there is additional damage to health and neighboring property. Currently, since the state governments are largely responsible for monitoring and maintaining de-listed sites, PRP's do not generally bear the full burden of residual costs.

PRP is responsible for only a fraction of total cleanup costs, and for this reason will *underweight* the benefits of uncertainty reduction. The net effect of these differences is thus indeterminate.

In order to determine the optimal investigation period, the regulator must make an initial estimate of the uncertainty associated with each site. In order to make this estimate, it must rely on the PRPs' documents and reports. By strategically misreporting their private information, PRPs can thus manipulate the regulator's decision and either hasten or delay proceedings. Rausser, Simon and Zhao (1998) (henceforth RSZ1) studies this strategic interaction between the PRPs and the regulator. In that paper PRPs strategically determine what to report about the accuracy of their information about their individual contributions. Using a relatively informal model, the paper identifies incentives for misreporting of accuracy levels and suggests several Bayesian mechanisms which would enable the regulator to extract the truth from the PRPs. In this paper, we formalize the interaction among the PRPs as an incomplete information game, given the regulator's policy. To sharpen analysis, we make two simplifying assumptions. First, we assume that PRP's liability shares are predetermined. Second, we assume that the regulator treats PRPs' reports as truthful, even though they are in fact strategically determined. While the latter assumption implies that the regulator's behavior is not fully rational, it can be defended on several grounds. The first is institutional: it seems quite consistent with actual regulator behavior: governmental bureaucrats typically take reports from agents under their jurisdiction at face value, rather than attempting to reverse engineer "the truth" from these reports, based on what the bureaucrats know about the agents' motivations. The second is pragmatic: if the regulator is modelled as acting fully rationally, then no pure strategy equilibrium will exist for the resulting game. Moreover, the simplification allows us to single out other factors which affect PRPs' incentives to delay, to focus upon the strategies the PRPs can pursue, and to identify the implications of some widespread government practices.

Each PRP in our game has information about the precision of its own records relating to the site. We identify this information with the agent's *type*. A strategy for a PRP is a function mapping its type into announced levels of precision. Once types have been realized, the regulator aggregates PRP's announcements and imposes the investigation schedule that is optimal relative to this aggregated announcement.

Assuming that the type space is nonatomic, we apply a methodology developed in Athey (1997) to establish the existence of a pure strategy Nash equilibrium in which each PRP's report is monotone in its type. We then study the role of liability shares: we show that bigger PRPs (those with higher shares) tend to over-report more, and PRPs tend to report more when their liability shares increase. We also show that under certain conditions, the aggregate report is expected to increase as the PRPs become more heterogenous in their liability shares.

In Section 2, we formulate the incomplete information game among the PRPs. Section 3 shows the existence of a pure strategy Nash equilibrium in which the PRPs' strategies are non-decreasing in their types. Section 4 shows that PRPs with higher liability shares tend to report a higher variance than those with lower shares. In Section 5, we show (for the case of two PRPs) that under certain conditions, when their liability shares become more heterogeneous, the PRPs' expected aggregate report increases. Section 6 extends this result to the case of multiple PRPs, but imposes significant restrictions on function forms. Section 7 studies the implications of some important aspects of the cleanup process: *de minimis* PRP buyouts and the formation of PRP steering committees. All of the proofs are gathered together in Appendix 8.

2. A model of Superfund cleanup

In the presentation that follows, we will use the following notational convention: given a vector $\boldsymbol{x} = (x_1, ..., x_n)$ or function $\boldsymbol{f(x)} = (f_1(x_1), ..., f_1(x_n))$ we denote the sum of the elements of $\boldsymbol{x}$ ($\boldsymbol{f(x)}$) by Σx (resp. $\Sigma f(x)$ or Σf) and the sum of all but the i'th element by Σx_{-i} (resp. Σf_{-i}).

In general, contaminated sites are characterized by a multiplicity of attributes, including concentration, toxicity, dispersion, etc. In this paper, we will abstract from these complexities and assume that sites are fully characterized by the *volume* of total contamination. Let n denote the number of PRPs, indexed by $i = 1, \ldots, n$. Prior to any investigation of the site, each PRP has private information about its contribution level. This information is assumed to be imperfect, however, due perhaps to the incompleteness of PRP records or to movement of the contaminant. Accordingly, we assume that i's contribution m_i is a random variable with mean $\bar{m}_i$ and variance θ_i. Assume further that $\bar{m}_i$ is common knowledge, while θ_i is known only by agent i. We assume that the m_i's are independent of each other, so that no PRP can infer from its own information the extent to which other PRP's contributions are uncertain. The total volume of contamination at the site is denoted by Σm. Clearly, Σm is a random variable with commonly known mean $\Sigma \bar{m}$. The variance of Σm, $\Sigma\theta$ is not commonly known: each PRP has partial information about $\Sigma\theta$ (i.e. information about the variance of its own contribution), while the regulator has no independent information at all about $\Sigma\theta$.

At the beginning of the planning stage, the regulator requests each PRP i to report the variance of its own contribution θ_i. Based on the reports, the regulator determines an optimal schedule of field investigation to generate further information (i.e. to reduce the variance). Since it is common knowledge that the regulator will choose an investigation length based on reported θ_i's, each PRP can strategically misreport its uncertainty to manipulate the regulator's decision. Intuitively, if

PRP i prefers to delay the cleanup beyond the socially optimal investigation period, it will report a value for θ_i that exceeds the truth. To highlight the role of uncertainties, we assume that PRP's have previously agreed to bear liability in proportion to their commonly known *expected* contributions. That is, agent i agrees to bear the share k_i of total liability, where $k_i = \frac{m_i}{\bar{m}}$. We will refer to a vector of liability shares $\boldsymbol{k} = (k_1, ..., k_n)$, with $k_i \geq 0$ and $\sum_{i=1}^{n} k_i = 1$, as a *liability profile.* We assume that $\boldsymbol{k}$ is common knowledge.

We formalize the interaction among the PRPs as an incomplete information game. Let the variance θ_i be the *type* of PRP i, which is known by PRP i only. We assume that the θ_i's are identically and independently distributed on the interval $[\theta^l, \theta^u]$, where $\theta^l > 0$. Let $g(\cdot)$ denote the density of agents' types. We assume that $g(\cdot)$ is nonatomic on $[\theta^l, \theta^u]$. Let $\boldsymbol{\Theta} = [\theta^l, \theta^u]^n$ and $\boldsymbol{\Theta}_{-i} = [\theta^l, \theta^u]^{n-1}$. For $\boldsymbol{\theta}_{-i} \in \boldsymbol{\Theta}_{-i}$, let $\boldsymbol{g}_{-i}(\boldsymbol{\theta}_{-i}) = \prod_{j \neq i} g(\theta_j)$.

The PRP's, acting strategically, simultaneously declare their types. The regulator then chooses the investigation schedule that would be optimal if the type revelations by PRP's were truthful.[4] The regulator's response function is common knowledge. Thus the problem we have posed can be analyzed as a single-stage, incomplete information game with continuous payoffs.[5]

A *pure strategy* for the i'th PRP is a function $s_i : [\theta^l, \theta^u] \to H = [\underline{\eta}, \bar{\eta}]$. A natural lower bound for the announced variance is zero, but for generality, we assume that $\underline{\eta}$ is nonnegative. $\bar{\eta}$ is chosen to exceed the optimal report for a PRP of type θ^u with liability share 1, assuming that all other PRPs report $\underline{\eta}$ with probability one.[6] The number $s_i(\theta_i)$ represents i's declared (as opposed to actual) type.[7] Since PRPs' strategies are bounded below by $\underline{\eta}$ and are nonincreasing in other agents'

---

4 See page 4 for a discussion on this assumption.

5 An alternative formulation would place the regulator on an equal informational footing with the PRP's, i.e., to assume that the regulator knew PRPs' strategies, though not their realized types. The present formulation is chosen for reasons of tractability. In particular, it seems very unlikely that a pure-strategy equilibrium would exist under the more complex formulation. As a practical matter, it seems reasonable to presume that PRPs might know each other's strategies, while the regulator, whose involvement with any given site is peripheral, would not have access to this information.

6 Clearly in any equilibrium for any i, $s_i(\cdot)$ must be strictly dominated by the scalar $\bar{\eta}$. Thus we can avoid the corner solution at $\bar{\eta}$ in calculating a PRP's optimal strategies.

strategies, the range of any PRP's best response correspondence will be contained in the interval $[\underline{\eta}, \bar{\eta}]$. Let $\boldsymbol{\theta}$ denote a vector of type-realizations $(\theta_i, \cdots, \theta_n)$ and let $\boldsymbol{k}$ and $\boldsymbol{s}$ denote, analogously, a vector of liability shares and strategies. A vector of strategies $\boldsymbol{s}$ is called a *pure strategy profile.*

We turn now to specify the payoff functions of the PRPs. Let $V(t, \nu, k) : R_+ \times R_+ \times [0, 1] \rightarrow R$ be the expected utility loss of a PRP who bears responsibility for a fraction k of the total liability and has a subjective estimate, ν, of the variance of total volume Σm, when the length of investigation is t. (Σm itself is not an argument of V, since it is unobservable; the mean value of Σm, $\Sigma \bar{m}$, is held constant throughout and is hence suppressed.) We assume that V is up to third order continuously differentiable in all three of its arguments.

We assume that V is increasing in both k and ν. The first relationship reflects the fact that PRP's with higher liability shares bear a larger share of the total cleanup burden. The second reflects both PRP risk aversion—intuitively, uncertainty raises the *expected* cost for risk averse decision makers.[8]—and the inherent characteristics of the cleanup cost function. Specifically, since the cleanup method has to be specified at the end of the investigation period, an increase in uncertainty increases the probability that a given method will prove inadequate and that the cleanup will have to be repeated.[9] Since the cost of repeating the cleanup is much higher than the cost saved from choosing a less expensive cleanup method, expected costs increase with uncertainty. Thus even if a PRP were risk neutral, its expected utility loss, V, would increase with ν

The relationship between V and the length of the investigation period, t, is less straightfoward. A longer investigation period reduces the level of uncertainty and delays the commencement of

[7] Of course, the notion that a PRP would declare a number representing the variance of its contribution is no more than a convenient abstraction. In reality, a PRP's report might consist of an upper and lower bound to its contributions. If this were the case, then a high (near η) value of s_i would be interpreted as a wide (negligible) gap between the two reported bounds.

[8] See RSZ1 for details in specifying the expected utility losses for risk averse PRPs.

[9] This point is developed in Zimmerman (1988)

cleanup activities, thus reducing the present value of cleanup costs. Both of these effects lead to a reduction in V. On the other hand, delays in the commencement of cleanup proceedings increase the period during which surrounding areas are exposed to toxic substances. Moreover, if these substances are migrating over time, cleanup delays will also increase the total area exposed to contamination. Reflecting these considerations, we assume that V is convex in t and that for any values of ν and k, an optimal investigation time exists.

We assume that $V_{t\nu} < 0$, i.e., that the greater is uncertainty, the greater is the marginal benefit from extending the investigation period. This condition would be satisfied, for example, if investigation reduced uncertainty proportionally. It has the (natural implication that the optimal investigation period increases with uncertainty about the degree of contamination.

We assume that $V(t, \nu, k)$ can be decomposed into two elements, $V(t, \nu, k) = U(t, \nu, k)d(k)$ with $d(\cdot) > 0$ and $U_{tk} < 0$. This is a sufficient condition for $-V$ to satisfy the "single crossing property (SCP) in (t, k)" (see Milgrom and Shannon (1994)): for all t, if $V_t(t, k_L) < 0$ and $k_L < k_H$, then $V_t(t, k_H) < 0$. From Milgrom and Shannon (1994), the SCP is a necessary and sufficient condition for the conclusion that a PRP's optimal investigation length increases in its liability share. PRPs with larger liability shares prefer longer investigation time for multiple reasons. If they are risk averse, the effect of uncertainty on expected cost is proportional to k^2, implying a longer optimal investigation period (see RSZ1 for a specific example). This point is also made by Zimmerman (1988), who points out that larger PRPs prefer less uncertainty because they will be held responsible for a larger share of the of the Superfund cleanup. Again, k enters the payoff function more than proportionally.

Finally, we assume that the regulator's response function $t(\cdot)$ is positive and thrice continuously differentiable. Further, we assume that both $t(\cdot)$ and $t'(\cdot)$ increase with Σs, while $t''(\cdot)$ does not decrease with Σs.

Let $W((\eta, \boldsymbol{s}_{-i}), (\theta_i, \boldsymbol{\theta}_{-i}), k_i) \equiv V(t(\eta + \Sigma s_{-i}(\boldsymbol{\theta}_{-i})), (\theta_i, \boldsymbol{\theta}_{-i}), k_i)$ be the cost to PRP i when the type-profile is $(\theta_i, \boldsymbol{\theta}_{-i})$, when i announces η and other PRP's are playing $\boldsymbol{s}_{-i}(\cdot)$. Hence $\int_{\boldsymbol{\Theta}_{-i}} -W((\eta, \boldsymbol{s}_{-i}), (\theta_i, \boldsymbol{\theta}_{-i}); k_i) \boldsymbol{g}_{-i}(\boldsymbol{\theta}_{-i}) d\boldsymbol{\theta}_{-i}$ is i's *expected payoff function* in the game between PRP's.

Define the function $f : \mathbb{R}^3 \to \mathbb{R}$ by $f(\eta + \boldsymbol{s}_{-i}(\boldsymbol{\theta}_{-i}), \Sigma\theta, k_i) = \frac{1}{d(k_i)} \frac{dW((\eta, \boldsymbol{s}_{-i}), (\theta_i, \boldsymbol{\theta}_{-i}), k_i)}{d\eta}$. f is a measure of the incremental marginal cost to i of increasing his announcement, given the realized types of all other PRP's. If $\boldsymbol{s}$ is a pure-strategy equilibrium profile, then PRP i's strategy must satisfy the following condition: for each θ_i, $s_i(\theta_i)$ must minimize $E_{\boldsymbol{\theta}_{-i}} W((\eta, \boldsymbol{s}_{-i}), (\theta_i, \boldsymbol{\theta}_{-i}); k_i)$, on $[\underline{\eta}, \bar{\eta}]$. That is, a necessary condition for $\boldsymbol{s}$ to be an equilibrium profile is that:

$$\text{for all } i \text{ and all } \theta_i \text{ ,} 0 \leq \int_{\boldsymbol{\Theta}_{-i}} f\left(s_i(\theta_i) + \Sigma s_{-i}(\boldsymbol{\theta}_{-i}), \theta + \Sigma\theta_{-i}, k_i\right) \boldsymbol{g}_{-i}(\boldsymbol{\theta}_{-i}) d\boldsymbol{\theta}_{-i} \qquad (1)$$

$$\text{with equality holding whenever } s_i(\theta_i) > \underline{\eta}.$$

3. Monotone Pure Strategy Nash Equilibrium

In this section, we establish that our incomplete information game has a pure strategy Nash equilibrium (PSNE), using the methodology introduced in Athey (1997). Athey identifies conditions under which there exists a PSNE satisfying a natural monotonicity property: each agent's strategy is a nondecreasing function of its "type." In our case, PRP i's type, θ_i, is the variance of its contribution, m_i, to the contamination of the site in question. The monotonicity property of the

PSNE is thus intuitive: other things equal, PRP i would prefer a longer investigation period (i.e. report a higher s_i) when θ_i increases.

Athey's existence theorem (Theorem 3.1 in Athey (1997)) states that if payoffs are continuous and satisfy a natural integrability condition, if type distributions are nonatomic and if player i's expected payoff satisfies SCP in (η_i, θ_i) *provided that all other players' strategies are nondecreasing in their types*, then a pure-strategy Nash equilibrium exists in which each player's strategy is nondecreasing in its type. In our case, Athey's qualified SCP condition is trivially satisfied: regardless of the strategies chosen by other players, PRP i's expected payoff function $-\int_{\Theta_{-i}} W((\eta, \boldsymbol{s}_{-i}), (\theta_i, \boldsymbol{\theta}_{-i}); k_i)\boldsymbol{g}_{-i}(\boldsymbol{\theta}_{-i})d\boldsymbol{\theta}_{-i}$ satisfies:

$$\frac{d^2(-E_{\boldsymbol{\theta}_{-i}}W)}{d\eta d\theta_i} = -E_{\boldsymbol{\theta}_{-i}}\frac{d^2V}{d\theta dt}\frac{dt}{d\Sigma s} > 0 \tag{2}$$

since $\frac{dt}{d\Sigma s}$ is positive and $V_{t\nu} < 0$. Thus $-E_{\boldsymbol{\theta}_{-i}}W$ satisfies SCP in (η, θ_i), leading to the existence of a PSNE.

We can also establish that the PSNE satisfies certain uniform differentiability conditions. For the applications that follow, it is important that the bounds we establish below hold uniformly across all possible distributions of agents' types.

Proposition 1 (Existence of a monotone PSNE). *Every game has a pure-strategy Nash equilibrium. Moreover, for any pure strategy profile $\boldsymbol{s}$ and any i, there exists $\tilde{\theta}_i \geq \theta^l$ such that s_i equals η on $[\theta^l, \tilde{\theta}_i)$, and strictly exceeds η on $(\tilde{\theta}_i, \theta^u]$. Moreover, s_i is increasing and continuously differentiable on $(\tilde{\theta}_i, \theta^u]$.*

4. Liability shares and the tendency to over-report variance

Proposition 1 established the existence of an equilibrium in which players' announcements are monotone with respect to their types. We now establish that in any such equilibrium, players' strategies will be monotone with respect to their liability shares. That is, players who bear more responsibility for the cleanup will have a greater tendency to over-report than players with smaller shares. While the proof of this result is quite technical, the basic idea is straightforward. Loosely, PRP's with higher liability shares are more sensitive to changes in the level of uncertainty than those with smaller shares (see Lemma 1.1 below). Hence those with higher shares will, on the margin, prefer longer investigation periods, and will, therefore, be more inclined to overreport in order to induce the regulator to select a higher level of t. Our task in this section is to formalize this intuition. We first identify certain properties of the function f measuring the marginal benefit of an announcement for a given PRP, when the types and strategies of other PRPs are given.

Lemma 1.1. *(a) For each i, $\boldsymbol{s}$ and each $\boldsymbol{\theta}$, $f(\eta + \boldsymbol{s}_{-i}(\boldsymbol{\theta}_{-i}), \Sigma\theta, k_i)$ decreases w.r.t. k_i. (b) There exists $\epsilon > 0$ such that if either $|V_t| < \epsilon$ or $|t''(\cdot)| < \epsilon$, then $f(\eta + \boldsymbol{s}_{-i}(\boldsymbol{\theta}_{-i}), \Sigma\theta, k_i)$ increases in η.*

Part (a) simply says that an increase in an agent's liability share raises the marginal benefit of over-reporting when a PRP knows the types of all other PRPs. That is, higher liability shares lead to higher reports if $\boldsymbol{\theta}_{-i}$ is known. Part (b) provides a sufficient condition for a result that we need, namely, that W is convex in η. The condition is that *either* the regulator's response function is sufficiently close to linear in the PRPs' aggregate report, *or* that the the distribution of types is sufficiently concentrated (that is, for each $j \neq i$, the variance of g_j must be sufficiently small).[10]

The conditions in (b) are quite restrictive; they are, however, only sufficient, not necessary, for the convexity of W. We will henceforth simply assume (Assumption 1 below) that W is convex in η. This assumption ensures that if a PRP knew for certain the types of all other PRP's, the Kuhn-Tucker conditions would be both necessary and sufficient for cost minimization.

[10] To see this, observe that if the variances of the g_j's were zero, then in any pure-strategy equilibrium V_t would be zero with probability one. By continuity, for any positive ϵ there must exist $\delta > 0$ such that if the variance of g_j is less than δ, then the absolute value of V_t will be less than ϵ with probability $1 - \epsilon$.

Assumption 1. *For each i, $\frac{1}{d(k_i)}\frac{dW}{d\eta}$ increases with η.*

Now consider a given monotone PSNE strategy profile. Lemma 1.1 implies the following result relating the strategies of any two PRPs i and j, with $k_i > k_j$: at any point θ at which the difference between $s_i(\cdot)$ and $s_j(\cdot)$ is minimized, the constraint that $s_j \geq \underline{\eta}$ is binding.

Lemma 1.2. *Assume that $k_i > k_j$. Let $\Theta_{ij} = \{\theta \in [\theta^l, \theta^u] : \big(s_i(\theta) - s_j(\theta)\big) = \min_{\vartheta \in [\theta^l, \theta^u]} \big(s_i(\vartheta) - s_j(\vartheta)\big)\}$. (Since s_i and s_j are continuous [Proposition 1], the set Θ_{ij} is closed.) For all $\theta \in \Theta_{ij}$, there exists a neighborhood U of θ such that $s_j(\cdot) = \underline{\eta}$ on U.*

This Lemma has three immediate implications: (a) at any value of θ such that $s_i(\theta) > \underline{\eta}$, $s_i(\theta) > s_j(\theta)$; (b) there exists an interval on which $s_i(\cdot) > \underline{\eta}$ but $s_j(\cdot) = \underline{\eta}$ and (c) only those agents' whose liability shares weakly exceed all other PRPs' shares report numbers that strictly exceed $\underline{\eta}$ with probability one.

Proposition 2 follows almost immediately Lemma 1.2. It states that in a PSNE, PRPs with larger liability shares have a greater tendency to over-report their uncertainty than smaller PRPs.

Proposition 2 (Monotonicity w.r.t. liability shares). *Let s be a PSNE strategy profile. If $k_i > k_j$ and if $\tilde{\theta}_i < \theta^u$, then (a) $\tilde{\theta}_i < \tilde{\theta}_j$ and (b) $s_i(\cdot) > s_j(\cdot)$ on $(\tilde{\theta}_i, \theta^u]$,*

5. Heterogeneity and Delay

In the preceding section we compared the strategies of different PRPs within a given game. In particular, Proposition 2 above suggests that large PRPs tend to over-report while small PRPs tend to under-report. In this section, we extend this result by comparing games in which PRPs' liability shares are different. We begin by showing that in a game with only two PRPs, an increase in a given PRP's liability share results in an increase in the report it makes. This result cannot, however, be extended in general to the n-PRP case. Moreover, even when there are only two PRPs, we cannot make any general statements about the *net* effect on expected delay of a change in liability shares. We can, however, show that if certain kinds of conditions are satisfied then a transfer of liability from small PRP's to larger ones will result in an increase in the expected investigation period. Propositions 4 and 5 below identify two such sets of conditions.

Assume that there are only two PRPs, i and j, and consider two liability profiles (k_i, k_j) and (k'_i, k'_j), with $k'_i > k_i$ and $k_j > k'_j$. We will show that in the latter profile, PRP i's announcement will be higher, and that of PRP j will be lower than in the earlier profile, except when either is restricted by the lower bound $\underline{\eta}$ of the action space. Note that this result holds regardless of whether k_i is larger or smaller than k_j.

Proposition 3 (Heterogeneity: two PRPs). *Consider two liability profiles (k_i, k_j) and (k'_i, k'_j) such that $k'_i > k_i$ and $k'_j < k_j$. Then $\tilde{\theta}'_i < \tilde{\theta}_i$ and $s'_i(\theta) > s_i(\theta)$ for all $\theta > \tilde{\theta}'_i$. Similarly, $\tilde{\theta}'_j > \tilde{\theta}_j$ and $s'_j(\theta) < s_i(\theta)$ for all $\theta > \tilde{\theta}_j$.*

For some intuition for this result, observe that as a PRP's liability share increases, its preferred investigation length increases also, so that *holding all other PRPs' strategies constant*, an increase in k_i will result in an increase in i's announcement. Similarly, holding other strategies constant, a decrease in k_j must result in a decrease in j's announcement. Now if $s'_i(\theta) \leq s_i(\theta)$ for some $\theta \geq \tilde{\theta}_i$, it must be the case that PRP j's announcement has increased for some θ_j. Indeed, the largest decrement in i's announcement (which depends on i's type) must be more than offset by the largest increment in j's announcement. On the other hand, if $s'_j(\theta) > s_j(\theta)$ the largest increment in j's announcement must be more than offset by the largest decrement in i's announcement. But these two requirements are mutually contradictory and the proposition now follows.

We now consider the effects of increasing the degree of hetorogeneity among PRP's. Specifically, we will compare two liability profile $\boldsymbol{k}$ and $\boldsymbol{k}'$, with the property that liability shares in the latter profile are more dispersed. The preceding result established that when there are only two PRPs, if the larger PRPs' liability shares become larger it will over-report more, while the smaller PRP who share has diminished will over-report less. We now address the issue of which of these effects dominates. That is, what will be the *net* effect of increased heterogeneity on the expected average report and hence the expected length of investigation?

In general, the answer to this question depends on the second order characteristics of the PRPs' equilibrium strategies. Roughly speaking, the issue is whether on average a PRP's announcement changes with its liability share at an *increasing* or *decreasing* rate. In general, it is not possible to answer this question definitively without an extensive specification of the third order derivatives of the PRPs' payoff functions. We can, however, obtain a determinate result under certain special circumstances. If these third order derivatives are relatively insignificant, if the regulator's response function is sufficiently close to linear and if the PRPs' liability share are initially sufficiently heterogeneous, then a further increase in heterogeneity will increase the expected length of investigation.

The key to this result is that the smaller PRP's strategic options are significantly limited by the fact that the action space is bounded below by $\underline{\eta}$. To see this, observe once again that as a PRP's liability share increases (decreases), its preferred investigation length increases (decreases) also. Any decrease in the smaller PRP's announcement can be more than offset by the larger PRP, who can simply increase its own announcement. The reverse is not true, however: the smaller PRP cannot offset the larger ones' increased announcement, because the smallest possible announcement it can make is $\underline{\eta}$. (More concretely, there is no natural upper bound to the level of uncertainty one can profess. One cannot, however, assert that the level of uncertainty is negative!)

To make the above ideas precise, we need some preliminary constructions. Fix two liability profiles $\boldsymbol{k}$ and $\boldsymbol{k}'$, with $k_1' > k_1 > k_2$. Define the function $\boldsymbol{\kappa} : [0,1] \to \mathbb{R}^2$ by, for $\alpha \in [0,1]$, $\boldsymbol{\kappa}(\alpha) = \alpha\boldsymbol{k}' + (1-\alpha)\boldsymbol{k}$. For each α, let $\boldsymbol{s}(\cdot;\alpha)$ be a PSNE corresponding to the liability share $\boldsymbol{\kappa}(\alpha)$. For each $r \in \{1,2\}$ and $\alpha \in [0,1]$, define $\tilde{\theta}_r(\alpha)$ as follows (cf. the definition of $\tilde{\theta}$ in Proposition 1.) if there exists $\theta_r' \in \Theta$ satisfying equation (3) below, set $\tilde{\theta}_r(\alpha)$ equal to θ_r'; otherwise set $\tilde{\theta}_r(\alpha)$ equal to θ^u.

$$0 \quad = \quad \int_\Theta \left\{ V_t\Big(t\big(\underline{\eta} + s_{\neg r}(\vartheta_{\neg r},\alpha)\big), (\theta_r', \vartheta_{\neg r}), \kappa_r(\alpha)\Big) \frac{d\left[t\big(\underline{\eta} + s_{\neg r}(\vartheta_{\neg r},\alpha)\big)\right]}{d\eta} \right\} g(\vartheta_{\neg r}) d\vartheta_{\neg r} \qquad (3)$$

where $\neg r = 2$, if $r = 1$, and 1 if $r = 2$. We will assume that for $\alpha \in (0,1)$ and each θ *except* $\tilde{\theta}_r(\alpha)$, $s_r(\theta;\alpha)$ is differentiable w.r.t. α. From Prop 3, we know that $\tilde{\theta}_1(\alpha)$ decreases, while $\tilde{\theta}_2(\alpha)$ increases with α. To focus on the more interesting case, we choose $\boldsymbol{k}'$ so that $\tilde{\theta}_1(\alpha) > \theta^l$, for all $\alpha < 1$, but $\tilde{\theta}_1(1) = \theta^l$. Next, let $Et(\alpha) = \int_{\Theta^2} t(\Sigma s(\boldsymbol{\vartheta};\alpha))\boldsymbol{g}(\boldsymbol{\vartheta})d\boldsymbol{\vartheta}$ denote the expected length of the investigation period associated with the PSNE $\boldsymbol{s}(\cdot;\alpha)$. Finally, given any continuous function f mapping a compact set to $\mathbb{R}_{++}$, we will say that f is ϵ-flat if for some positive scalar $\underline{f}$, $f(\cdot) \in [\underline{f}, (1+\epsilon)\underline{f}]$, We can now state the above result formally:

Proposition 4 (Heterogeneity and delay: two PRPs). *There exists $\epsilon_V > 0$, $\epsilon_t > 0$ such that if (a) V_{tt} is ϵ_V-flat*[11] *and (b) $\frac{\max\{t''(\Sigma s): s_r \in H\}}{\min\{t'(\Sigma s): s_r \in H\}} < \epsilon_t$ then for some $\bar{\alpha} < 1$, $\frac{dEt(\cdot)}{d\alpha}$ is positive on $[\bar{\alpha}, 1]$.*

While the proof of Proposition 4 is somewhat complex, the basic idea is quite straightforward. The conclusion of the proposition can be rewritten as:

$$0 < \int_{\theta^l}^{\theta^u} \left\{ \int_\Theta t'(s_1(\vartheta_1,\alpha) + s_2(\vartheta_2,\alpha)) \left(\frac{d\left[s_1(\vartheta_1,\alpha)\right]}{d\alpha} + \frac{d\left[s_2(\vartheta_2,\alpha)\right]}{d\alpha} \right) g(\vartheta_2) d\vartheta_2 \right\} g(\vartheta_1) d\vartheta_1. \quad (4)$$

[11] Recall from page 9 that V is assumed to be convex.

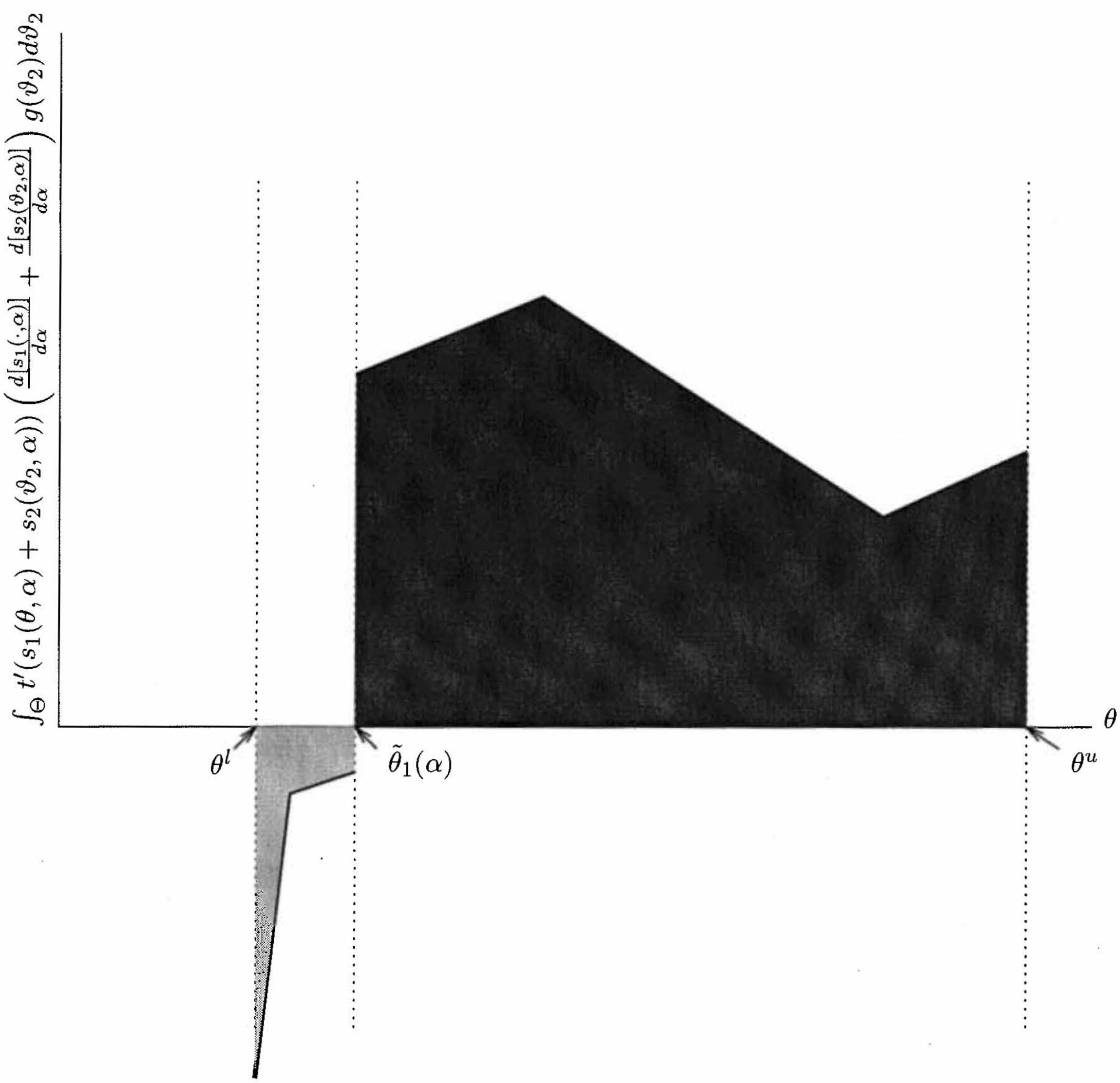

FIGURE 1. Total shaded area must be positive

This inequality corresponds graphically to the requirement that the lightly shaded area below the axis in figure 1 be dominated by the more heavily shaded area above the axis. To prove that this requirement is satisfied, we establish two properties of the graph: (a) to the right of $\tilde{\theta}_1(\alpha)$, it is positive; and (b) to the left of this point, it is bounded below. Inequality (4) follows from these properties, provided that $\tilde{\theta}_1(\alpha)$ is sufficiently close to θ^l.

More specifically, note from (1) and (3) that for each α and $\theta_1 > \tilde{\theta}_r(\alpha)$,

$$0 \quad = \quad \int_\Theta \left\{ V_t\Big(t\big(s_1(\theta_1,\alpha)+s_2(\vartheta_2,\alpha)\big),(\theta_1,\vartheta_2),\kappa_1(\alpha)\Big)\frac{d\left[t\big(s_1(\theta_1,\alpha)+s_2(\vartheta_2,\alpha)\big)\right]}{d\Sigma s}\right\} g(\vartheta_2)d\vartheta_2.$$

Hence for each $\theta_1 > \tilde{\theta}_1(\alpha)$:

$$0 \quad = \quad \frac{d}{d\alpha}\left[\int_\Theta \left\{ V_t\Big(t\big(s_1(\theta_1,\alpha)+s_2(\vartheta_2,\alpha)\big),(\theta_1,\vartheta_2),\kappa_1(\alpha)\Big)\frac{d\left[t\big(s_1(\theta_1,\alpha)+s_2(\vartheta_2,\alpha)\big)\right]}{d\Sigma s}\right\} g(\vartheta_2)d\vartheta_2\right]$$

Now $\kappa_1(\alpha)$ increases with α and V_t decreases with κ. It follows that for each $\theta_1 > \tilde{\theta}_1(\alpha)$, there exists $\delta_1(\theta_1,\alpha) > 0$ such that:

$$\begin{aligned}
\delta_1(\theta_1,\alpha) \quad &= \quad -\int_\Theta \left\{ \frac{\partial}{\partial k}\left[V_t\Big(t\big(s_1(\theta_1,\alpha)+s_2(\vartheta_2,\alpha)\big),(\theta_1,\vartheta_2),\kappa_1(\alpha)\Big)\frac{d\left[t\big(s_1(\theta_1,\alpha)+s_2(\vartheta_2,\alpha)\big)\right]}{d\Sigma s}\right] \times \right.\\
&\qquad \left.\frac{d\kappa_1(\alpha)}{d\alpha}\right\} g(\vartheta_2)d\vartheta_2 \\
&= \quad \int_\Theta \left\{ \frac{\partial}{\partial t}\left[V_t\Big(t\big(s_1(\theta_1,\alpha)+s_2(\vartheta_2,\alpha)\big),(\theta_1,\vartheta_2),\kappa_1(\alpha)\Big)\frac{d\left[t\big(s_1(\theta_1,\alpha)+s_2(\vartheta_2,\alpha)\big)\right]}{d\Sigma s}\right] \times \right.\\
&\qquad \left.\left(\frac{d\left[s_1(\theta_1,\alpha)\right]}{d\alpha} + \frac{d\left[s_2(\vartheta_2,\alpha)\right]}{d\alpha}\right)\right\} g(\vartheta_2)d\vartheta_2 \qquad (5)
\end{aligned}$$

Moreover continuity and compactness ensure that $\delta_1(\cdot,\cdot)$ is both bounded away from zero and bounded above on $\left\{(\alpha,\theta) : \theta \in [\tilde{\theta}_1(\alpha),\theta^u]\right\}$. If $\frac{\partial}{\partial t}\left[V_t(\cdot,\cdot,\cdot)t'(\cdot)\right] = (V_{tt}t' + V_t t'')$ *were constant* (and positive), then certainly $\int_\Theta \left(\frac{d[s_1(\cdot,\alpha)]}{d\alpha} + \frac{d[s_2(\vartheta_2,\alpha)]}{d\alpha}\right) g(\vartheta_2)d\vartheta_2$ would be bounded away from zero on the interval $[\tilde{\theta}_1(\alpha),\theta^u]$. Our assumptions—i.e., V_{tt} is sufficiently flat and t'' is sufficiently small relative to t'—together with continuity ensure that this integral is positive on the required interval. It follows that if $t'(\cdot)$ is sufficiently flat, then on the same interval:

$$0 < \int_\Theta t'(s_1(\cdot,\alpha)+s_2(\vartheta_2,\alpha))\left(\frac{d\left[s_1(\cdot,\alpha)\right]}{d\alpha} + \frac{d\left[s_2(\vartheta_2,\alpha)\right]}{d\alpha}\right) g(\vartheta_2)d\vartheta_2, \qquad (6)$$

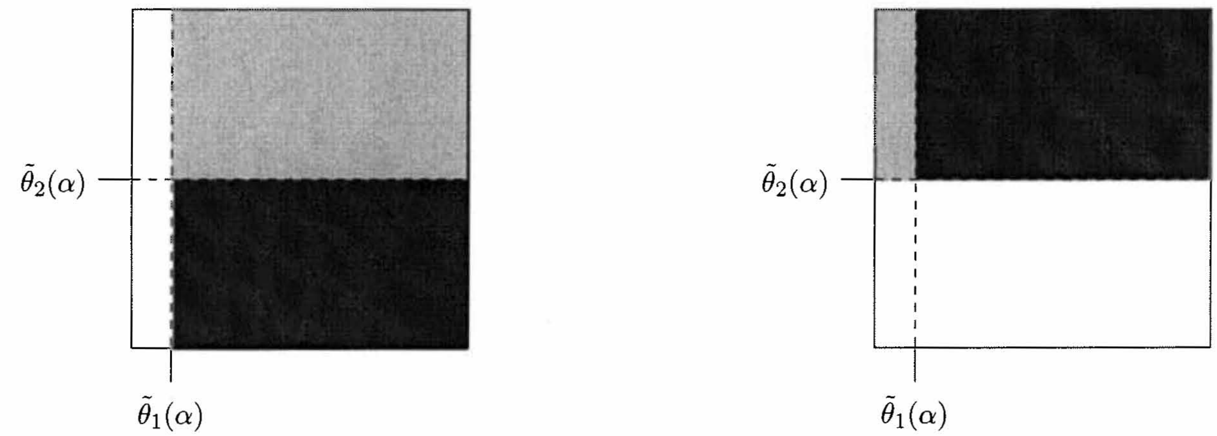

FIGURE 2. Intuition for Lemma 4.2.

as depicted in figure 1. Now since $s_1(\cdot,\alpha) \equiv \underline{\eta}$ on $[\theta^l, \tilde\theta_1)$, $\frac{d[s_1(\cdot,\alpha)]}{d\alpha}$ is identically zero on $[\theta^l, \tilde\theta_1(\alpha))$. Moreover, $\tilde\theta_1(\alpha) \searrow \theta^l$ as $\alpha \nearrow 1$. Hence (4) will follow from (6), for α sufficiently close to 1, *provided that* $\frac{d[s_2(\cdot,\alpha)]}{d\alpha}$ is bounded below by a number that is independent of α. To establish this fact, however, requires a considerable amount of work.

The formal argument that $\frac{d[s_2(\cdot,\alpha)]}{d\alpha}$ is bounded requires two steps (see lemma 4.2). First, in a "preliminary step", we show that for $r = 1, 2$, if $\frac{d[s_r(\cdot,\alpha)]}{d\alpha}$ is *not* bounded *above* (resp. not bounded *below*) independently of α, then for any given α, $\frac{d[s_r(\cdot,\alpha)]}{d\alpha}$ is a nonnegative (resp. nonpositive), function of ϑ_r. Boundedness then follows from the following argument. Assume that for some sequence $\{\alpha^m\}$, $\sup\frac{d[s_1(\cdot,\alpha^m)]}{d\alpha}$ increases without bound. Since $\delta_1(\cdot,\cdot)$ in (5) *is* bounded above, this assumption implies that $\frac{d[s_2(\cdot,\alpha^m)]}{d\alpha}$ must *decrease* without bound on some open subset of Θ. Since $\frac{\partial}{\partial t}\left[V_t(\cdot,\cdot,\cdot)t'(\cdot)\right]$ is nearly constant, it follows from the preliminary step and the boundedness of $\delta_1(\cdot,\cdot)$ in (5) that as $m \to \infty$, the ratio of $\left|E\left[\frac{d[s_2(\cdot,\alpha^m)]}{d\alpha}|\theta_1 > \tilde\theta_1(\alpha^m)\right]\right|$ to $\left|E\left[\frac{d[s_1(\cdot,\alpha^m)]}{d\alpha}|\theta_1 > \tilde\theta_1(\alpha^m)\right]\right|$ must converge to a number close to unity.

At this point, consider the left panel of figure 2. The square represents $[\theta^l, \theta^u]^2$. It is important to note (see Proposition 3) that since $\tilde\theta_1(\alpha) \searrow \theta^l$ as $\alpha \nearrow 1$, $\tilde\theta_2(\cdot)$ must be bounded away from θ^l, as depicted in the figure. Now, we established above that the integrals of $\frac{d[s_1(\cdot,\alpha^m)]}{d\alpha}$ and $\frac{d[s_2(\cdot,\alpha^m)]}{d\alpha}$

on the shaded region to the right of $\tilde{\theta}_1(\alpha)$ must roughly offset each other. But $\frac{d[s_2(\cdot,\alpha^m)]}{d\alpha}$ is zero on the region below $\tilde{\theta}_2(\alpha)$. Hence $\left|\frac{d[s_2(\cdot,\alpha^m)]}{d\alpha}\right|$ must be significantly larger on average than $\left|\frac{d[s_1(\cdot,\alpha^m)]}{d\alpha}\right|$ on the cross-hatched region *above* $\tilde{\theta}_2(\alpha)$ and to the right of $\tilde{\theta}_1(\alpha)$. However, by an exactly parallel argument, which starts by reversing 1's and 2's in expression (5), the integrals of $\frac{d[s_1(\cdot,\alpha^m)]}{d\alpha}$ and $\frac{d[s_2(\cdot,\alpha^m)]}{d\alpha}$ must also roughly offset each other on the shaded region above $\tilde{\theta}_2(\alpha)$ (see the right panel of figure 2). Since $\frac{d[s_1(\cdot,\alpha^m)]}{d\alpha}$ is zero to the left of $\tilde{\theta}_1(\alpha)$, this requirement implies that $\left|\frac{d[s_2(\cdot,\alpha^m)]}{d\alpha}\right|$ *is not*, on average, significantly larger than $\left|\frac{d[s_1(\cdot,\alpha^m)]}{d\alpha}\right|$ on the region above $\tilde{\theta}_2(\alpha)$ and to the right of $\tilde{\theta}_1(\alpha)$. But now we have reached a contradiction, which completes the proof.

6. Multiple PRPs and the effect of heterogeneity

Matters are more complex when there are multiple PRP's. Accordingly, we will analyze the special case in which the social and private utility loss functions are quadratic forms in investigation time and the sum of agents' private uncertainty parameters. This specification considerably oversimplifies the nature of our problem, ignoring, in particular, factors such as time discounting and all but one dimension of agent heterogeneity. Its offsetting benefit is that the resulting solution is quite transparent, providing insights into the structure of our model. In addition, the example demonstrates that under certain conditions the two-player heterogeneity result obtained in the previous subsection can be generalized. Finally, as will be clear from the structure of its proof, the conclusion of Proposition 5 below will hold more generally, provided that third-order effects are sufficiently small relative to lower-order effects.

Let Y denote the loss in social welfare due to cleanup when the length of the investigation period is t and the vector of PRP types is $\boldsymbol{\theta} = (\theta_1, ..., \theta_n)$, where $\theta_i \in \Theta = [\theta^l, \theta^u]$:

$$Y(t, \boldsymbol{\theta}) \quad = \quad 0.5\beta_{11}t^2 \ - \ \beta_{12}t\Sigma\theta \ + \ \beta_{22}\left(\Sigma\theta\right)^2$$

$$\text{where } \beta_{11}, \ \beta_{12}, \ \beta_{22} > 0.$$

The first term measures the welfare loss associated with the expected cost of cleanup, while the remaining terms reflect the loss arising from uncertainty over the magnitude of the cleanup task. That is, β_{12} may be interpreted as a measure of social risk aversion.[12] Note that $Y_t, Y_{tt} > 0$ while $Y_{t\Sigma\theta} < 0$. Clearly $Y(\cdot, \boldsymbol{\theta})$ is minimized at $t^*(\Sigma\theta) = \beta_{12}\Sigma\theta/\beta_{11}$. It follow that $Y_{\Sigma\theta}$ will be positive for all t in the relevant range if $\theta^l \geq 1$ and $\beta_{22} > n\beta_{12}^2\theta^u/\beta_{11}$.

As discussed on page 4 above, we model our regulator as assuming that agents' type announcements are truthful. That is, the regulator in fact minimizes the following function of Σs:

$$Y(\Sigma s, \boldsymbol{\theta}) \quad = \quad 0.5\beta_{11}t^2 \ - \ \beta_{12}t\Sigma s \ + \ \beta_{22}\left(\Sigma s\right)^2$$

setting $t^*(\Sigma s) = \beta_{12}\Sigma s/\beta_{11}$. It follows that for each i, $\frac{dt^*(\Sigma s)}{ds_i} = \beta_{12}/\beta_{11}$.

Now consider the problem facing PRP i, whose liability share is $k_i \in (0, 1)$. We assume that the PRPs' expected loss functions have the same form as those of the regulator, except that the parameters β_{11}, β_{12} and β_{22} are replaced by γ_{11}, γ_{12} and γ_{22}. In particular, we assume that $\gamma_{11} < \beta_{11}$ and $\gamma_{22} > \beta_{22}$, reflecting the fact that PRPs are likely to overvalue the benefits of uncertainty reduction and to undervalue the costs of extending the investigation period beyond its socially optimal length (See page 3). On the other hand, each PRP is responsible for only a fraction

[12] Note that β_{12} might be positive even if society were risk neutral: for example, if the cleanup task proves more difficult than expected, costly extensions to the original schedule might be required, while if it proves less onerous than expected, it may prove too late to adopt a less expensive cleanup strategy. See Rausser et al. (1998) for a further discussion.

of total cleanup costs: firm i bears a fraction k_i of expected social cost but the variance of i's cost share is only k_i^2 times the variance of social costs. Hence, taking as given the regulator's response function, $t^*(\cdot)$ and the sum of other agents' strategies, $\Sigma s_{-i}(\cdot)$, PRP i's expected utility loss is

$$V(t^*(\Sigma s), \boldsymbol{\theta}, k_i) \;=\; 0.5 k_i \gamma_{11} \left(t^*(\Sigma s)\right)^2 \;-\; k_i^2 \gamma_{12} t^*(\Sigma s) \Sigma\theta \;+\; k_i^2 \gamma_{22} \left(\Sigma\theta\right)^2 \tag{7}$$

We will assume that even if all liability shares are equal (i.e., when $k_i = 1/n$, for all i), the expected investigation period that results when firms act strategically exceeds the socially optimal investigation period.

PRP i's strategy is a map $s_i : \Theta \to [\underline{\eta}, \bar{\eta}]$. We assume that $\underline{\eta} = 0$. Given the mapping $\Sigma s_{-i}(\cdot)$ from Θ to the sum of other agents strategies, $s_i(\theta)$ must satisfy the first order condition:

$$0 \;\leq\; \int_{\Theta} \left[\left(k_i \gamma_{11} t^* \left(s_i(\theta_i) + \Sigma s_{-i}(\boldsymbol{\theta}_{-i})\right) \;-\; k_i^2 \gamma_{12} (\theta_i + \Sigma\theta_{-i})\right) \frac{dt^*(\Sigma s)}{ds_i} \right] \boldsymbol{g}_{-i}(\boldsymbol{\theta}_{-i}) d\boldsymbol{\theta}_{-i}$$

with equality holding whenever $s_i(\theta_i) > \underline{\eta}$.

Dividing through by k_i, the constant $\frac{dt^*(\cdot)}{ds_i} = \beta_{12}/\beta_{11}$ and the constant γ_{12}/γ_{11}, and substituting for $t^*(\cdot)$, we obtain

$$0 \;=\; \int_{\Theta} \left[s_i(\theta_i) + \Sigma s_{-i}(\boldsymbol{\theta}_{-i}) \;-\; k_i \frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}} (\theta_i + \Sigma\theta_{-i}) \right] \boldsymbol{g}_{-i}(\boldsymbol{\theta}_{-i}) d\boldsymbol{\theta}_{-i}, \text{ for all } \theta_i \geq \tilde{\theta}_i \tag{8}$$

so that, defining $\tilde{\theta}_i$ implicitly by $k_i \frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}} (\tilde{\theta}_i + E_\Theta \Sigma\theta_{-i}) \;=\; \underline{\eta} + E_\Theta \Sigma s_{-i}(\boldsymbol{\theta}_{-i})$, i's strategy can be written as:

$$s_i(\theta_i) \;\dot{=}\; \begin{cases} \underline{\eta} & \text{if } \theta_i \leq \tilde{\theta}_i \\ k_i \frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}} (\theta_i + E_\Theta \Sigma\theta_{-i}) \;-\; E_\Theta \Sigma s_{-i}(\boldsymbol{\theta}_{-i}) & \text{if } \theta_i > \tilde{\theta}_i \end{cases}$$

The following proposition establishes that under this simple specification, a shift in liability from smaller PRPs to larger ones results in an increase in the expected investigation period. Since by assumption, the investigation period is too long even when firms' liability shares are equal, the liability shift further exacerbates the delay.

Proposition 5 (Heterogeneity and delay: multiple PRPs). *Consider a liability profile,* $\mathbf{k}$*, such that not all liability shares are equal and* $i > j$ *implies* $k_i \geq k_j$*. Now consider a shift in liability shares to* $\mathbf{k}' \neq \mathbf{k}$*. Assume that for some* $\bar{j} \in \{2, ..., n\}$*,* $k'_i \geq k_i$*, for all* $i > \bar{j}$ *and* $k'_i \leq k_i$ *otherwise. Let* $\mathbf{s}$ *and* $\mathbf{s}'$ *be PSNE's corresponding to these profiles. Then* $E_{\Theta^n}\ t(\Sigma s'(\boldsymbol{\theta})) > E_{\Theta^n}\ t(\Sigma s(\boldsymbol{\theta}))$*.*

7. Policy Analysis

In this section we apply our theoretical framework to evaluate two widespread practices. In particular, we consider the effect on cleanup delay of *de minimis* PRP buyouts and the formation of steering committees by PRP's. Each of these practices has been widely advocated as an effective way to reduce litigation and transaction costs, and thereby expedite the cleanup process. Our model suggests that each may have side-effects that have been hitherto ignored. It must be emphasized that the formal results we obtain in this section are valid only under the restrictive conditions we specify in section 6. As we noted above, however, the key comparative statics result we obtain in that section, Proposition 5, will hold more generally, provided that third-order effects do not dominate lower-order effects.

7.1. ***de minimis* PRP buyouts.** A PRP is classified as *de minimis* if its liability share falls below some (small) critical fraction. In a *de minimis* PRP buyout, small PRPs pay the regulator a fixed amount in exchange for relief from all future liabilities. Typically the buyout price that a given PRP will pay will more than cover its *expected* liability burden, because the PRP will be willing to pay a risk premium to avoid the uncertainty arising from continued liability. From a policy standpoint, a *de minimis* buyout serves several useful purposes. In particular, it is a source of immediate liquidity to fund short-term expenses. Moreover, by reducing the number of PRPs involved in the negotiation process, it may lower transactions costs and thus expedite the cleanup process.

After a buyout, on the other hand, all of the costs of uncertainty that the small PRPs originally had to bear will be transferred to the larger PRPs that remain. Thus the *expected* liability burden borne by each remaining PRP will remain the same, or will actually decline if the *de minimis* parties pay a risk premium. On the other hand, the variance of this burden will increase, because each remaining party is now responsible for a larger share of total cleanup costs. In the context of our model, this shift will increase the marginal benefit to an individual PRP of further investigation and thus exacerbate its incentive to induce delay. A striking fact is that this negative consequence of buyouts cannot be mitigated by increasing the risk premiums extracted from the *de minimis* parties. From the perspective of a remaining PRP, buyout revenue is a lump-sum transfer. That is, it reduces the PRP's exposure by an amount that is independent of the variable (private uncertainty level) over which the PRP has discretion.

The preceding discussion is summarized by the following proposition.

Proposition 6 (*de minimis* buyouts and delay). *Consider a liability profile,* $\mathbf{k}$*, and fix* $\underline{k} > 0$ *such that for some* j*,* $k_j \leq \underline{k}$*. For the model specified in section 6, a* de minimis *buyout of* all *PRPs whose liability shares do not exceed* $\underline{k}$ *will increase expected delay. Moreover, the increase in delay will be invariant with respect to the magnitude of the buyout premiums.*

7.2. **PRP steering committees.** Just as regulators have encouraged small PRPs to settle early via *de minimis* buyouts, they have also encouraged large PRPs to form *steering committees* that will negotiate with the regulators on behalf of their members. Such committees are socially useful to the extent that they encourage cooperative behavior between PRP's, and reduce contentious litigation over liability shares. Once again, our model highlights an attribute of steering committees that could have potentially major negative social consequences.

As we have noted, the extent of cleanup delay is determined by balancing the marginal expected cost of increasing the investigation period against the marginal benefit of the reduction in uncertainty resulting from a more thorough investigation. As we have noted above, an individual PRP's share of total expected cost increases in proportion to its liability share, while the uncertainty associated with its liability exposure increases with the *square* of its liability share. Thus, larger PRPs assign a greater relative importance to uncertainty than smaller ones. Consequently cleanup sites with many small PRPs are likely to be less subject to delay than ones with a few large PRPs. When a steering committee is formed, the interests of several PRP's are coalesced: effectively, several smaller PRP's are replaced by a single large one. From the PRPs' perspective, the formation of a steering committee thus internalizes an important externality. From a *social* perspective, however, this externality is a positive one, since it reduces delay. When a steering committee is formed, this externality is mitigated and social welfare is reduced. Of course, the social benefits of reduced litigation may offset this loss, but in this paper, these benefits are unmodeled.

The formation of a steering committee can be represented as a change in liability profile. Consider a liability profile $\boldsymbol{k}$, ordered as usual so that $i > j$ implies $k_i \geq k_j$. Suppose that a subset $J \in \{1, ..., n\}$ of PRPs form a committee, and let $\bar{j}$ be the PRP in J whose liability share is the largest. After the steering committee is formed, the liability profile can be represented by $\boldsymbol{k}'$ where $\bar{j}$'s share is now equal to the sum of the shares of all PRP's in J, and the shares of all PRP's in J except $\bar{j}$'s are reduced to zero. That is, let $\bar{j} = \min\{j \in J\}$ and define $k_i' = \sum_{j \in J} k_j$, if $i = \bar{j}$, k_i, if $i \in \{1, ..., n\} \setminus J$ and 0 otherwise. Observe also that $\boldsymbol{k}$' satisfies the condition of Proposition 5, since $k_i' > k_i$ for $i < \bar{j} + 1$, and $k_i' \leq k_i$ otherwise. Consequently, the following proposition is an immediate corollary of Proposition 5.

Proposition 7 (Steering Committees and delay). *For the model specified in section 6, if at least two PRP's form a steering committee, then expected delay will increase.*

8. Proofs

Proof of Proposition 1: Existence and the fact that s_i is $\underline{\eta}$ on an interval (which may be null) and then increasing, follows immediately from Athey's theorem. ■

Proof of Lemma 1.1: Part (a) follows from the fact that $f = \frac{1}{g}\frac{dW}{d\eta} = \frac{1}{g}\frac{dV}{dt}t' = U_t t'$ and $U_{t\kappa} < 0$. To prove (b), observe that $\frac{d^2W}{ds_i{}^2} = V_{tt}t' + V_t t''$ with $V_{tt} > 0$ and $t' > 0$. The sign of $f_{\Sigma s} = \frac{1}{d(k_i)}\frac{d^2W}{ds_i{}^2}$ will be determined by the first of these terms, provided that either $|V_t|$ or $|\frac{d^2t}{d\Sigma s^2}|$ is sufficiently small in absolute value. This proves part (b) of the Lemma. ■

Proof of Lemma 1.2: Pick $\underline{\theta}_i \in \Theta_{ij}$ and let $\hat{\gamma} = s_i(\underline{\theta}_i) - s_j(\underline{\theta}_i)$. We have

$$\begin{aligned}
0 &\le \frac{1}{d(k_i)}\frac{dE_{\boldsymbol{\theta}_{-i}}W((s_i(\underline{\theta}_i), \boldsymbol{s}_{-i}), (\underline{\theta}_i, \boldsymbol{\theta}_{-i}), k_i)}{d\eta} \\
&= \int_{[\theta^l,\theta^u]} E_{\boldsymbol{\theta}_{-ij}}\left[f(s_i(\underline{\theta}_i) + s_j(\vartheta) + \Sigma s_{-ij}, \underline{\theta}_i + \vartheta + \Sigma\theta_{-ij}, k_i)\right] g(\vartheta)d\vartheta \\
&= \int_{[\theta^l,\theta^u]} E_{\boldsymbol{\theta}_{-ij}}\left[f(s_j(\underline{\theta}_i) + \hat{\gamma} + s_j(\vartheta) + \Sigma s_{-ij}, \underline{\theta}_i + \vartheta + \Sigma\theta_{-ij}, k_i)\right] g(\vartheta)d\vartheta \\
&< \int_{[\theta^l,\theta^u]} E_{\boldsymbol{\theta}_{-ij}}\left[f(s_j(\underline{\theta}_i) + (\hat{\gamma} + s_j(\vartheta)) + \Sigma s_{-ij}, \underline{\theta}_i + \vartheta + \Sigma\theta_{-ij}, k_j)\right] g(\vartheta)d\vartheta \\
&\le \int_{[\theta^l,\theta^u]} E_{\boldsymbol{\theta}_{-ij}}\left[f(s_j(\underline{\theta}_i) + s_i(\vartheta) + \Sigma s_{-ij}, \underline{\theta}_i + \vartheta + \Sigma\theta_{-ij}, k_j)\right] g(\vartheta)d\vartheta \\
&= \frac{1}{d(k_j)}\frac{dE_{\boldsymbol{\theta}_{-j}}W((s_j(\underline{\theta}_i), \boldsymbol{s}_{-j}), (\underline{\theta}_i, \boldsymbol{\theta}_{-j}), k_j)}{d\eta}
\end{aligned}$$

The strict inequality follows from Lemma 1.1(a), i.e., f decreases w.r.t. k_i; the weak inequality holds because for all ϑ, $s_i(\vartheta) \ge s_j(\vartheta) + \hat{\gamma}$ and, by Lemma 1.1(b), f increases w.r.t. t, which in turn increases w.r.t. Σs. Since f is continuous, there exists a neighborhood U of $\underline{\theta}_i$ such that for all $\theta \in U$, $0 < \frac{dE_{\boldsymbol{\theta}_{-j}}W((s_j(\theta),\boldsymbol{s}_{-j}),(\theta,\boldsymbol{\theta}_{-j}),k_j)}{d\eta}$ and hence, from (1), $s_j(\theta) = \underline{\eta}$. ■

Proof of Proposition 2: To prove part (a), suppose $\tilde{\theta}_i \ge \tilde{\theta}_j$. If $\min_{\vartheta\in[\theta^l,\theta^u]}(s_i(\vartheta) - s_j(\vartheta)) = 0$, then since $s_j(\tilde{\theta}_j) = s_i(\tilde{\theta}_j) = \underline{\eta}$, $\tilde{\theta}_j \in \operatorname{argmin}_{\vartheta\in[\theta^l,\theta^u]}(s_i(\vartheta) - s_j(\vartheta))$. If $\min_{\vartheta\in[\theta^l,\theta^u]}(s_i(\vartheta) - s_j(\vartheta)) < 0$, then since $s_j(\cdot) = s_i(\cdot) = \underline{\eta}$ on $[\theta^l, \tilde{\theta}_j]$, $\operatorname{argmin}_{\vartheta\in[\theta^l,\theta^u]}(s_i(\vartheta) - s_j(\vartheta)) \subset (\tilde{\theta}_j, \theta^u]$. In either case, $\operatorname{argmin}_{\vartheta\in[\theta^l,\theta^u]}(s_i(\vartheta) - s_j(\vartheta)) \cap [\tilde{\theta}_j, \theta^u] \ne \emptyset$. But for all $\theta > \tilde{\theta}_j$, $s_j(\theta) > \underline{\eta}$, contradicting Lemma 1.2.

To prove part (b), assume that for some $\underline{\theta}_i \in (\tilde{\theta}_i, \theta^u]$, $s_i(\underline{\theta}_i) \le s_j(\underline{\theta}_i)$. From part (a), $\tilde{\theta}_j > \tilde{\theta}_i$. Since $s_i(\cdot) > s_j(\cdot)$ on $(\tilde{\theta}_i, \tilde{\theta}_j]$, it follows that $\underline{\theta}_i \in (\tilde{\theta}_j, \theta^u]$. Assume w.l.o.g. that $\underline{\theta}_i \in \operatorname{argmin}_{\vartheta\in(\tilde{\theta}_j,\theta^u]}(s_i(\vartheta) - s_j(\vartheta))$. Since $s_i(\cdot) \ge \underline{\eta} = s_j(\cdot)$ on $[\theta^l, \tilde{\theta}_j]$, $\underline{\theta}_i \in \operatorname{argmin}_{\vartheta\in(\theta^l,\theta^u]}(s_i(\vartheta) - s_j(\vartheta))$. But by definition of $\tilde{\theta}_j$, $s_j(\underline{\theta}_i) > \underline{\eta}$, contradicting Lemma 1.2. ■

Proof of Proposition 3: We first consider PRP i and argue that

$$\tilde{\theta}_i' < \tilde{\theta}_i \quad and \quad s_i'(\theta) > s_i(\theta) \quad \text{for} \quad \theta \geq \tilde{\theta}_i. \tag{9}$$

An exactly analogous argument establishes the corresponding properties for PRP j. Let $\underline{\gamma}_i = \min_{\vartheta \in [\tilde{\theta}_i, \theta^u]} \left(s_i'(\vartheta) - s_i(\vartheta)\right)$ and $\bar{\gamma}_j = \max_{\vartheta \in [\theta^l, \theta^u]} \left(s_j'(\vartheta) - s_j(\vartheta)\right)$. Suppose that $\underline{\gamma}_i \leq 0$ and pick $\underline{\theta}_i \in [\tilde{\theta}_i, \theta^u]$ such that $s_i'(\underline{\theta}_i) - s_i(\underline{\theta}_i) = \underline{\gamma}_i$. We first establish that $\bar{\gamma}_j + \underline{\gamma}_i$ must be positive. If not then,

$$\begin{aligned}
0 &\leq \int_{[\theta^l, \theta^u]} \left[f(s_i'(\underline{\theta}_i) + s_j'(\theta_j), \underline{\theta}_i + \theta_j, k_i')\right] g(\theta_j) d\theta_j \\
&< \int_{[\theta^l, \theta^u]} \left[f(s_i'(\underline{\theta}_i) + s_j'(\theta_j), \underline{\theta}_i + \theta_j, k_i)\right] g(\theta_j) d\theta_j \\
&\leq \int_{[\theta^l, \theta^u]} \left[f(s_i(\underline{\theta}_i) + \underline{\gamma}_i + s_j(\theta_j) + \bar{\gamma}_j, \underline{\theta}_i + \theta_j, k_i)\right] g(\theta_j) d\theta_j \\
&\leq \int_{[\theta^l, \theta^u]} \left[f(s_i(\underline{\theta}_i) + s_j(\theta_j), \underline{\theta}_i + \theta_j, k_i)\right] g(\theta_j) d\theta_j
\end{aligned} \tag{10}$$

The strict inequality follows from Lemma 1.1(a), i.e., f decreases w.r.t. k_i. The first weak inequality holds because $s_i'(\underline{\theta}_i) = s_i(\underline{\theta}_i) + \underline{\gamma}_i$, $s_j'(\cdot) \leq s_j(\cdot) + \bar{\gamma}_j$ and, by Lemma 1.1(b), f increases w.r.t. Σs. The second weak inequality holds because by assumption, $\bar{\gamma}_j + \underline{\gamma}_i \leq 0$ and f increases w.r.t. Σs. But inequality (10) is impossible since $\underline{\theta}_i \geq \tilde{\theta}_i$ implies

$$0 = \int_{[\theta^l, \theta^u]} \left[f(s_i(\underline{\theta}_i) + s_j(\theta_j), \underline{\theta}_i + \theta_j, k_i)\right] g(\theta_j) d\theta_j \tag{11}$$

This establishes that $\underline{\gamma}_i \leq 0$ implies $\bar{\gamma}_j + \underline{\gamma}_i > 0$. Now pick $\bar{\theta}_j$ such that $s_j'(\bar{\theta}_j) - s_j(\bar{\theta}_j) = \bar{\gamma}_j$. Since $\bar{\gamma}_j > -\underline{\gamma}_i \geq 0$, it follows that $\bar{\theta}_j > \tilde{\theta}_j'$. Hence

$$\begin{aligned}
0 &= \int_{[\theta^l, \theta^u]} \left[f(s_j'(\bar{\theta}_j) + s_i'(\theta_i), \bar{\theta}_j + \theta_i, k_j')\right] g(\theta_i) d\theta_i \\
&\geq \int_{[\theta^l, \theta^u]} \left[f(s_j'(\bar{\theta}_j) + s_i'(\theta_i), \bar{\theta}_j + \theta_i, k_j)\right] g(\theta_i) d\theta_i \\
&\geq \int_{[\theta^l, \theta^u]} \left[f(s_j(\bar{\theta}_j) + \bar{\gamma}_j + s_i(\theta_i) + \underline{\gamma}_i, \bar{\theta}_j + \theta_i, k_j)\right] g(\theta_i) d\theta_i \\
&> \int_{[\theta^l, \theta^u]} \left[f(s_j(\bar{\theta}_j) + s_i(\theta_i), \bar{\theta}_j + \theta_i, k_j)\right] g(\theta_i) d\theta_i
\end{aligned} \tag{12}$$

Once again, the first weak inequality hold because $f(\cdot)$ decreases with k and $k_j' \leq k_j$. The second weak inequality holds because $s_j'(\bar{\theta}_j) = s_j(\bar{\theta}_j) + \bar{\gamma}_j$, $s_i'(\cdot) \geq s_i(\cdot) + \underline{\gamma}_i$ and f increases w.r.t. Σs. The strict inequality holds because $\bar{\gamma}_j + \underline{\gamma}_i > 0$ and, once again, because f increases w.r.t. Σs. But inequality (12) is impossible because

$$0 \leq \int_{[\theta^l, \theta^u]} \left[f(s_j(\bar{\theta}_j) + s_i(\theta_i), \bar{\theta}_j + \theta_i, k_j) \right] g(\theta_i) d\theta_i$$

Thus we know $\underline{\gamma}_i > 0$, or $s_i'(\theta) > s_i(\theta) \quad \text{for} \quad \theta \geq \tilde{\theta}_i$, and from continuity of $s(\cdot)$, $\tilde{\theta}_i' < \tilde{\theta}_i$. ■

Proof of Proposition 4: The proof relies on two Lemmas. The first gathers together several results about ϵ-flatness.

Lemma 4.1. *(a): If $X \subset \mathbb{R}$ and for all $x, x' \in X$, $\left| \frac{f'(x)}{f(x')} \right| < \frac{\epsilon}{\max(x \in X) - \min(x \in X)}$, then f is ϵ-flat.*

Now fix an integer n, scalars $\alpha, \Omega \in \mathbb{R}_+$, and an integrable function $y : X \to \mathbb{R}$ such that $|y(\cdot)| < \Omega$.

(b): If f is ϵ-flat, for $\epsilon < \frac{\alpha\delta}{\Omega}$ and $\int_X f(x) y(x) g(x) dx \geq \alpha n \delta \underline{f}$ then $\int_X y(x) g(x) dx \geq \alpha(n-1)\delta$
(c): If f is ϵ-flat, for $\epsilon < \frac{\alpha\delta}{\Omega \underline{f}}$ and $\int_X y(x) g(x) dx \geq \frac{\alpha n \delta}{\underline{f}}$ then $\int_X f(x) y(x) g(x) dx \geq \alpha(n-1)\delta$.

The second lemma establishes a uniform boundedness property on a family of functions, paramaterized by $\alpha \in [0,1]$. Let $\Theta = [\theta^l, \theta^u]$ and let $\boldsymbol{\Theta} = \Theta^n$. For each α, let $\boldsymbol{f}(\cdot, \alpha)$ be a continuous function mapping $\boldsymbol{\Theta}$ to $\mathbb{R}^n_{++}$ and let $\boldsymbol{x}(\cdot, \ alpha) = (x_1(\cdot, \alpha, x_2(\cdot, \alpha)$ be an integrable function mapping $\boldsymbol{\Theta}$ to $\mathbb{R}^n$.

Lemma 4.2. *Assume that there exists $\omega \in \mathbb{R}_+$ and $\boldsymbol{\theta}^* \in \boldsymbol{\Theta}$ with $\theta_2^* > \theta^l$ such that for all α, the following conditions are satisfied: (a) for each r, $x_r(\cdot, \alpha)$ depends only on θ_r; there exists $\tilde{\boldsymbol{\theta}}(\alpha) \in \boldsymbol{\Theta}$, $\tilde{\boldsymbol{\theta}}(\alpha) \geq \boldsymbol{\theta}^*$, such that (b) $x_r(\cdot, \alpha) = 0$ on $[\theta^l, \tilde{\theta}_r(\alpha))$; (c) for each $\theta_r \in [\tilde{\theta}_r(\alpha), \theta^u]$, $E[f_r(\boldsymbol{\vartheta}, \alpha) \sum_{2=1}^n x_2(\boldsymbol{\vartheta}, \alpha) | \vartheta_r = \theta_r] \in (-\omega, \omega)$. For any $\epsilon > 0$ satisfying $\int_{\tilde{\theta}_2}^{\theta^l} g(\vartheta) d\vartheta < 1/(1+\epsilon)^2$, there exists $\Omega \in \mathbb{R}$ such that if f is ϵ-flat, then for $r = 1, 2$, $\sup(|x_r(\cdot, \alpha)|) < \Omega$, for all $\alpha \in [0,1]$.*

We can now proceed with the proof of Proposition 4. We need to prove that for α sufficiently close to unity,

$$0 \quad \leq \quad \int_{\boldsymbol{\Theta}} \frac{dt(\boldsymbol{\vartheta}, \alpha)}{d\alpha} \boldsymbol{g}(\boldsymbol{\vartheta}) d\boldsymbol{\vartheta} \quad = \quad \int_{\boldsymbol{\Theta}} \left(t'(\Sigma s(\boldsymbol{\vartheta}, \alpha)) \right) \sum_{r=1}^n \frac{d\left[s_r(\vartheta_r, \alpha) \right]}{d\alpha} \boldsymbol{g}(\boldsymbol{\vartheta}) d\boldsymbol{\vartheta} \tag{13}$$

We will proceed as follows.

1. For fixed α, partition $\boldsymbol{\Theta}$ into $\underline{\boldsymbol{\Theta}}(\alpha) = \left\{ \boldsymbol{\vartheta} \in \boldsymbol{\Theta} : \vartheta < \tilde{\theta}_1(\alpha) \right\}$ and $\overline{\boldsymbol{\Theta}}(\alpha) = \left\{ \boldsymbol{\vartheta} \in \boldsymbol{\Theta} : \vartheta \geq \tilde{\theta}_1(\alpha) \right\}$. In expression (17) below, we define a function $\psi_{1t} : \boldsymbol{\Theta} \times [0,1] \to \mathbb{R}_+$ below and prove that there exists $\delta > 0$ such that

$$\int_{\overline{\boldsymbol{\Theta}}(\alpha)} \psi_{1t}(\boldsymbol{\vartheta}, \alpha) \left(\sum_{r=1}^n \frac{d\left[s_r(\vartheta_r, \alpha) \right]}{d\alpha} \right) g(\boldsymbol{\vartheta}) d\boldsymbol{\vartheta} \quad > \quad \frac{3\delta \underline{\psi}_{1t}}{t'(n\underline{\eta})} \tag{14}$$

where $\underline{\psi}_{1t} = \min \left\{ \psi_{1t}(\boldsymbol{\vartheta}, \alpha) : \boldsymbol{\vartheta} \in \boldsymbol{\Theta} \right\}$.

2. Invoke Lemma 4.2 to establish that there exists $\Omega \in \mathbb{R}_+$ such that for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and all $\alpha \in [0,1]$, $\sup \left| \frac{d[s_r(\boldsymbol{\theta},\alpha)]}{d\alpha} \right| < \Omega$.
3. If $\psi_{1t}(\cdot,\alpha)$ is γ-flat, for $\gamma < \frac{\delta\psi_{1t}}{\Omega t'}$ it follows from Step 1 and Part (b) of Lemma 4.1 that

$$\int_{\underline{\boldsymbol{\Theta}}(\alpha)} \left(\sum_{r=1}^{n} \frac{d\left[s_r(\vartheta_r,\alpha)\right]}{d\alpha} \right) g(\boldsymbol{\vartheta}) d\boldsymbol{\vartheta} \quad > \quad \frac{2\delta}{\underline{t}'}.$$

4. If V_{tt} is ϵ_{V1}-flat, for $\epsilon_{V1} < (1+\gamma)^{1/3} - 1$ and $\frac{t''(n\bar{\eta})}{t'(n\underline{\eta})} < \epsilon_t$, for $\epsilon_t < \min\left\{ \frac{(1+\gamma)\min(V_t)}{\max(V_t)}, \frac{(1+\gamma)^{1/3}-1}{2(\bar{\eta}-\underline{\eta})} \right\}$, then $\psi_{1t}(\cdot,\alpha)$ is γ-flat.
5. If t' is ϵ_{V2}-flat, for $\epsilon_{V2} < \frac{\delta}{\Omega t'(n\underline{\eta})}$ it follows from Step 3 and Part (c) of Lemma 4.1 that

$$\int_{\underline{\boldsymbol{\Theta}}(\alpha)} \left(t'\left(\Sigma s(\boldsymbol{\vartheta},\alpha)\right) \right) \sum_{r=1}^{n} \frac{d\left[s_r(\vartheta_r,\alpha)\right]}{d\alpha} g(\boldsymbol{\vartheta}) d\boldsymbol{\vartheta} > \delta.$$

6. Since $\tilde{\theta}_1(\alpha) \searrow \theta^l$ as $\alpha \to 1$, and since $t'(\cdot)$ is increasing in Σs, we can pick $\bar{\alpha} < 1$ sufficiently large that $\int_{\underline{\boldsymbol{\Theta}}(\alpha)(\alpha)} g(\boldsymbol{\vartheta}) d\boldsymbol{\vartheta} < \frac{\delta}{t'(n\bar{\eta})}$. thus ensuring that for $\alpha > \bar{\alpha}$, $\int_{\underline{\boldsymbol{\Theta}}(\alpha)} \left(t'\left(\Sigma s(\boldsymbol{\vartheta},\alpha)\right) \right) \sum_{r=1}^{n} \frac{d[s_r(\vartheta_r,\alpha)]}{d\alpha} g(\boldsymbol{\vartheta}) d\boldsymbol{\vartheta} > -\delta$. This will complete the proof.

Of these steps, 3, 5 and 6 require no further work.

Step 1: Fix $\alpha \in [0,1]$, $r \in \{1,2\}$ and $\theta'_r \in (\tilde{\theta}_r(\alpha), \theta^u]$. It follows from Proposition 1 and (1) that

$$\begin{aligned} 0 &= \int_{\Theta} \left[V_t\Big(t\big(s_r(\theta'_r,\alpha) + s_{\neg r}(\vartheta_{\neg r},\alpha)\big), (\theta'_r, \vartheta_{\neg r}), \kappa_r(\alpha)\Big) \frac{d\left[t\big(s_r(\theta'_r,\alpha) + s_{\neg r}(\vartheta_{\neg r},\alpha)\big)\right]}{ds_r} \right] g(\vartheta_{\neg r}) d\vartheta_{\neg r} \\ &= \int_{\Theta} \psi_r((\theta'_r, \vartheta_{\neg r}),\alpha) g(\vartheta_{\neg r}) d\vartheta_{\neg r} \qquad (15) \end{aligned}$$

where

$$\begin{aligned} \psi_r((\theta'_r, \vartheta_{\neg r}),\alpha) &= V_t\Big(t\big(\Sigma s((\theta'_r, \vartheta_{\neg r}),\alpha)\big), (\theta'_r, \vartheta_{\neg r}), \kappa_r(\alpha)\Big) t'\big(\Sigma s((\theta'_r, \vartheta_{\neg r}),\alpha)\big) \frac{d\Sigma s((\theta'_r, \vartheta_{\neg r}),\alpha)}{ds_r} \\ &= V_t\Big(t\big(\Sigma s((\theta'_r, \vartheta_{\neg r}),\alpha)\big), (\theta'_r, \vartheta_{\neg r}), \kappa_r(\alpha)\Big) t'\big(\Sigma s((\theta'_r, \vartheta_{\neg r}),\alpha)\big) \end{aligned}$$

Since (15) holds for all α, the total derivative w.r.t. α of the right hand side of (15) must be identically zero. To expand this derivative, we must first compute the total derivative of $\psi_r((\theta'_r,\cdot),\alpha)$ with respect to α. For each $\vartheta_{\neg r}$, we have:

$$\frac{d}{d\alpha}\left[\psi_r((\theta'_r, \vartheta_{\neg r}),\alpha)\right] = \psi_{rt}((\theta'_r, \vartheta_{\neg r}),\alpha) \sum_{2=1}^{n} \frac{d\left[s_2((\theta'_r, \vartheta_{\neg r}),\alpha)\right]}{d\alpha} + \psi_{rk}((\theta'_r, \vartheta_{\neg r}),\alpha) \frac{d\kappa_r(\alpha)}{d\alpha} \qquad (16)$$

where

$$\psi_{rt}((\theta_r', \vartheta_{\neg r}), \alpha) = \Big(V_{tt}\Big(t\big(\Sigma s((\theta_r', \vartheta_{\neg r}), \alpha)\big), (\theta_r', \vartheta_{\neg r}), \kappa_r(\alpha)\Big) \left(t'\big(\Sigma s((\theta_r', \vartheta_{\neg r}), \alpha)\big)\right)^2 + V_t\Big(t\big(\Sigma s((\theta_r', \vartheta_{\neg r}), \alpha)\big), (\theta_r', \vartheta_{\neg r}), \kappa_r(\alpha)\Big) t''\big(\Sigma s((\theta_r', \vartheta_{\neg r}), \alpha)\big) \Big) \quad (17)$$

and

$$\psi_{rk}((\theta_r', \vartheta_{\neg r}), \alpha) = V_{tk}\Big(t\big(\Sigma s((\theta_r', \vartheta_{\neg r}), \alpha)\big), (\theta_r', \vartheta_{\neg r}), \kappa_r(\alpha)\Big) \frac{d\left[t\big(\Sigma s((\theta_r', \vartheta_{\neg r}), \alpha)\big)\right]}{ds_r}$$

Our assumptions on V and t ($V_t, V_{tt}, t', t'' > 0$, $V_{tk} < 0$) guarantee that $\psi_{rt}(\cdot) > 0$ and $\psi_{rk}(\cdot) < 0$. Now, taking the total derivative of (15) w.r.t. α we obtain:

$$\begin{aligned} 0 &= \frac{d}{d\alpha}\left[\int_\Theta \psi_r((\theta_r', \vartheta_{\neg r}), \alpha) g(\vartheta_{\neg r}) d\vartheta_{\neg r}\right] \\ &= \int_\Theta \psi_{rt}((\theta_r', \vartheta_{\neg r}), \alpha) \sum_{2=1}^{n} \left[\frac{d\left[s_2((\theta_r', \vartheta_{\neg r})\alpha)\right]}{d\alpha}\right] g(\vartheta_{\neg r}) d\vartheta_{\neg r} \\ &\quad + \frac{d\kappa_r(\alpha)}{d\alpha} \int_\Theta \psi_{rk}((\theta_r', \vartheta_{\neg r}), \alpha) g(\vartheta_{\neg r}) d\vartheta_{\neg r} \end{aligned} \quad (18)$$

For future reference, note that for all r, α and $\theta_r' \in (\tilde{\theta}_r(\alpha), \theta^u]$,

$$\begin{aligned} \int_\Theta \psi_{rt}((\theta_r', \vartheta_{\neg r}), \alpha) \sum_{2=1}^{n} \left[\frac{d\left[s_2((\theta_r', \vartheta_{\neg r})\alpha)\right]}{d\alpha}\right] g(\vartheta_{\neg r}) d\vartheta_{\neg r} &\in -\left(k_r' - k_r\right) \overline{\psi}_{rk} \\ &\subset [-\overline{\psi}_{rk}, \overline{\psi}_{rk}] \end{aligned} \quad (19)$$

where $\overline{\psi}_{rk} = \max\{\psi_{rk}(\boldsymbol{\vartheta}, \alpha) : \boldsymbol{\vartheta} \in \boldsymbol{\Theta}, \alpha \in [0,1]\}$. Now set $r = 1$ and integrate (18) over the interval $(\tilde{\theta}_r(\alpha), \theta^u]$ on which this equality holds to obtain:

$$\begin{aligned} \int_{\overline{\boldsymbol{\Theta}}(\alpha)} \psi_{1t}(\boldsymbol{\vartheta}, \alpha) \left(\sum_{r=1}^{n} \frac{d\left[s_r(\vartheta_r, \alpha)\right]}{d\alpha}\right) g(\boldsymbol{\vartheta}) d\boldsymbol{\vartheta} - &= \frac{d\kappa_1(\alpha)}{d\alpha} \int_{\overline{\boldsymbol{\Theta}}(\alpha)} \psi_{1k}(\boldsymbol{\vartheta}, \alpha) g(\boldsymbol{\vartheta}) d\boldsymbol{\vartheta} \\ &\geq -\left(k_1' - k_1\right) \underline{\psi}_{1k} > 0 \end{aligned}$$

where $\underline{\psi}_{1k} = \min\{\psi_{1k}(\boldsymbol{\vartheta}, \alpha) : \boldsymbol{\vartheta} \in \boldsymbol{\Theta}, \alpha \in [0,1]\}$. Hence $\delta > 0$ can be chosen sufficiently small to ensure that inequality (14) is satisfied. This completes the proof of Step 1.

<u>Step 2</u>: Consider the family of functions $\boldsymbol{x}(\cdot, \alpha)$, where $x_r(\cdot, \alpha) = \frac{d[s_r(\cdot,\alpha)]}{d\alpha}$, and $\boldsymbol{f}(\cdot, \alpha)$ where $f_r(\cdot, \alpha) = \psi_{rt}(\cdot, \alpha)$. Let $\omega = \overline{\psi}_{rk}$ and $\boldsymbol{\theta}^* = (\theta^l, \tilde{\theta}_2(0))$. Note that for each α, $(\tilde{\theta}_1(\alpha), \tilde{\theta}_2(\alpha)) \geq (\theta^l, \tilde{\theta}_2(0)) = \boldsymbol{\theta}^*$. Observe that (a) for each r, $x_r(\cdot, \alpha)$ depends only on θ_r; for $\tilde{\boldsymbol{\theta}}(\alpha) = (\tilde{\theta}_1(\alpha), \tilde{\theta}_2(\alpha))$ $\tilde{\boldsymbol{\theta}}(\alpha) \geq \boldsymbol{\theta}^*$ and the following conditions are satisfied for each r, (b) (cf. Proposition 1 and (1)) $x_r(\cdot, \cdot) = 0$ on $[\theta^l, \tilde{\theta}_r(\alpha))$; (c) (cf. (19)). for each $\theta_r \in [\tilde{\theta}_r(\alpha), \theta^u]$, $E[f_r(\boldsymbol{\vartheta}) \sum_{2=1}^{n} x_2(\boldsymbol{\vartheta}, \alpha) | \vartheta_r = \theta_r] \in (-\omega, \omega)$. Hence the conclusion of Lemma 4.2 applies.

Step 4: From (17)

$$\psi_{1t}(\cdot,\cdot) \quad = \quad V_{tt}(\cdot,\cdot,\cdot)(t'(\cdot))^2 \quad + \quad V_t(\cdot,\cdot,\cdot)t''(\cdot)$$

From Part (a) of Lemma 4.1, t'' is $((1+\gamma)^{1/3}-1)$-flat. By assumption, V_{tt} is also. Hence $\frac{\max[V_{tt}(\cdot,\cdot,\cdot)(t'(\cdot))^2]}{\min[V_{tt}(\cdot,\cdot,\cdot)(t'(\cdot))^2]} \leq (1+\gamma)$. Moreover, since $t''(\cdot) < (1+\gamma)\frac{\min(V_t)}{\max(V_t)}$, $\frac{\max[V_t(\cdot,\cdot,\cdot)t''(\cdot)]}{\min[V_t(\cdot,\cdot,\cdot)t''(\cdot)]} \leq (1+\gamma)$ also. Hence $\frac{\max[\psi_{1t}(\cdot,\cdot)]}{\min[\psi_{1t}(\cdot,\cdot)]} \leq (1+\gamma)$. ■

We now prove the two lemmas.
Proof of Lemma 4.1:

(a): Let $\underline{f} = \min(f(x) : x \in X)$ and $\overline{f} = \max(f(x) : x \in X)$. Note that $|f'(x)| < \frac{\epsilon \underline{f}}{\max(x\in X)-\min(x\in X)}$. Pick $\overline{x}$ and $\underline{x}$ in X such that $f(\overline{x}) = \overline{f}$ and $f(\underline{x}) = \underline{f}$. Assume (w.l.o.g.) that $\overline{x} > \underline{x}$.

$$\begin{aligned}\overline{f} - \underline{f} \quad &= \quad \int_{\underline{x}}^{\overline{x}} f'(x)dx \quad \leq \quad \sup(f')\int_{\underline{x}}^{\overline{x}} dx \quad \leq \quad \sup(f')\int_{\max(x\in X)}^{\min(x\in X)} dx \\ &\leq \quad \frac{\epsilon \underline{f}}{[\max(x\in X)-\min(x\in X)]}[\max(x\in X)-\min(x\in X)] \quad = \quad \epsilon \underline{f}\end{aligned}$$

(b):

$$\begin{aligned}\alpha n \delta \underline{f} \quad &\leq \quad \int_X f(x)y(x)g(x)dx \\ &\leq \quad \underline{f}\int_X (y(x))^- g(x)dx \quad + \quad \underline{f}\left(1+\frac{\alpha\delta}{\Omega}\right)\int_X (y(x))^+ g(x)dx \\ &= \quad \underline{f}\left[\int_X y(x)g(x)dx \quad + \quad \frac{\alpha\delta}{\Omega}\Omega\right]\end{aligned}$$

Hence

$$\alpha(n-1)\delta \quad \leq \quad \int_X y(x)g(x)dx$$

(c):

$$\begin{aligned}\int_X f(x)y(x)g(x)dx \quad &\geq \quad \underline{f}\left[\int_X (y(x))^+ g(x)dx \quad + \quad \left(1+\frac{\alpha\delta}{\underline{f}\Omega}\right)\int_X (y(x))^- g(x)dx\right] \\ &\geq \quad \underline{f}\left[\frac{\alpha n\delta}{\underline{f}} \quad - \quad \frac{\alpha\delta}{\Omega \underline{f}}\Omega\right] \quad = \quad \alpha(n-1)\delta\end{aligned}$$

■

Proof of Lemma 4.2:
Preliminary Step: If $\epsilon < 1$, there exists Ω such that for all $\alpha \in [0,1]$ and $r = \{1,2\}$, if $|x_r(\theta)| > \Omega$ for some $\theta \in [\tilde{\theta}_r(\alpha), \theta^u]$, then $x_r(\theta)x_r(\theta') > 0$. for all $\theta' \in [\tilde{\theta}_r(\alpha), \theta^u]$.

Proof of the Preliminary Step: Suppose that the preliminary step is false, i.e., that there exists a sequence $\{\alpha^m\} \in [0,1]$ and two sequences $\{\overline{\theta}_1^m\}$ and $\{\underline{\theta}_1^m\}$ such that $x_1^m(\overline{\theta}_1^m) \to \infty$ while for all m, $x_1^m(\underline{\theta}_1^m) \leq 0$, where $x_1^m(\cdot) = x_1(\cdot, \alpha^m)$. It follows that $\Delta x_1^m \to \infty$, where $\Delta x_1^m = \sup\Big\{x_1^m(\vartheta) : \vartheta \in [\tilde{\theta}_1(\alpha), \theta^u)\Big\} - \inf\Big\{x_1^m(\vartheta) : \vartheta \in [\tilde{\theta}_1(\alpha), \theta^u]\Big\}$. Letting $f_1^m = f_1(\cdot, \alpha^m)$, we have[13]

$$
\begin{aligned}
\omega \quad &\geq \quad x_1^m(\overline{\theta}_1^m)\int_\Theta f_1^m(\overline{\theta}_1^m, \vartheta_2)g_2(\vartheta_2)d\vartheta_2 + \int_\Theta f_1^m(\overline{\theta}_1^m, \vartheta_2)x_2^m(\vartheta_2)g_2(\vartheta_2)d\vartheta_2 \\
&\geq \quad \underline{f}_1 x_1^m(\underline{\theta}_1^m) \quad + \quad \underline{f}_1 E\Big[\Big(x_2^m\Big)^+\Big] \quad + \quad \overline{f}_1 E\Big[\Big(x_2^m\Big)^-\Big] \\
&\geq \quad \underline{f}_1\Big\{x_1^m(\overline{\theta}_1^m) \quad + \quad E\Big[x_2^m\Big] \quad + \quad \epsilon E\Big[\Big(x_2^m\Big)^-\Big]\Big\}
\end{aligned}
\tag{20}
$$

Similarly,

$$
\begin{aligned}
-\omega \quad &\leq \quad x_1^m(\underline{\theta}_1^m)\int_\Theta f_1^m(\underline{\theta}_1^m, \vartheta_2)g_2(\vartheta_2)d\vartheta_2 + \int_\Theta f_1^m(\underline{\theta}_1^m, \vartheta_2)x_2^m(\vartheta_2)g_2(\vartheta_2)d\vartheta_2 \\
&\leq \quad \underline{f}_1 x_1^m(\underline{\theta}_1^m) \quad + \quad \overline{f}_1 E\Big[\Big(x_2^m\Big)^+\Big] \quad + \quad \underline{f}_1 E\Big[\Big(x_2^m\Big)^-\Big] \\
&\leq \quad \underline{f}_1\Big\{x_1^m(\underline{\theta}_1^m) \quad + \quad E\Big[x_2^m\Big] \quad + \quad \epsilon E\Big[\Big(x_2^m\Big)^+\Big]\Big\}
\end{aligned}
\tag{21}
$$

Let $\Delta x^m = \sum_{r=1}^n \Delta x_r^m$. Since $\Delta x_1^m \to \infty$ and for all m, $\Delta x^m \geq \Delta x_1^m$, it follows that $\Delta x^m \to \infty$. Dividing both sides of (20) and (21) by $\underline{f}_1 \Delta x^m$ we obtain:

$$
\frac{x_1^m(\overline{\theta}_1^m)}{\Delta x^m} \quad + \quad \frac{E\Big[x_2^m\Big]}{\Delta x^m} \quad + \quad \frac{\epsilon E\Big[\Big(x_2^m\Big)^-\Big]}{\Delta x^m} \quad \to \quad 0 \tag{22}
$$

$$
\frac{x_1^m(\underline{\theta}_1^m)}{\Delta x^m} \quad + \quad \frac{E\Big[x_2^m\Big]}{\Delta x^m} \quad + \quad \frac{\epsilon E\Big[\Big(x_2^m\Big)^+\Big]}{\Delta x^m} \quad \to \quad 0 \tag{23}
$$

Note that (22), (23) and the fact that $\lim_{m\to\infty} x_1^m(\underline{\theta}_1^m) \leq 0$ together imply that

$$
\begin{aligned}
\inf\Big\{x_2^m(\vartheta) : \boldsymbol{\vartheta} \in [\tilde{\theta}_2^m, \theta^u]\Big\} \quad &\leq \quad E\Big[\Big(x_2^m\Big)^-\Big] \quad \to \quad -\infty \qquad \text{while} \\
\sup\Big\{x_2^m(\vartheta) : \boldsymbol{\vartheta} \in [\tilde{\theta}_2^m, \theta^u]\Big\} \quad &\geq \quad E\Big[\Big(x_2^m\Big)^+\Big] \quad \to \quad \infty
\end{aligned}
\tag{24}
$$

[13] For any variable or function x taking values in $\mathbb{R}$, let $x^+ = \max(x, 0)$ and $x^- = \min(x, 0)$.

Subtracting (23) from (22), we obtain

$$\frac{\Delta x_1^m - \epsilon\left(E\left[\left(x_2^m\right)^+\right] - E\left[\left(x_2^m\right)^-\right]\right)}{\Delta x^m} \quad \rightarrow \quad 0 \tag{25}$$

Hence from (24)

$$\lim_{m\to\infty} \frac{\Delta x_1^m \quad - \quad \epsilon\Delta x_2^m}{\Delta x^m} \quad \leq \quad 0 \tag{26}$$

It follows that $\Delta x_2^m \rightarrow \infty$. Repeating the argument that gave rise to (25) for 1, we obtain (27) below:

$$\frac{\Delta x_2^m \quad - \quad \epsilon\left(E\left[\left(x_1^m\right)^+\right] - E\left[\left(x_1^m\right)^-\right]\right)}{\Delta x^m} \quad \rightarrow \quad 0 \tag{27}$$

Hence, reasoning as above:

$$\lim_{m\to\infty} \frac{\Delta x_2^m \quad - \quad \epsilon\Delta x_1^m}{\Delta x^m} \quad \leq \quad 0 \tag{28}$$

But if $\epsilon < 1$, (26) and (28) cannot hold simultaneously, proving that the Preliminary Step is true.

Now suppose (w.l.o.g.) that the conclusion of the lemma is false for 1,, i.e., that there exists a sequence $\{\alpha^m\} \in [0,1]$ and a sequence $\{\theta_1^m\}$ such that for all m, $|x_1^m(\theta_1^m)| > m$, where $x_1^m(\cdot) = x_1(\cdot, \alpha^m)$. Assume w.l.o.g. that $x_1^m(\theta_1^m)$ is positive for all m. Moreover, we can clearly choose θ_1^m so that for each m, $x_1^m(\theta_1^m) \geq E[x_2^m(\vartheta_1)|\vartheta_1 \geq \tilde{\theta}_1^m]$, where $\tilde{\theta}_1^m = \tilde{\theta}_1(\alpha^m)$. From the preliminary step, it follows that if m is sufficiently large, $x_1^m(\cdot)$ is positive on $[\tilde{\theta}_1^m, \theta^u]$. From condition (c) in the statement of the Lemma,

$$\frac{\omega}{x_1^m(\theta_1^m)E(f_1^m(\theta_1^m, \vartheta_2))} \quad \geq \quad 1 \quad + \quad \frac{E[f_1^m(\boldsymbol{\vartheta})x_2^m(\boldsymbol{\vartheta})|\vartheta_1 = \theta_1^m]}{x_1^m(\theta_1^m)E(f_1^m(\theta_1^m, \vartheta_2))} \quad \geq \quad \frac{-\omega}{x_1^m(\theta_1^m)E(f_1^m(\theta_1^m, \vartheta_2))}$$

Since $x_2^m(\cdot)$ depends only on ϑ_2 and is identically zero on $[\theta^l, \tilde{\theta}_2^m)$, the above expression can be rewritten as

$$\frac{\omega}{x_1^m(\theta_1^m)E(f_1^m(\theta_1^m, \vartheta_2))} \quad \geq \quad 1 \quad + \quad \frac{p_2(\tilde{\boldsymbol{\theta}}^m)E[f_1^m(\boldsymbol{\vartheta})x_2^m(\boldsymbol{\vartheta})|\vartheta_1 \geq \tilde{\theta}_2^m]}{x_1^m(\theta_1^m)E(f_1^m(\theta_1^m, \vartheta_2))} \quad \geq \quad \frac{-\omega}{x_1^m(\theta_1^m)E(f_1^m(\theta_1^m, \vartheta_2))}$$

Clearly both of the outer bounds converge to zero. Hence

$$\lim_{m\to\infty} \frac{E[f_1^m(\boldsymbol{\vartheta})x_2^m(\boldsymbol{\vartheta})|\vartheta_1 \geq \tilde{\theta}_2^m]}{x_1^m(\theta_1^m)E(f_1^m(\theta_1^m, \vartheta_2))} \quad = \quad -\frac{1}{p_2(\tilde{\boldsymbol{\theta}}^m)} \tag{29}$$

Necessarily, there exists a sequence $\{\theta_2^m\}$ in $[\tilde{\theta}_2^m, \theta^u]$, such that $x_2^m(\theta_2^m) \to -\infty$. From the preliminary step, it follows if m is sufficiently large, $x_2^m(\cdot)$ is negative on $[\tilde{\theta}_2^m, \theta^u]$. Hence, since $x_1^m(\theta_1^m) \geq E[x_2^m(\vartheta_1)|\vartheta_1 \geq \tilde{\theta}_1^m]$:

$$\lim_{m\to\infty} \frac{E[f_1^m(\boldsymbol{\vartheta})x_2^m(\boldsymbol{\vartheta})|\vartheta_1 \geq \tilde{\theta}_2^m]}{x_1^m(\theta_1^m)E(f_1^m(\theta_1^m, \vartheta_2))} \geq \lim_{m\to\infty} \frac{\overline{f}_1 E[x_2^m(\boldsymbol{\vartheta})|\vartheta_1 \geq \tilde{\theta}_2^m]}{x_1^m(\theta_1^m)E(f_1^m(\theta_1^m, \vartheta_2))}$$
$$\geq \lim_{m\to\infty} \frac{\overline{f}_1 E[x_2^m(\boldsymbol{\vartheta})|\vartheta_1 \geq \tilde{\theta}_2^m]}{\underline{f}_1 E[x_1^m(\boldsymbol{\vartheta})|\vartheta_1 \geq \tilde{\theta}_1^m]} \quad (30)$$

Since $\overline{f}_1 = (1+\epsilon)\underline{f}_1$, (29) and (30) imply that

$$\lim_{m\to\infty} \frac{E[x_2^m(\vartheta_2)|\vartheta_2 \geq \tilde{\theta}_2^m]}{E[x_1^m(\vartheta_1)|\vartheta_1 \geq \tilde{\theta}_1^m]} \leq \frac{-1}{p_2(\tilde{\boldsymbol{\theta}}^m)(1+\epsilon)} \quad (31)$$

Note, however, that the chain of reasoning we have just applied leads to the following, exact counterpart of (31):

$$\lim_{m\to\infty} \frac{E[x_1^m(\vartheta_2)|\vartheta_1 \geq \tilde{\theta}_1^m]}{E[x_2^m(\vartheta_1)|\vartheta_2 \geq \tilde{\theta}_2^m]} \leq \frac{-1}{p_1(\tilde{\boldsymbol{\theta}}^m)(1+\epsilon)} \quad (32)$$

Since $p_1(\tilde{\boldsymbol{\theta}}^m) \leq 1$, it follows that $\lim_{m\to\infty} \frac{E[x_2^m(\vartheta_2)|\vartheta_1 \geq \tilde{\theta}_1^m]}{E[x_1^m(\vartheta_1)|\vartheta_2 \geq \tilde{\theta}_2^m]} \geq -(1+\epsilon)$. But since $p_2(\tilde{\boldsymbol{\theta}}^m) \leq \boldsymbol{p}(\boldsymbol{\theta}^*) < 1/(1+\epsilon)^2$, this inequality and (31) cannot hold simultaneously. ■

Proof of Proposition 5: Define the function $\kappa : [0,1] \to \mathbb{R}^n$ by, for $\alpha \in [0,1]$, $\kappa(\alpha) = \alpha \boldsymbol{k}' + (1-\alpha)\boldsymbol{k}$. Observe that for all $i > \overline{j}$ and all $\alpha \in [0,1]$, $\frac{d\kappa_i(\cdot)}{d\alpha} > 0$, while $\frac{d\kappa_i(\cdot)}{d\alpha} < 0$ otherwise. For each $\alpha \in [0,1]$, let $\boldsymbol{s}(\cdot;\alpha)$ be the PSNE corresponding to α. We will prove that $\frac{d}{d\alpha}\left[E_{\Theta^n}\ (t(\Sigma s(\boldsymbol{\theta},\alpha)))\right] > 0$. Integrating this derivative w.r.t. α from zero to one, it will then follow $E_{\Theta^n}\ t(\Sigma s'(\boldsymbol{\theta})) > E_{\Theta^n}\ t(\Sigma s(\boldsymbol{\theta}))$.

For each i define $\tilde{\theta}_i(\alpha)$ implicitly by the equation $\kappa_i(\alpha)\frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}}(\tilde{\theta}_i(\alpha)+E_\Theta \Sigma\theta_{-i}) = \underline{\eta}+E_\Theta \Sigma s_{-i}(\boldsymbol{\theta}_{-i},\alpha)$. Observe from (8) that for i and $\theta_i > \tilde{\theta}_i(\alpha)$,

$$0 = \int_\Theta \left[s_i(\theta_i,\alpha) + \Sigma s_{-i}(\boldsymbol{\theta}_{-i},\alpha) - \kappa_i(\alpha)\frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}}(\theta_i + \Sigma\theta_{-i})\right] \boldsymbol{g}_{-i}(\boldsymbol{\theta}_{-i})d\boldsymbol{\theta}_{-i}$$

It follows that for all i and all α:

$$\frac{ds_i(\theta_i,\alpha)}{d\alpha} = \begin{cases} 0 & \text{if } \theta_i < \tilde{\theta}_i(\alpha) \\ \frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}}(\theta_i + E_\Theta\Sigma\theta_{-i})\frac{d\kappa_i(\alpha)}{d\alpha} - E_\Theta\frac{d\Sigma s_{-i}(\boldsymbol{\theta}_{-i},\alpha)}{d\alpha} & \text{if } \theta_i > \tilde{\theta}_i(\alpha) \end{cases} \quad (33)$$

Applying Leibniz's rule to (33) and noting that $s_i(\tilde{\theta}_i) = \underline{\eta} = 0$, we obtain

$$\frac{d}{d\alpha}\left[E_\Theta s_i(\cdot,\alpha)\right] = -s_i(\tilde{\theta}_i,\alpha)\frac{d\tilde{\theta}_i(\alpha)}{d\alpha} + \int_{\tilde{\theta}_i(\alpha)}^{\theta^u} \frac{ds_i(\theta_i,\alpha)}{d\alpha}g(\theta_i)d\theta_i = \int_{\tilde{\theta}_i(\alpha)}^{\theta^u} \frac{ds_i(\theta_i,\alpha)}{d\alpha}g(\theta_i)d\theta_i.$$

For $\vartheta \in \Theta$, let $\Psi(\vartheta) = \frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}} \int_{\vartheta}^{\theta^u} (\vartheta + E_\Theta \Sigma\theta_{-i})\, g(\vartheta) d\vartheta$. Also let $p_i(\alpha) = \int_{\tilde{\theta}_i(\alpha)}^{\theta^u} g(\theta_i) d\theta_i$ denote the probability that $s_i(\cdot, \alpha) > \underline{\eta}$. Integrating (33) over Θ, we obtain

$$\begin{aligned}
\frac{d}{d\alpha}\left[E_\Theta s_i(\theta_i, \alpha)\right] &= \Psi(\tilde{\theta}_i(\alpha)) \frac{d\kappa_i(\alpha)}{d\alpha} - p_i(\alpha) \frac{d}{d\alpha}\left[E_{\Theta^n} \Sigma s(\boldsymbol{\theta}, \alpha)\right] + p_i(\alpha) \frac{d}{d\alpha}\left[E_\Theta s_i(\theta_i, \alpha)\right] \\
&= \frac{\Psi(\tilde{\theta}_i(\alpha))}{1 - p_i(\alpha)} \frac{d\kappa_i(\alpha)}{d\alpha} - \frac{p_i(\alpha)}{1 - p_i(\alpha)} \frac{d}{d\alpha}\left[E_{\Theta^n} \Sigma s(\boldsymbol{\theta}, \alpha)\right]. \qquad (34)
\end{aligned}$$

Now by assumption, there exists $\bar{j} \in \{2, ..., n\}$ such that $\frac{d\kappa_i(\alpha)}{d\alpha} \geq 0$ iff $i \leq \bar{j}$ and $\frac{d\kappa_i(\alpha)}{d\alpha} \leq 0$ otherwise. Moreover, since $i > \bar{j}$ implies $k_i(\alpha) \geq k_{\bar{j}}(\alpha)$, it follows from Proposition 2 that $\tilde{\theta}_i(\alpha) \leq \tilde{\theta}_{\bar{j}}(\alpha)$ and $p_i(\alpha) \geq p_{\bar{j}}(\alpha)$ if $i < \bar{j}$, while $\tilde{\theta}_i(\alpha) \geq \tilde{\theta}_{\bar{j}}(\alpha)$ and $p_i(\alpha) \leq p_{\bar{j}}(\alpha)$ otherwise, with equality holding only if $\kappa_i(\alpha) = \kappa_{\bar{j}}(\alpha)$. It follows that $\frac{\Psi(\tilde{\theta}_i(\alpha))}{1-p_i(\alpha)} \geq \frac{\Psi(\tilde{\theta}_{\bar{j}}(\alpha))}{1-p_{\bar{j}}(\alpha)}$ if $i \leq \bar{j}$, while $\frac{\Psi(\tilde{\theta}_i(\alpha))}{1-p_i(\alpha)} \leq \frac{\Psi(\tilde{\theta}_{\bar{j}}(\alpha))}{1-p_{\bar{j}}(\alpha)}$, otherwise. Moreover, by assumption not all of the k_i's are equal, so that for at least one i, the above inequality holds strictly. These relationships imply that for all i,

$$E_\Theta d s_i(\theta_i) \geq \frac{\Psi(\tilde{\theta}_{\bar{j}}(\alpha))}{1 - p_{\bar{j}}(\alpha)} \frac{d\kappa_i(\alpha)}{d\alpha} - \frac{p_i(\alpha)}{1 - p_i(\alpha)} \frac{d}{d\alpha}\left[E_{\Theta^n} \Sigma s(\boldsymbol{\theta}, \alpha)\right].$$

with strictly inequality holding for at least one i. Summing over all i, we obtain:

$$\begin{aligned}
\frac{d}{d\alpha}\left[E_{\Theta^n} \Sigma s(\boldsymbol{\theta}, \alpha)\right] &> \frac{\Psi(\tilde{\theta}_{\bar{j}}(\alpha))}{1 - p_{\bar{j}}(\alpha)} \sum_{i=1}^{n} \frac{d\kappa_i(\alpha)}{d\alpha} - \sum_{i=1}^{n} \frac{p_i(\alpha)}{1 - p_i(\alpha)} \frac{d}{d\alpha}\left[E_{\Theta^n} \Sigma s(\boldsymbol{\theta}, \alpha)\right] \\
&= \left(1 + \sum_{i=1}^{n} \frac{p_i(\alpha)}{1 - p_i(\alpha)}\right)^{-1} \frac{\Psi(\tilde{\theta}_{\bar{j}}(\alpha))}{1 - p_{\bar{j}}(\alpha)} \sum_{i=1}^{n} \frac{d\kappa_i(\alpha)}{d\alpha} \qquad (35)
\end{aligned}$$

Since $\sum_{i=1}^{n} \frac{d\kappa_i(\alpha)}{d\alpha} = 0$, it follows that $\frac{d}{d\alpha}\left[E_{\Theta^n} \Sigma s(\boldsymbol{\theta}, \alpha)\right]$ is positive. ■

Proof of Proposition 6: Let $\boldsymbol{k}'$ denote the liability profile after the buyout. That is, $k_i' = k_i \left(\sum_{\{i:k_i > \underline{k}\}}\right)^{-1}$ if $k_i \leq \underline{k}$ and 0 otherwise. Clearly, the conditions of Proposition 5 are satisfied, except for the fact that each remaining PRP's objective function (see expression (7)) is decremented by a constant number (the sum of *de minimis* PRPs' buyout payments). Specifically, assume that liability shares are ordered as usual so that $i > j$ implies $k_i \geq k_j$. Let $\bar{j} = \max\{j \in \{1, ..., n\} : k_j \leq \underline{k}\}$ and observe that $k_i' > k_i$ if $i > \bar{j}$ while $k_i' \leq k_i$ otherwise. Thus for each i the first order condition (8) remains valid and is independent of buyout payments. The conclusion of the proposition now follows immediately from Proposition 5. ■

References

Athey, S. (1997). Single crossing properties and the existence of pure strategy equilibria in games of incomplete information, *Technical report*, MIT.

Birdsall, T. H. and Salah, D. (1993). Prejudgment interest on superfund costs: Cercla's running meter, *Environmental Law Reporter* **23**: 10424+.

Church, T. W. and Nakamura, R. T. (1993). *Cleaning Up the Mess: Implementation Strategies in Superfund*, The Brookings Institution, Washington, D.C.

Dixon, L. S. (1994). Fixing superfund: The effects of the proposed superfund reform act of 1994 on transaction costs, *Technical Report RAND/MR-455-ICJ*, RAND.

Milgrom, P. and Shannon, C. (1994). Monotone comparative statics, *Econometrica* **62**(1): 1089–1122.

Rausser, G., Simon, L. and Zhao, J. (1998). Information asymmetries, uncertainties, and cleanup delays at superfund sites, *Journal of Environmental Economics and Management (forthcoming)* .

Zimmerman, R. (1988). Federal-state hazardous waste management policy implementation in the context of risk uncertainties, *in* C. E. Davis and J. P. Lester (eds), *Dimensions of Hazardous Waste Politics and Policy*, Greenwood Press, chapter 11, pp. 177–201.