ENVIRONMENTAL REMEDIES: AN INCOMPLETE INFORMATION AGGREGATION GAME

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

May 11, 2000

Abstract. The burden of resolving an environmental problem is typically shared among several responsible parties. To clarify the nature and extent of the problem, these parties must provide information to the regulator. Based on this information, the regulator will instigate an investigation of the problem, to determine an appropriate remedy. This paper investigates the incentives facing agents to promote excessive investigation and postpone remediation. Our incomplete information game-theoretic model may be of general interest to game theorists: we apply a new theorem guaranteeing pure-strategy equilbria and introduce a class of games called "aggregation games" which have interesting properties and are widely applicable.

JEL classification: D82, Q28

Keywords: Environmental economics; environmental remediation; Superfund; hazardous waste cleanups; incomplete information games; pure-strategy equilibria; aggregation games; strategic information transmission; strategic delay.

Corresponding author: Leo Simon, Department of Agricultural Economics, 207 Giannini Hall, University of California at Berkeley, Berkeley, CA 94709-3301. Email: simon@are.berkeley.edu. Phone: 510 642 8230.
Rausser: Dean and Robert Gordon Sproul Distinguished Professor, College of Natural Resources, University of California, Berkeley, CA 94720 and member, Giannini Foundation of Agricultural Economics; Simon: Adjunct Professor, Department of Agricultural Economics, University of California, Berkeley, CA 94720 and member, Giannini Foundation of Agricultural Economics; Zhao: Assistant Professor, Department of Economics, Iowa State University, Ames, IA 50011. The authors are grateful to seminar participants at the University of Washington, at the Economic Theory Seminar Series at Iowa State University, and at the Graduate Seminar, Department of Agricultural and Resource Economics, University of California at Berkeley. Susan Athey and Chris Shannon provided very helpful advice relating to the application of Athey's methodology to our problem. Larry Karp provided useful comments on an earlier draft.

We consider a class of problems in which several agents contribute to the creation of an environmental problem and are required to share the financial burden of remediation. In many instances, there is a natural information asymmetry between these Responsible Parties (RP's) and the regulatory authority responsible for implementing a remediation strategy. The asymmetry arises when RP's have private information regarding their individual contributions to the problem. For example, if the problem is a toxic contamination site, only the RP's themselves will have detailed information about the nature, volume and spatial/intertemporal distribution of the substances they have contributed. Alternatively, if the problem involves overuse of a common property resource, only the users themselves will be able to provide detailed information about their individual use-patterns. Typically, even individuals' own private information will be imperfect, because, for example, historical records may be incomplete and not readily accessible. However, RPs' private information about their own contribution will generally be more precise than the information that is directly available to any regulatory authority.

For this class of problems, a tradeoff arises between time and information. On one hand, it is obviously preferable that environmental problems be remediated sooner rather than later, especially when contamination is spreading and the public is exposed to health hazards. On the other hand, because, typically, the nature and magnitude of any given environmental problem is highly uncertain, it will generally be suboptimal to proceed rapidly to the implementation phase of the remediation process. Rather, time is needed to conduct field investigations that will provide more precise information about the characteristics and extent of the remediation task. If the investigation period is too short, the severity of the problem may be over- or underestimated: in the first instance, excessive resources will be allocated to mitigation; in the second, the remediation plan may be inadequate, resulting in costly revisions to the original remediation schedule and possibly exacerbated health risks.

The optimal length of the investigation period depends on the regulator's initial information: the greater the uncertainty, the longer is the optimal period, and hence the longer is the optimal delay before the problem is resolved. The issue of strategic information revelation naturally arises because the authority has to rely on the RPs for its initial information about the contamination. In order to determine an appropriate investigation program, the regulator must make an preliminary estimate of the uncertainty associated with the site. Since this estimate will be based on RP reports, RPs can, by strategically misreporting their private information, manipulate the regulator's decision and either hasten or delay the proceedings.

Even when the social costs of the environmental problem are fully internalized by the RP's as a group, individual RP's views about the optimal timing of remediation will typically differ from each other as well as from those of the regulator. One reason, of course, is that individual firms' intertemporal discount rates are higher than the social rate. For this reason alone, an RP has an incentive to manipulate the information transmission process, and thereby delay the remediation process without incurring legal sanctions. The Superfund program provides a dramatic illustration of delayed implementation. While for this program there are other factors leading to delay,[1] it has been observed that RPs benefit considerably from delayed implementation because it reduces their *discounted* cleanup costs (see, for example, Dixon (1994)). Cost savings due to discounting may be significant: as reported by Birdsall and Salah (1993), prejudgment interest is typically the single largest cost item at a Superfund site and accounts for nearly one-third of the total costs.[2]

Apart from differences between their respective rates of discount, there are other factors that lead RPs to prefer remediation schedules that differ from the socially optimal one. In particular, they have different degrees of risk aversion and face liabilities that differ in both their nature and extent.

---

[1] Primarily the enormously contentious issue of how to apportion burden shares among RP's.

[2] Prejudgment interest is accumulated when the government or an RP sues other RPs for past cleanup costs.

Typically, the regulator will be less risk averse than the RPs, because the former can pool risks across the many sites that it oversees. Moreover, even when they are in principle liable for all costs that are incurred as a result of delayed remediation, in practice RP's typically bear only a fraction of the incremental costs resulting from delay.[3] For most RP's, there is also the issue of insurance indemnification and the reluctance of insurance carriers to admit liability. For all these reasons, RPs 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 RP is responsible for only a fraction of total remediation costs, and for this reason will *underweight* the benefits of uncertainty reduction. The net effect of these differences is thus indeterminate.

In a previous paper (Rausser, Simon and Zhao (1998), henceforth RSZ), we examined the strategic interaction between RPs and the regulator in the specific context of Superfund cleanup. In that paper RP's strategically report the quality of their information. It identifies incentives for misreporting of accuracy levels and suggests several Bayesian mechanisms which would enable the regulator to extract the truth from the RPs. In the present paper, we investigate the abstract problem of delayed implementation, and formalize the interaction among the RPs as an incomplete information game, given the regulator's policy. To sharpen the analysis, we make three simplifying assumptions. First, we assume that while the magnitude of the environmental problem is a random variable, the *mean* of the distribution governing this variable is commonly known. Second, we assume that RPs' burden shares are predetermined. Third, we do not impose as an equilibrium condition that the regulator's decision rule elicits truthful behavior from the RP's. These simplifications allow us to single out other factors which affect PRPs' incentives to delay, to focus upon the

[3] For example, while health damage increases with exposure time, affected individuals and communities can obtain compensation for this damage only through a private cost recovery action. Such actions tend to be very costly to initiate and may not be successful. Accordingly the effect of delay on the extent of health damages is in practice unlikely to be fully internalized by RP's.

strategies the PRPs can pursue, and to identify the implications of some widespread government practices.

The second of these assumptions is problematic in some contexts—for example, as we have observed, burden shares are an enormously contentious issue in the context of environmental remediation—but perfectly reasonable in others—for example, in many instances, remediation activities are funded by special-purpose taxes levied on the basis of observable criteria.[4] The third assumption sets this paper apart from the vast agency-theoretic literature which focuses on the design of mechanisms that efficiently induce agents to truthfully reveal their private information. On the other hand, our approach is more consistent with actual regulator behavior. In most instances governmental bureaucrats typically accept reports from agents under their jurisdiction at face value, rather than attempting to reverse engineer "the truth" from these reports, based on what they know about the agents' motivations.

The paper is organized as follows. In Section 1, we formulate the incomplete information aggregation game among the RPs. Each RP 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* and assume that types are nonatomically distributed. A strategy for an RP is a function mapping its type into announced levels of precision. Once types have been realized, the regulator aggregates RP's announcements and imposes the investigation schedule that is optimal relative to this aggregated announcement. Applying Athey's methodology, we establish the existence of a pure strategy Nash equilibrium in which each RP's report is monotone in its type. In section 2 we study the class of "linear-quadratic" aggregation games and preview our subsequent results in this special context. Section 3 examines the role of burden shares. We demonstrate that RPs with higher burden shares on average report a higher level of uncertainty than those with lower shares. In Section 4, we show

[4] One recent example is the famous "penny-per-pound" tax on sugar production, proposed (although not ultimately adopted) as a way of financing the $700 million cost of restoring the Florida Everglades.

(for the case of two RPs) that under certain conditions, when their burden shares become more heterogeneous, the RPs' expected aggregate report increases. Section 5 extends this result to the case of multiple RPs, but imposes significant restrictions on functional forms. In Section 6 we apply our theory to the topic of Superfund cleanups, and study the implications of two widespread policy options: *de minimis* RP buyouts and the formation of RP steering committees. We discuss and conclude the paper in Section 7. Proofs are presented in the appendix.

1. The Model

We assume that there are n RPs, indexed by $i = 1, \ldots, n$, who have contributed to an environmental problem which must now be remediated. Parties' contributions are additive, i.e., the total magnitude of the problem is the sum of individual contributions. Prior to any evaluation of the problem, each RP has private information about its contribution level. This information is assumed to be imperfect, because, for example, historical records may be incomplete and not readily accessible. More precisely, we assume that i's contribution m_i is a random variable with mean $\bar{m}_i$ and variance θ_i. We assume that the $\bar{m}_i$'s are commonly known, while θ_i is known only by agent i. We assume that the m_i's are independent of each other, so that no RP can infer from its own uncertainty anything about the degree of uncertainty about other RPs' contributions. The total magnitude of the problem is denoted by Σm.[5] Thus, Σm is a random variable with commonly known mean $\Sigma \bar{m}$. Each RP has partial information about the variance of Σm, denoted by $\Sigma\theta$; that is, it knows the variance of its own contribution only. The regulator has no independent information at all about $\Sigma\theta$.

Each RP is required to make a report to the regulator about θ_i. Based on the aggregate of RPs' reports, the regulator determines a field investigation schedule that will generate further information

[5] Throughout this paper, 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 $\boldsymbol{x}$ $(resp. \boldsymbol{f(x)})$ by Σx (resp. $\Sigma f(x)$ or Σf) and the sum of all but the i'th element by Σx_{-i} (resp. Σf_{-i}).

about the magnitude of the problem, i.e., will reduce the variance of its estimate of Σm. Since it is common knowledge that the regulator's decision is a function of reported variances, each RP can, by strategically misreporting its uncertainty, manipulate this decision. Intuitively, if RP i prefers to delay the start of remediation beyond the socially optimal date, it will report a value of θ_i that exceeds its true value and thus prolong the investigation period.

Both the investigation and the remediation processes will be financed entirely by the RPs as a group. We assume that there is a predetermined vector of *burden shares*, or *burden profile*, $\boldsymbol{k} = (k_1, ..., k_n)$, with $k_i \geq 0$ and $\sum_{i=1}^{n} k_i = 1$: agent i is held responsible for the share k_i of total costs, whatever these costs turn out to be. (In some cases—e.g., Superfund cleanups,—the k_i's reflect individual agents' relative expected or estimated contributions to the toal problem. In others—e.g., when costs are financed by output taxation—burden shares reflect other considerations, such as market share.)

We formalize the interaction among the RPs as an incomplete information, simultaneous-move game. The i'th RP's *type* is identified with the variance, θ_i, of its contribution, which is known only by RP i. We assume that the θ_i's are identically and independently distributed on the interval $[\theta^l, \theta^u]$, where $\theta^l \geq 0$. Let $g(\cdot)$ denote the density of agents' types. We assume that $g(\cdot)$ is nonatomic on $[\theta^l, \theta^u]$. Let $\boldsymbol{\Theta} = [\theta^l, \theta^u]^n$ with generic element $\boldsymbol{\theta}$. Similarly, let $\boldsymbol{\Theta}_{-i} = [\theta^l, \theta^u]^{n-1}$. For $\boldsymbol{\theta}_{-i} \in \boldsymbol{\Theta}_{-i}$, let $\boldsymbol{g}_{-i}(\boldsymbol{\theta}_{-i}) = \prod_{j \neq i} g(\theta_j)$.

A *pure strategy* for the i'th RP is a function $s_i : [\theta^l, \theta^u] \rightarrow H = [\underline{\eta}, \bar{\eta}] \subset \mathbb{R}_+$ where H is the set of admissible variance announcements: $s_i(\theta_i)$ is the level of individual uncertainty declared by the i'th RP when its actual level of uncertainty is θ_i. The strategy vector $\boldsymbol{s} = (s_1, ..., s_n)$ is called a *pure strategy profile.* Thus a strategy profile is a mapping from $\boldsymbol{\Theta}$ to $\boldsymbol{H} = H^n$. The scalar $s_i(\theta_i)$ represents i's declared (as opposed to its actual) type. Of course, the notion that an RP would declare a number representing the variance of its contribution is no more than a convenient abstraction. In reality, an RP might report an upper and lower bound to its contributions. In this case, a high (resp. low) value of s_i would be interpreted as a wide (resp. negligible) gap between the two reported bounds.

Strategy profiles are mapped into *outcomes*. An outcome is a scalar, t, representing the length of the investigation period selected by the regulator. An important restriction on the outcome function, $t : \boldsymbol{H} \to \mathbb{R}_+$, is that this mapping depends only on the *sum* of agents' announcements. We shall refer to games that satisfy this restriction as *aggregation games*. Formally, the restriction is that if $\sum_{i=1}^n s_i(\theta_i) = \sum_{i=1}^n s_i(\theta_i')$ then $t\Big(s_1(\theta_1), ..., s_n(\theta_n)\Big) = t\Big(s_1(\theta_1'), ..., s_n(\theta_n')\Big)$. This restriction formalizes the idea that the regulator aggregates individual RPs' declarations before choosing an investigation schedule. To streamline notation, we will henceforth write t as a mapping from $\mathbb{R}$ to $\mathbb{R}$, with argument Σs. We assume that $t(\Sigma s)$ is thrice continuously differentiable and nondecreasing with Σs. We assume also that $t'(\cdot)$ is nonnegative and nondecreasing. Note that while the aggregation assumption is very natural in our context, the class of aggregation games is very special. In Cournot games, for example, payoffs depend both on the sum of agents' actions—because this sum determines prices—and on agent's individual actions—because these determine quantities. In a Bertrand, or auction game, on the other hand, what matters is the whole *profile* of agents' individual actions; the sum of agents' actions is immaterial.

As will soon become apparent, the bounds θ^l and θ^u on agents' strategies play a very important role in any aggregation game. A natural value for the lower bound, $\underline{\eta}$, on variance announcements would be zero, but for generality, we assume that $\underline{\eta}$ is nonnegative. On the other hand, we will choose $\bar{\eta}$ to exceed the optimal report for an RP whose type is θ^u and whose burden share is one, assuming that all other RPs report $\underline{\eta}$ with probability one. Now RPs' strategies are bounded below by $\underline{\eta} \geq 0$. Moreover it can be shown that their strategies both increase in their burden share and decrease in other agents' strategies. It follows that regardless of its type or burden-share, if an RP responds optimally to other agents' strategies, then this response must lie strictly below $\bar{\eta}$. Thus, while $\underline{\eta}$ will often constrain agents' behavior, $\bar{\eta}$ will never do so. This asymmetry will turn out to have important consequences for our model, and will drive our comparative statics results.

Before specifying the *payoff function*, F, for the game, we define a preliminary function, C, which represents the RPs' anticipated financial exposure. The only distinction between C and F is that the former depends on the outcome of the game (i.e., the length of the investigation period) while the latter depends on agents' strategies. Specifically, $C(t, \Sigma\theta, k) : \mathbb{R}_+ \times \mathbb{R}_+ \times [0,1] \to \mathbb{R}_+$ represents the anticipated financial exposure of an RP when the length of the investigation period is t, the variance of the contamination is $\Sigma\theta$ and the RP is responsible for a fraction k of total costs.[6] C is assumed to be thrice continuously differentiable in each of its arguments.

Obviously, C must increase with k, the RP's share of anticipated costs. We also assume that C increases with $\Sigma\theta$. This reflects RP risk aversion—intuitively, for risk-averse agents, anticipated costs will increase with uncertainty.[7] Even if RP's were risk neutral, however, anticipated costs would, typically, still increase with $\Sigma\theta$. To see this, consider the following, simple example: suppose, for example, that the regulator's decision rule is is to impose a remediation plan that "will be adequate 90% of the time." As uncertainty over the magnitude of the underlying problem increases, so also will the range of problems that lies within this percentile range, and, in turn, so also will the minimal remediation plan that will be satisfactory for all these possibilities[8]

The relationship between C 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

---

[6] We use the term *anticipated* exposure to distinguish C from the *expected* exposure function, which would depend on the random variable Σm rather than the (unobserved) statistic $\Sigma\theta$.

[7] In RSZ, we provide an explicit functional form for anticipated costs.

[8] This point is developed in more detail in Zimmerman (1988), in the context of Superfund cleanups.

remediation activities, thus reducing the present value of remediation costs. Both these effects tend to reduce C. On the other hand, more lengthy investigations are more costly. Moreover, they delay the remediation process, and thus increase the period during which society is exposed to the original environmental problem. Reflecting these considerations, we assume that C is convex in t and that for any values of $\Sigma\theta$ and k, an optimal investigation time (that minimizes C) exists.

We assume that $C_{t\Sigma\theta} < 0$, 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 given the types and strategies of other RPs, an RP's optimal investigation period increases with its uncertainty about the degree of contamination.

We also assume that $C_{tk} < 0$, so that $-C$ satisfies Milgrom and Shannon (1994)'s "single crossing property (SCP) in (t, k)," i.e., for all t, if $C_t(t, k_L) < 0$ and $k_L < k_H$, then $C_t(t, k_H) < 0$. The SCP is a necessary and sufficient condition for the conclusion that an RP's optimal investigation length increases in its burden share, provided that that other agents' strategies are monotone increasing in their types. RPs with larger burden shares prefer longer investigation periods because they assign more weight to uncertainty reduction. As Zimmerman (1988) point out, larger RPs weight uncertainty reduction more heavily because they will be held responsible for a larger share of the remediation. To illustrate with a simple example, suppose that RP i's anticipated exposure is jointly linear in the mean and the variance of $(k_i \Sigma m)$, that is, i's burden-share times the magnitude of the environmental problem. The mean of $(k_i \Sigma \bar{m})$ increases linearly with Σm, but the variance increases quadratically. Hence as k_i increases, the importance of the second moment relative to the first increases also. While this mean-variance example is illuminating, our model is much more general. For the comparative statics results we present below, all we need is that $C_{tk} < 0$. The mean-variance framework implies this condition but is not implied by it.

The payoff function for the game can now be written as a simple transformation of C. For $\eta \in H$, let $F((\eta, \boldsymbol{s}_{-i}), (\theta_i, \boldsymbol{\theta}_{-i}), k_i) \equiv C(t(\eta + \Sigma s_{-i}(\boldsymbol{\theta}_{-i})), (\theta_i, \boldsymbol{\theta}_{-i}), k_i)$ be the anticipated cost to RP i when its burden share is k_i, when the realized type-profile is $(\theta_i, \boldsymbol{\theta}_{-i})$, when i announces η and other RP's are playing $\boldsymbol{s}_{-i}(\cdot)$. Hence $\int_{\boldsymbol{\Theta}_{-i}} -F((\eta, \boldsymbol{s}_{-i}), (\theta_i, \boldsymbol{\vartheta}_{-i}), k_i)\boldsymbol{g}_{-i}(\boldsymbol{\vartheta}_{-i})d\boldsymbol{\vartheta}_{-i}$ is i's *expected payoff function* in the game between RP's.

Define the function $f : \mathbb{R}^3 \to \mathbb{R}$ by $f(\eta + \Sigma s_{-i}(\boldsymbol{\theta}_{-i}), \Sigma\theta, k_i) = \frac{dF((\eta, \boldsymbol{s}_{-i}), (\theta_i, \boldsymbol{\theta}_{-i}), k_i)}{d\eta}$. f is the marginal cost to i of increasing his announcement, given the realized types of all other RP's. If $\boldsymbol{s}$ is a pure-strategy equilibrium profile, then for each i and each θ_i, $s_i(\theta_i)$ must minimize $E_{\boldsymbol{\theta}_{-i}} F((\cdot, \boldsymbol{s}_{-i}), (\theta_i, \boldsymbol{\vartheta}_{-i}), k_i)$ on $[\underline{\eta}, \bar{\eta}]$. Since by construction, the upper bound $\bar{\eta}$ is never binding, a necessary condition for $\boldsymbol{s}$ to be an equilibrium profile is that:

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

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

To conclude this section, we establish that our incomplete information game has a pure strategy Nash equilibrium (PSNE). 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* (cf. the definition of SCP in (t, k) on p 9), 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, RP i's expected payoff function $-\int_{\boldsymbol{\Theta}_{-i}} F((\eta, \boldsymbol{s}_{-i}), (\theta_i, \boldsymbol{\theta}_{-i}), k_i)\boldsymbol{g}_{-i}(\boldsymbol{\theta}_{-i})d\boldsymbol{\theta}_{-i}$ satisfies:

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

since $\frac{dt}{d\Sigma s} > 0$ and $C_{t\Sigma\theta} < 0$. Thus $-E_{\boldsymbol{\theta}_{-i}} F$ satisfies SCP in (η, θ_i), implying that a PSNE exists.

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* $\underline{\eta}$ *on* $[\theta^l, \tilde{\theta}_i)$*, and strictly exceeds* $\underline{\eta}$ *on* $(\tilde{\theta}_i, \theta^u]$.[9] *Moreover,* s_i *is increasing and continuously differentiable on* $(\tilde{\theta}_i, \theta^u]$.

2. Aggregation Games: the linear-quadratic case

In this section, we abstract from the specific details of the model developed above and consider a particularly tractable subclass of aggregation games. We will say that an aggregation game is *linear-quadratic* if its *outcome function* is affine in the sum of agents' actions and if agents' *payoffs* are quadratic in the outcome of the game. Preserving where possible the notation in section 1, we assume that the i'th agent's pure strategies map Θ_i to $H_i = [\underline{\eta}_i, \bar{\eta}_i]$ and that the outcome function t is an affine, increasing function from $\boldsymbol{H} = \prod_{i=1}^n H_i$ to $\mathbb{R}$ such that for any two profiles $\boldsymbol{\eta}$ and $\boldsymbol{\eta}'$ such that $\Sigma\eta = \Sigma\eta'$, we have $t(\boldsymbol{\eta}) = t(\boldsymbol{\eta}')$. We assume further that the payoff function for the θ'th type of the i'th agent, $F_i(\cdot, \theta) = F(\cdot, \theta, k_i)$ is quadratic in t, with $F_i'' < 0$. Finally, we assume that payoffs satisfy Athey's qualified SCP condition (see page 10), so that the existence of a pure-strategy equilibrium is guaranteed. Note that this class of games is more general in some respects, but less general in others, than the class of games defined in section 1.

We impose one additional assumption, which corresponds to the condition on page 7 relating to the upper bound, $\bar{\eta}$, on actions: we assume that for each i and θ, there exists $\tilde{\eta}_i \in H_i$ such that $F_i(\cdot, \theta)$ is globally maximized at $t\big(\tilde{\eta}_i + \sum_{j \neq i} \underline{\eta}_j\big)$. This last assumption implies that the *upper* bound on i's set of admissible actions will *never* be a binding constraint on i's behavior: even if for all $j \neq i$, the j's agent selects its lowest possible action with probability one, i can, by selecting $\tilde{\eta}_i$,

[9] Note that if $s_i(\cdot) > \underline{\eta}$ on Θ then the interval $[\theta^l, \tilde{\theta}_i)$ will be empty.

singlehandedly induce i's most preferred outcome; since the outcome function is monotone, i will thus never wish to select an action that exceeds $\tilde{\eta}_i$.

Linear-quadratic aggregation games exhibit a very striking property. Assume that in a pure-strategy equilibrium, the action taken by type θ of player i exceeds the lower bound on i's action space. Then the expected outcome of the game coincides with the outcome that globally maximizes i's payoff function. Accordingly, with the exception of player-types who bid their lower bounds, every player of every type achieves, in expectation, the very best outcome he can possibly achieve! Formally:

Proposition 2 (Linear-quadratic aggregation games). *Suppose that* $\boldsymbol{s}$ *is a pure-strategy equilibrium for a linear quadratic aggregation game and that for some* i *and* $\theta \in \Theta_i$, $s_i(\theta) > \underline{\eta}_i$. *Then* $E_{\boldsymbol{\Theta}}\left[t\big(\Sigma s(\cdot)\big)|\theta_i = \theta\right] \;=\; \text{argmax } F_i(\cdot, \theta)$.

We include the proof in the text since it is both simple and instructive. Note in particular the critical role of the linear-quadratic specification: it allows us to pass the expectation operator through two sets of parentheses.

Proof of Proposition 2: As argued above, the upper bound on i's strategy space is never binding. Hence a necessary condition for $s_i(\theta) > \underline{\eta}_i$ to be optimal is that $E_{\boldsymbol{\Theta}_{-i}}\left[F_i'\Big(t\big(s_i(\theta) + \Sigma s_{-i}(\cdot)\big), \theta\Big)\right] = 0$. By assumption $t(\cdot)$ is affine in Σs. Moreover, since F is quadratic, F' is affine also. Hence $F_i'\Big(t\big(s_i(\theta) + E_{\boldsymbol{\Theta}_{-i}}\left[\Sigma s_{-i}(\cdot)\right]\big), \theta\Big) = 0$. Since $F'' < 0$, there is a unique $\tilde{t}$ such that $F_i'(\tilde{t}, \theta) = 0$. Hence $E_{\boldsymbol{\Theta}}\left[t\big(\Sigma s(\cdot)\big)|\theta_i = \theta\right] = t\big(s_i(\theta) + E_{\boldsymbol{\Theta}_{-i}}\left[\Sigma s_{-i}(\cdot)\right]\big) = \tilde{t}$. ■

This result is puzzling at first sight. If the optimal outcome for i is strictly greater when i's type is θ' than when it is θ, how can the realized outcome be optimal for both types? In the same vein, if the optimal outcome for type θ_i of agent i differs from that for type θ_j of player j, how can the outcome when both types are realized be optimal for both players *simultaneously*? In particular, if

j's preferred outcome is lower than i's, why is there not an infinite tug-of-war, in which each agent tries to counteract the other's effect on the aggregate action?

These puzzles are, of course, only artificial. The answer to the second puzzle is that the proposition refers to *conditional* outcomes, and each type of i and j are conditioning on different events. To illustrate, suppose that both types are uniformly distributed on the unit interval and that the set of admissible actions is $[0, 2]$. Suppose furthermore that argmax $F_i(\cdot, \theta) = \theta + 1/4$, while argmax $F_j(\cdot, \theta) = \theta$. Now consider a game with outcome function $t = \Sigma s$. In this case, the following pair is (the unique) pure-strategy equilibrium: $s_i(\theta) = \theta$, while $s_j = 0$, if $\theta < 1/2$, and $\theta - 1/2$, otherwise. To see that these strategies are optimal, observe that the expected outcome conditional on $\theta_i = \theta$ is $\theta + 1/4$, while the expected outcome conditional on $\theta_j = \theta$ is $1/2$, if $\theta < 1/2$, and θ otherwise. Hence for each type of player i and each type θ of player j, $\theta \leq 1/2$, the expected outcome conditional on each type being realized is globally optimal. For future reference, note that the (unconditional) expected outcome of the game is $3/4$.

The preceding example also demonstrates why an infinite tug-of-war between i and j does not arise. In general j cannot tug hard enough against i, because j's action space is bounded below by $\underline{\eta}_j$. Because we chose $\bar{\eta}_i$ sufficiently high (relative, of course, to $\underline{\eta}_j$) i is not similarly constrained. In our example, types of player j below $1/2$ would prefer to select negative actions to offset i's expected action, but negative actions are infeasible. Thus the asymmetry between the upper and lower bounds on agents' action spaces ensures that i will always win the tug-of-war. Note that this asymmetry is not simply a game-theoretic contrivance. In many applications, the natural action space is the set of nonnegative numbers: that is, agents can reasonably select arbitrarily large positive actions, while negative actions have no natural interpretation.

As a preview to the heterogeneity results in sections 4 and 5 below, consider the following "mean-preserving spread" of the payoffs in our original example: argmax $F_i(\cdot,\theta) = \theta + 1/2$, while argmax $F_j(\cdot,\theta) = \theta - 1/4$. That is, all types of player i want more of t than originally, while all types of player j want less. The pure-strategy equilibrium for the perturbed game is: $s_i(\theta) = \theta + 1/3$, while $s_j = 0$, if $\theta < 2/3$, and $\theta - 2/3$, otherwise. The (unconditional) expected outcome of this game is unity. That is, a mean preserving spread in preferences *increases* the expected outcome of the game. Note that all types of i continue to get exactly what they want, while more types of player j are thwarted, and thwarted to a greater extent. This asymmetry arises, of course, because of the asymmetry in the bounds on agents' action spaces.

3. Burden shares and preferences over investigation lengths

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 burden shares. That is, players who bear more responsibility for the remediation will select larger actions than players of the same type with smaller shares. While the proof of this result is quite technical, the basic idea is straightforward. Loosely, RP's with higher burden shares are more sensitive to changes in the level of uncertainty than those with smaller shares (see Lemma 2.1 below). Hence those with higher shares will prefer longer investigation periods, and will, therefore, choose larger values of η 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 RP, when the types and strategies of other RPs are given.

Lemma 2.1. *(a) For each i, $\boldsymbol{s}$ and each $\boldsymbol{\theta}$, $f(\eta + \Sigma s_{-i}(\boldsymbol{\theta}_{-i}), \Sigma\theta, k_i)$ decreases w.r.t. k_i. (b) There exists $\epsilon > 0$ such that if either $|C_t| < \epsilon$ or $|t''(\cdot)| < \epsilon$, then $f(\eta + \Sigma s_{-i}(\boldsymbol{\theta}_{-i}), \Sigma\theta, k_i)$ increases in η.*

Part (a) simply says that an increase in an agent's burden share raises the marginal benefit of over-reporting when an RP knows the types of all other RPs. That is, higher burden 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 F is convex in η. The condition is that *either* the regulator's response function is sufficiently close to linear in the RPs' 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 F. We will henceforth simply assume (Assumption 1 below) that F is convex in η. This assumption ensures that if an RP knew for certain the types of all other RP's, the Kuhn-Tucker conditions in (1) would be both necessary and sufficient for cost minimization.

Assumption 1. *For each i, $\frac{dF}{d\eta}$ increases with η.*

Now consider a given monotone PSNE strategy profile. Lemma 2.1 implies the following result relating the strategies of any two RPs i and j, with $k_i > k_j$: $s_j(\theta) = \underline{\eta}$ at any point θ at which the difference between $s_i(\cdot)$ and $s_j(\cdot)$ is minimized.

Lemma 2.2. *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) for any 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) if all types of a given agent select actions greater than $\underline{\eta}$, then that agent's burden share must weakly exceed all other RPs' shares.

[10] To see this, observe that if the variances of the g_j's were zero, then in any pure-strategy equilibrium C_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 C_t will be less than ϵ with probability $1 - \epsilon$.

Proposition 3 follows almost immediately from Lemma 2.2. It states that in a PSNE, RPs with larger burden shares have a greater tendency to over-report their uncertainty than smaller RPs.

Proposition 3 (**Monotonicity w.r.t. burden shares).** *Let* $\mathbf{s}$ *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* $(\tilde{\theta}_i, \theta^u]$,

4. Comparative Statics analysis: two RP's

In section 3 we compared the strategies of different RPs within a given game. In this section, we compare two-player games with different burden profiles. We begin by showing (Prop 4) that a transfer of burden from one RP to another results in an increase in the latter's strategy and a reduction in the former's. We are, of course, ultimately more interested the *net* effect of these changes on the expected length of the investigation period. While we cannot make any general statements about this, we can (Prop 5) identify conditions under which a transfer of burden from the smaller RP to larger one will result in an increase in the expected investigation time.

Consider two burden profiles (k_1, k_2) and (k'_1, k'_2), with $k'_1 > k_1$ and $k'_2 < k_2$. We will show that in the latter profile, 1's announcement will be higher, and 2's will be lower, than in the former one, except when either is equal to $\underline{\eta}$. This result holds regardless of relative sizes of k_1 and k_2.

Proposition 4 (**Heterogeneity: two RPs).** *Consider two burden profiles* (k_1, k_2) *and* (k'_1, k'_2) *such that* $k'_1 > k_1$ *and* $k'_2 < k_2$. *Then* $\tilde{\theta}'_1 < \tilde{\theta}_1$ *and* $s'_1(\theta) > s_1(\theta)$ *for all* $\theta > \tilde{\theta}'_1$. *Similarly,* $\tilde{\theta}'_2 > \tilde{\theta}_2$ *and* $s'_2(\theta) < s_1(\theta)$ *for all* $\theta > \tilde{\theta}_2$.

This result is quite intuitive. As an RP's burden share increases, its preferred investigation length increases also, so that *holding 2's strategy constant*, an increase in k_1 will result in an increase in 1's announcement. Similarly, holding 1's strategy constant, a decrease in k_2 must result in a decrease in 2's announcement. Thus if $s'_1(\theta) \leq s_1(\theta)$ for some $\theta \geq \tilde{\theta}_1$, then there must be some type of 2 whose announcement has increased. Indeed, the largest decrement in 1's announcement (which depends on 1's type) must be more than offset by the largest increment in 2's announcement. On the other hand, if $s_2'(\theta) > s_2(\theta)$ the largest increment in 2's announcement must be more than offset by the largest decrement in 1's announcement. But these two requirements are mutually contradictory and the proposition now follows.

We now consider the *net* effect on investigation length of increasing the degree of hetorogeneity between RP's. In general, this issue cannot be resolved without an extensive specification of the third order derivatives of the RPs' payoff functions. We can, however, obtain a determinate result for games that are "nearly" linear-quadratic (cf. section 2), provided that one additional condition is satisfied. Specifically, suppose that: (a) the third order derivatives of agents' payoffs are relatively insignificant; (b) the regulator's response function is sufficiently close to linear; (c) for most types of the larger RP, the lower bound $\underline{\eta}$ is not binding. In this case, an increase in heterogeneity will increase the expected length of investigation (cf. the example on page 14). The basic idea is extremely simple: as the larger RP's share increases, it prefers longer investigations; under (c), most of the time, it gets more-or-less what it wants (cf. Proposition 2). Under conditions (a) and (b), we can "almost" pass the expectation operator through the two integral signs, as we did in the proof of Prop 2.

To fully understand this result, the "tug-of-war" discussion on page 13 is critical. The smaller RP's strategic options are significantly limited by the lower bound on the action space, while the larger RP is not comparably constrained. To see this, observe once again that as an RP's burden share increases (decreases), its preferred investigation length increases (decreases) also. Any decrease in the smaller RP's announcement can be more than counterbalanced by the larger RP, who can simply increase its own announcement. The reverse is not true: the smaller RP cannot counterbalance the larger's increased announcement, because the smallest possible announcement it can make is $\underline{\eta}$.[11]

To make the above ideas precise, we need some preliminary constructions. Fix two burden profiles $\boldsymbol{k}$ and $\boldsymbol{k}'$, with $k_1' > k_1 > k_2 > k_2'$. 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 burden profile $\boldsymbol{\kappa}(\alpha)$. 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[12]

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

then set $\tilde{\theta}_r(\alpha)$ equal to θ_r'; otherwise set $\tilde{\theta}_r(\alpha)$ equal to θ^u. 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 4, we know that $\tilde{\theta}_1(\alpha)$ decreases, while $\tilde{\theta}_2(\alpha)$ increases with α. To implement condition (c) above, we choose $\boldsymbol{k}'$ so that $\tilde{\theta}_1(\alpha) > \theta^l$, for all $\alpha < 1$, but $\tilde{\theta}_1(1) = \theta^l$. That is, as α approaches one, the lower bound on actions binds for fewer and fewer types of RP 1. 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 5 (Heterogeneity and delay: two RPs). *There exists $\epsilon_C > 0$ and $\epsilon_t > 0$ such that if (a) C_{tt} is ϵ_C-flat*[13] *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 $[\bar{\alpha}, 1]$.*

Because the proof of Proposition 5 is long and complex, we provide an outline in the text. 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] More concretely, there is no natural upper bound to the level of uncertainty one can declare. One cannot, however, assert that the level of uncertainty is negative!

[12] $\neg r = 2$, if $r = 1$, and 1 if $r = 2$.

[13] Recall from page 9 that C is assumed to be convex.

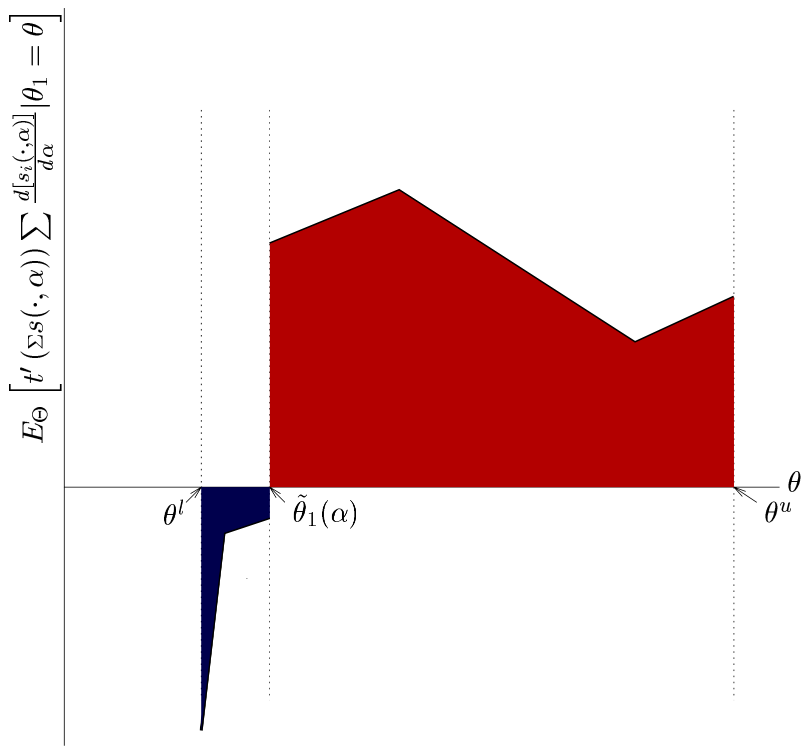

FIGURE 1. Total shaded area must be positive

Graphically, this inequality amounts to the requirement that the shaded area below the axis in figure 1 is dominated by the shaded area above it. To prove (4), we establish two properties of the graph: (i) to the right of $\tilde{\theta}_1(\alpha)$, it is positive; and (ii) to the left of this point, it is bounded below. The inequality follows from these properties, provided that $\tilde{\theta}_1(\alpha)$ is sufficiently close to θ^l. Conditions (a) and (b) above are used to prove (i); (c) corresponds to the proviso. By far the hardest task is to prove (ii): informally, we need to rule out the possibility of a tug-of-war between 1 and 2 *in derivative space*; in this case we cannot invoke a natural lower bound analogous to $\underline{\eta}$.[14]

Returning to the details, note from (1) and (3) that for each α and $\theta_1 > \tilde{\theta}_r(\alpha)$,

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

[14] We were unable to generalize step (ii) to the n-person case. Had we been able to do so, the remainder of the proof would have generalized quite straightfowardly.

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

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

Now $\kappa_1(\alpha)$ increases with α and C_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:

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

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 $\frac{\partial}{\partial t}\left[C_t(\cdot,\cdot,\cdot)t'(\cdot)\right] = (C_{tt}t' + C_t t'')$ *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., C_{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:

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

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.

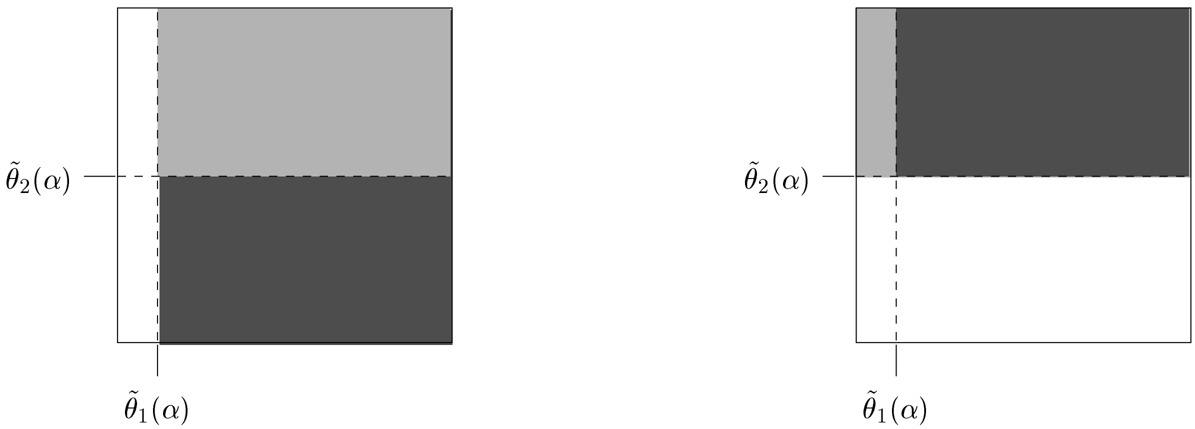

FIGURE 2. Intuition for Lemma 5.2.

The formal argument that $\frac{d[s_2(\cdot,\alpha)]}{d\alpha}$ is bounded requires two steps (see lemma 5.2). 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 $\frac{\partial}{\partial t}\left[C_t(\cdot,\cdot,\cdot)t'(\cdot)\right]$ 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 4) 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.

5. Comparative Statics analysis: multiple RP's

Matters are more complex when there are multiple RP's. Accordingly, we will analyze the special case in which (a) the social and private anticipated financial exposure functions are quadratic in investigation time and the sum of agents' private uncertainty parameters, and (b) the regulator's response function would minimize society's expected financial exposure, if RPs were to truthfully revealed their types. Under these restrictions, the regulator's response function will be linear in agents' actions. Consequently, the induced incomplete information game will belong to the class of linear-quadratic aggregation games, defined in section 2.

The linear-quadratic specification clearly oversimplifies the nature of our time-information tradeoff. Its offsetting benefit is that the comparative statics result we obtain is quite transparent, providing useful insights into the structure of our model. Moreover, as will be clear from the structure of its proof, the conclusion of Proposition 6 below will hold more generally, provided that third-order effects are sufficiently small relative to lower-order effects.

Let C^s denote society's anticipated financial exposure when the length of the investigation period is t and the vector of RP types is $\boldsymbol{\theta} = (\theta_1, ..., \theta_n)$, with $\theta_i \in \Theta = [\theta^l, \theta^u]$:

$$C^s(t, \boldsymbol{\theta}) \quad = \quad 0.5\beta_{11}t^2 \; - \; \beta_{12}t\Sigma\theta \; + \; \beta_{22}\left(\Sigma\theta\right)^2, \text{ where } \beta_{11},\ \beta_{12},\ \beta_{22} > 0.$$

Note that $C^s_{tt} > 0$ while $C^s_{t\Sigma\theta} < 0$. Thus, C^s satisfies all of the conditions we imposed on the RPs' anticipated exposure function C (see page 8), except that, of course, burden shares are not involved, and $C^s_{\Sigma\theta}$ is not globally positive. In a moment, we will identify a restriction which will ensure that $C^s_{\Sigma\theta}$ *is* positive on the relevant range of the function.

We model our regulator as choosing t to minimize society's anticipated financial exposure, assuming that agents' type announcements are truthful (see page 4 for a discussion of this assumption). That is, the outcome function $t^*(\cdot)$ is defined as follows:

$$t^*(\Sigma s) \quad = \quad \text{argmin}_t\ C^s(t, \Sigma s) \quad = \quad \text{argmin}_t\ \left\{0.5\beta_{11}t^2 \ - \ \beta_{12}t\Sigma s \ + \ \beta_{22}\left(\Sigma s\right)^2\right\}$$

Clearly $t^*(\Sigma s) = \beta_{12}\Sigma s/\beta_{11}$ so that for each i, $\frac{dt^*(\Sigma s)}{ds_i} = \beta_{12}/\beta_{11}$. Note that $t^*(\Sigma s)$ is bounded above by $n\beta_{12}\theta^u/\beta_{11}$.

The regulator's decision rule induces a private financial exposure function for a RP with burden share $k \in (0,1)$. We assume that the RPs' exposure 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 RPs tend 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 RP is responsible for only a fraction of total remediation costs: an RP with burden share k is exposed to the fraction k of anticipated social exposure; the variance of the RP's exposure, however, is only k^2 times the variance of social exposure. Hence, taking as given the regulator's response function, $t^*(\cdot)$ and the sum of other agents' strategies, $\Sigma s_{-i}(\cdot)$, this RP's anticipated financial exposure is:

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

We will assume that even if all burden 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.

RP 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^{n-1} to the sum of other agents strategies, $s_i(\theta)$ must satisfy the first order condition:

$$0 \;\leq\; \int_{\Theta^{n-1}} \left[\left(k_i \gamma_{11} t^* \left(s_i(\theta_i) + \Sigma s_{-i}(\boldsymbol{\theta}_{-i}) \right) \;-\; k_i^2 \gamma_{12} (\theta_i + \Sigma\theta_{-i}) \right) \frac{dt^*(\Sigma s)}{ds_i} \right] \boldsymbol{g}_{-i}(\boldsymbol{\theta}_{-i}) d\boldsymbol{\theta}_{-i}$$

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

$$0 \;=\; \int_{\Theta^{n-1}} \left[s_i(\theta_i) + \Sigma s_{-i}(\boldsymbol{\theta}_{-i}) \;-\; k_i \frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}} (\theta_i + \Sigma\theta_{-i}) \right] \boldsymbol{g}_{-i}(\boldsymbol{\theta}_{-i}) d\boldsymbol{\theta}_{-i}, \text{ for all } \theta_i \geq \tilde{\theta}_i \qquad (8)$$

so that, defining $\tilde{\theta}_i$ implicitly by $k_i \frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}} (\tilde{\theta}_i + E_{\Theta^{n-1}} \Sigma\theta_{-i}) \;=\; \underline{\eta} + E_{\Theta^{n-1}} \Sigma s_{-i}(\boldsymbol{\theta}_{-i})$, i's strategy can be written as:

$$s_i(\theta_i) \;=\; \begin{cases} \underline{\eta} & \text{if } \theta_i \leq \tilde{\theta}_i \\ k_i \frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}} (\theta_i + E_{\Theta^{n-1}} \Sigma\theta_{-i}) \;-\; E_{\Theta^{n-1}} \Sigma s_{-i}(\boldsymbol{\theta}_{-i}) & \text{if } \theta_i > \tilde{\theta}_i \end{cases}$$

The following proposition establishes that under the linear-quadratic specification, any shift in burden from smaller RPs to larger ones results in an increase in the expected investigation period. More precisely, arrange RP's by burden share in decreasing order, so that RP's with larger burdens have smaller indices. Now partition the set $\{1, ..., n\}$ into two subsets, i.e., larger RP's are to the left and smaller are to the right. Now consider a new burden profile in which shares for RP's in the left subset are no smaller, while shares for those in the right subset are no larger, than originally. Then in the new equilibrium, the expected value of the outcome function increases. Since by assumption, the investigation period is too long even when firms' burden shares are equal, the burden shift further exacerbates the delay.

Proposition 6 (**Heterogeneity and delay: multiple RPs**). *Consider a burden profile,* $\boldsymbol{k}$*, such that not all burden shares are equal. Arrange RPs so that* $i < j$ *implies* $k_i \geq k_j$*. Now consider a shift in burden shares to* $\boldsymbol{k}' \neq \boldsymbol{k}$*. Assume that for some* $\bar{j} \in \{1, ..., n-1\}$*,* $k_i' \geq k_i$*, for all* $i \leq \bar{j}$ *and* $k_i' \leq k_i$ *otherwise. Let* $\boldsymbol{s}$ *and* $\boldsymbol{s}'$ *be PSNE's corresponding to these profiles. Then* $E_{\Theta^n}\, t(\Sigma s'(\boldsymbol{\theta})) > E_{\Theta^n}\, t(\Sigma s(\boldsymbol{\theta}))$*.*

6. A policy application: Superfund cleanups

In this section we apply our theoretical framework to evaluate two practices that are widespread in the management of Superfund cleanups. In particular, we consider the effect on cleanup delay of *de minimis* RP buyouts and the formation of steering committees by RP'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 linear-quadratic specification in section 5. As we noted above, however, the key comparative statics result we obtain in that section, Proposition 6, will hold more generally, provided that third-order effects do not dominate lower-order effects.

6.1. *de minimis* RP buyouts.

A RP is classified as *de minimis* if its burden share falls below some (small) critical fraction. In a *de minimis* RP buyout, small RPs pay the regulator a fixed amount in exchange for relief from all future liabilities. Typically the buyout price that a given RP will pay will more than cover its *expected* burden, because the RP 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 RPs 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 RPs originally had to bear will be transferred to the larger RPs that remain. Thus although the *expected* burden borne by each remaining RP will remain the same, or will actually decline if the *de minimis* parties pay a risk premium, 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 RP 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 RP, buyout revenue is a lump-sum transfer; it reduces the RP's exposure by an amount that is independent of the variable (private uncertainty level) over which the RP has discretion. The following proposition summarizes this discussion.

Proposition 7 (*de minimis* buyouts and delay). *Consider a burden profile,* $\boldsymbol{k}$*, and fix* $\underline{k} > 0$ *such that for some* j*,* $k_j \leq \underline{k}$*. For the model specified in section 5, a* de minimis *buyout of* all *RPs whose burden shares do not exceed* $\underline{k}$ *will increase the expected investigation length. Moreover, this increase will be invariant with respect to the magnitude of the buyout premiums.*

6.2. **RP steering committees.** Just as regulators have encouraged small RPs to settle early via *de minimis* buyouts, they have also encouraged large RPs 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 RP's, and reduce contentious litigation over burden shares. Once again, however, our model highlights an attribute of steering committees that could have potentially major negative social consequences.

As we have seen, 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 noted above, an individual RP's share of total expected cost increases in proportion to its burden share, while the uncertainty associated with its liability exposure increases with the *square* of its burden share. Thus, larger RPs assign a greater relative importance to uncertainty than smaller ones. Consequently cleanup sites with many small RPs are likely to be less subject to delay than ones with a few large RPs. When a steering committee is formed, the interests of several RP's are coalesced: effectively, several smaller RP's are replaced by a single large one. From the RPs' 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.[15]

The formation of a steering committee can be represented as a change in burden profile. Consider a burden profile $\boldsymbol{k}$, ordered as usual so that $i < j$ implies $k_i \geq k_j$. Suppose that a subset $J \subset \{1, ..., n\}$ of RPs, $\#J \geq 2$, form a committee, and let $\bar{j} = \min\{j \in J\}$ be the RP in J whose liability share is the largest. After the steering committee is formed, there is, effectively, a new liability profile in which $\bar{j}$'s share is now equal to the sum of the shares of all RP's in J, and the shares of all RP's in J except $\bar{j}$'s are reduced to zero. Formally, define the profile $\boldsymbol{k}' \neq \boldsymbol{k}$ as follows: $k'_i = \sum_{j \in J} k_j$, if $i = \bar{j}$; $k'_i = k_i$, if $i \notin J$; and $k'_i = 0$ otherwise. Observe that $\boldsymbol{k}'$ satisfies the condition of Proposition 6, since $\boldsymbol{k}' \neq \boldsymbol{k}$, $k'_i \geq k_i$ for $i \leq \bar{j}$, and $k'_i \leq k_i$ otherwise. Consequently, the following result is an immediate corollary of Proposition 6.

Proposition 8 (Steering Committees and delay). *For the model specified in section 5, if at least two RP's form a steering committee, then the expected investigation length will increase.*

[15] Of course, the social benefits of reduced litigation may offset this loss, but in this paper, these benefits are unmodeled.

7. Conclusion

In this paper, we study an incomplete information game among parties who have contributed to the creation of an environmental problem and are now responsible for its remediation. A key assumption is that these RPs are better informed about their individual contributions to the problem than the regulator. Provided that the distributions of RPs' types are nonatomic, there exists a pure-strategy Nash equilibrium in which each RP's report about its level of uncertainty is positively related to its true level. Moreover, larger RP's are more likely to over-report than smaller RP's. Under certain conditions, aggregate over-reporting increases with the degree to which RPs' burden shares are heterogeneous.

While this paper focuses primarily on environmental policy and regulation, the model we present may be of general interest to game-theorists. Our model applies a very powerful theorem by Athey (1997), which to our knowledge has not yet been applied to the theory of regulation. The theorem guarantees existence of pure strategy equilibria for incomplete information games satisfying a condition that is very natural in many applications. We can thus benefit for the well known and very considerable advantages that pure strategies have over mixed strategies. In addition, the game we model belongs to a class of games called *aggregation games*, which are applicable in a wide variety of economic and political contexts. This class has the property that each agent's payoff depends only on the sum of all agents' actions. Under certain conditions, games of this kind exhibit a property that is both very striking and facilitates comparative static analysis.

Modified versions of the game presented in this paper can be applied to a wide range of regulatory problems. In particular, our model may be applicable whenever there is incomplete information and an aggregation filter is applied to individual actions. One class of problems arise in competition or antitrust regulation of collusion. Here each firm has private information and firms' profits depend on the sum of all firms' pricing and output decisions (Athey, Bagwell and Sanchiroico, 1998). In the implementation of collusive pricing, the vector of market shares is the publicly known parameter, playing the role that the k_i's play in our model. Strategic behavior in the presence of an aggregation filter also arises in the context of group decision-making, when individual group members have incentives to manipulate the information transmission process. For example, in a model of committee voting behavior, the parameter k_i might represent the i'th committee member's publicly known indvidual stake in the issue at hand.

Finally, our results have important policy implications. Current regulatory policy does not include guidelines for strategically extracting accurate information from RPs. Furthermore, some widespread practices in the management of environmental remediation, such as the formation of RP steering committees and de minimis RP buyouts, may exacerbate strategic behavior. This paper demonstrates that the issue of strategic information transmission may be significant.

References

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

Athey, S., Bagwell, K. and Sanchiroico, C. (1998). Collusion and price rigidity, *Working paper*, MIT Department of Economics.

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

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. C., Simon, L. K. and Zhao, J. (1998). Information asymmetries, uncertainties, and cleanup delays at Superfund sites, *Journal of Environmental Economics and Management* **35**(1): 48–68.

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.

Appendix: 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 2.1: Part (a) follows from the fact that $f = \frac{dF}{d\eta} = \frac{dC}{dt}t'$ and $t' > 0$ while $C_{t\kappa} < 0$. To prove (b), observe that $\frac{d^2F}{d{s_i}^2} = C_{tt}t' + C_t t''$ with $C_{tt} > 0$ and $t' > 0$. The sign of $f_{\Sigma s} = \frac{d^2F}{d{s_i}^2}$ will be determined by the first of these terms, provided that either $|C_t|$ or $|\frac{d^2t}{d\Sigma s^2}|$ is sufficiently small. This proves part (b) of the Lemma. ■

Proof of Lemma 2.2: Pick $\underline{\theta}_i \in \Theta_{ij}$ and let $\hat{\gamma} = s_i(\underline{\theta}_i) - s_j(\underline{\theta}_i)$. We have

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

The strict inequality follows from Lemma 2.1(a), i.e., f decreases w.r.t. k_i; the weak inequality holds because for all ϑ, $s_i(\vartheta) \geq s_j(\vartheta) + \hat{\gamma}$ and, by Lemma 2.1(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 $\theta \in U$, $0 < \frac{dE_{\boldsymbol{\theta}_{-j}} F((s_j(\theta), \boldsymbol{s}_{-j}), (\theta, \boldsymbol{\theta}_{-j}), k_j)}{d\eta}$ and hence, from (1), $s_j(\theta) = \underline{\eta}$. ■

Proof of Proposition 3: To prove part (a), suppose $\tilde{\theta}_i \geq \tilde{\theta}_j$. 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, $\operatorname{argmin}_{\vartheta \in [\theta^l,\theta^u]}(s_i(\vartheta) - s_j(\vartheta)) \cap [\tilde{\theta}_j, \theta^u] \neq \emptyset$. But for all $\theta > \tilde{\theta}_j$, $s_j(\theta) > \underline{\eta}$, contradicting Lemma 2.2.

To prove part (b), assume that for some $\underline{\theta}_i \in (\tilde{\theta}_i, \theta^u]$, $s_i(\underline{\theta}_i) \leq s_j(\underline{\theta}_i)$. ¿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 $s_i(\cdot) \geq \underline{\eta} = s_j(\cdot)$ 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 2.2. ■

Proof of Proposition 4: We first consider RP 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 RP j. Let $\underline{\gamma}_1 = \min_{\vartheta \in [\tilde{\theta}_i, \theta^u]} \left(s_i'(\vartheta) - s_i(\vartheta) \right)$ and $\bar{\gamma}_2 = \max_{\vartheta \in [\theta^l, \theta^u]} \left(s_j'(\vartheta) - s_j(\vartheta) \right)$. Suppose that $\underline{\gamma}_1 \leq 0$ and pick $\underline{\theta}_1 \in [\tilde{\theta}_i, \theta^u]$ such that $s_i'(\underline{\theta}_1) - s_i(\underline{\theta}_1) = \underline{\gamma}_1$. We first establish that $\bar{\gamma}_2 + \underline{\gamma}_1$ must be positive. If not then,

$$\begin{aligned} 0 \ &\leq \int_{[\theta^l, \theta^u]} \left[f(s_i'(\underline{\theta}_1) + s_j'(\theta_j), \underline{\theta}_1 + \theta_j, k_i') \right] g(\theta_j) d\theta_j \\ &< \int_{[\theta^l, \theta^u]} \left[f(s_i'(\underline{\theta}_1) + s_j'(\theta_j), \underline{\theta}_1 + \theta_j, k_i) \right] g(\theta_j) d\theta_j \\ &\leq \int_{[\theta^l, \theta^u]} \left[f(s_i(\underline{\theta}_1) + \underline{\gamma}_1 + s_2(\theta_j) + \bar{\gamma}_2, \underline{\theta}_1 + \theta_j, k_i) \right] g(\theta_j) d\theta_j \\ &\leq \int_{[\theta^l, \theta^u]} \left[f(s_i(\underline{\theta}_1) + s_2(\theta_j), \underline{\theta}_1 + \theta_j, k_i) \right] g(\theta_j) d\theta_j \end{aligned} \tag{10}$$

The strict inequality follows from Lemma 2.1(a), i.e., f decreases w.r.t. k_i. The first weak inequality holds because $s_i'(\underline{\theta}_1) = s_i(\underline{\theta}_1) + \underline{\gamma}_1$, $s_j'(\cdot) \leq s_j(\cdot) + \bar{\gamma}_2$ and, by Lemma 2.1(b), f increases w.r.t. Σs. The second weak inequality holds because by assumption, $\bar{\gamma}_2 + \underline{\gamma}_1 \leq 0$ and f increases w.r.t. Σs. But inequality (10) is impossible since $\underline{\theta}_1 \geq \tilde{\theta}_i$ implies

$$0 \ = \int_{[\theta^l, \theta^u]} \left[f(s_i(\underline{\theta}_1) + s_2(\theta_j), \underline{\theta}_1 + \theta_j, k_i) \right] g(\theta_j) d\theta_j \tag{11}$$

This establishes that $\underline{\gamma}_1 \leq 0$ implies $\bar{\gamma}_2 + \underline{\gamma}_1 > 0$. Now pick $\bar{\theta}_2$ such that $s_j'(\bar{\theta}_2) - s_j(\bar{\theta}_2) = \bar{\gamma}_2$. Since $\bar{\gamma}_2 > -\underline{\gamma}_1 \geq 0$, it follows that $\bar{\theta}_2 > \tilde{\theta}_j'$. Hence

$$\begin{aligned} 0 \ &= \int_{[\theta^l, \theta^u]} \left[f(s_j'(\bar{\theta}_2) + s_i'(\theta_i), \bar{\theta}_2 + \theta_i, k_j') \right] g(\theta_i) d\theta_i \\ &\geq \int_{[\theta^l, \theta^u]} \left[f(s_j'(\bar{\theta}_2) + s_i'(\theta_i), \bar{\theta}_2 + \theta_i, k_j) \right] g(\theta_i) d\theta_i \\ &\geq \int_{[\theta^l, \theta^u]} \left[f(s_j(\bar{\theta}_2) + \bar{\gamma}_2 + s_i(\theta_i) + \underline{\gamma}_1, \bar{\theta}_2 + \theta_i, k_j) \right] g(\theta_i) d\theta_i \\ &> \int_{[\theta^l, \theta^u]} \left[f(s_j(\bar{\theta}_2) + s_i(\theta_i), \bar{\theta}_2 + \theta_i, k_j) \right] g(\theta_i) d\theta_i \end{aligned} \tag{12}$$

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 $s'_j(\bar{\theta}_2) = s_j(\bar{\theta}_2) + \bar{\gamma}_2$, $s'_i(\cdot) \geq s_i(\cdot) + \underline{\gamma}_1$ and f increases w.r.t. Σs. The strict inequality holds because $\bar{\gamma}_2 + \underline{\gamma}_1 > 0$ and, once again, because f increases w.r.t. Σs. But inequality (12) is impossible because

$$0 \leq \int_{[\theta^l,\theta^u]} \left[f(s_j(\bar{\theta}_2) + s_i(\theta_i), \bar{\theta}_2 + \theta_i, k_j)\right] g(\theta_i) d\theta_i$$

Thus we know $\underline{\gamma}_1 > 0$, or $s'_i(\theta) > s_i(\theta)$ for $\theta \geq \tilde{\theta}_i$, and from continuity of $s(\cdot)$, $\tilde{\theta}'_i < \tilde{\theta}_i$. ■

Proof of Proposition 5: The proof relies on two Lemmas. The first gathers together several results about ϵ-flatness.

Lemma 5.1. *(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 $\boldsymbol{\Theta} = \Theta^2$. For each α, let $\boldsymbol{f}(\cdot,\alpha)$ be a continuous function mapping $\boldsymbol{\Theta}$ to $\mathbb{R}^2_{++}$ and let $\boldsymbol{x}(\cdot,\alpha) = (x_1(\cdot,\alpha), x_2(\cdot,\alpha))$ be an integrable function mapping $\boldsymbol{\Theta}$ to $\mathbb{R}^2$.

Lemma 5.2. *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 $r = 1,2$, $x_r(\cdot,\alpha)$ depends only on θ_r; (b) there exists $\tilde{\boldsymbol{\theta}}(\alpha) \in \boldsymbol{\Theta}$, $\tilde{\boldsymbol{\theta}}(\alpha) \geq \boldsymbol{\theta}^*$, such that $x_r(\cdot,\alpha) = 0$ on $[\theta^l, \tilde{\theta}_r(\alpha))$; and (c) for each $\theta_r \in [\tilde{\theta}_r(\alpha), \theta^u]$, $E[f_r(\boldsymbol{\vartheta},\alpha)\sum_{2=1}^{2} x_2(\boldsymbol{\vartheta},\alpha)|\vartheta_r = \theta_r] \in (-\omega,\omega)$. 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 5. We need to prove that for α sufficiently close to unity,

$$0 \quad \leq \quad \int_{\boldsymbol{\Theta}} \frac{dt(\boldsymbol{\vartheta},\alpha)}{d\alpha}\boldsymbol{g}(\boldsymbol{\vartheta})d\boldsymbol{\vartheta} \quad = \quad \int_{\boldsymbol{\Theta}} \left(t'\left(\Sigma s(\boldsymbol{\vartheta},\alpha)\right)\right) \sum_{r=1}^{2} \frac{d\left[s_r(\vartheta_r,\alpha)\right]}{d\alpha}\boldsymbol{g}(\boldsymbol{\vartheta})d\boldsymbol{\vartheta} \qquad (13)$$

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

$$\int_{\overline{\boldsymbol{\Theta}}(\alpha)} \psi_{1t}(\boldsymbol{\vartheta},\alpha) \left(\sum_{r=1}^{2} \frac{d\left[s_r(\vartheta_r,\alpha)\right]}{d\alpha}\right) g(\boldsymbol{\vartheta})d\boldsymbol{\vartheta} \quad > \quad \frac{3\delta\underline{\psi}_{1t}}{t'(2\underline{\eta})} \qquad (14)$$

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

2. Invoke Lemma 5.2 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 $\gamma < \frac{\delta\underline{\psi}_{1t}}{\Omega \underline{t}'}$ it follows from Step 1 and Part (b) of Lemma 5.1 that

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

4. If C_{tt} is ϵ_{C1}-flat, for $\epsilon_{C1} < (1+\gamma)^{1/3}-1$ and $\frac{t''(2\bar{\eta})}{t'(2\underline{\eta})} < \epsilon_t$, for $\epsilon_t < \min\left\{\frac{(1+\gamma)\min(C_t)}{\max(C_t)}, \frac{(1+\gamma)^{1/3}-1}{2(\bar{\eta}-\underline{\eta})}\right\}$, then $\psi_{1t}(\cdot,\alpha)$ is γ-flat.
5. If t' is ϵ_{C2}-flat, for $\epsilon_{C2} < \frac{\delta}{\Omega t'(2\underline{\eta})}$ it follows from Step 3 and Part (c) of Lemma 5.1 that

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

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 $\int_{\underline{\boldsymbol{\Theta}}(\alpha)(\alpha)} g(\boldsymbol{\vartheta})d\boldsymbol{\vartheta} < \frac{\delta}{t'(2\bar{\eta})}$. thus ensuring that for $\alpha > \bar{\alpha}$, $\int_{\underline{\boldsymbol{\Theta}}(\alpha)} \left(t'\left(\Sigma s(\boldsymbol{\vartheta},\alpha)\right)\right) \sum_{r=1}^{2} \frac{d[s_r(\vartheta_r,\alpha)]}{d\alpha} \boldsymbol{g}(\boldsymbol{\vartheta})d\boldsymbol{\vartheta} > -\delta$. This, together with step 5, 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

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

where

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

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:

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

where

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

and

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

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

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

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

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

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:

$$\begin{aligned} \int_{\overline{\boldsymbol{\Theta}}(\alpha)} \psi_{1t}(\boldsymbol{\vartheta}, \alpha) \left(\sum_{r=1}^{2} \frac{d\left[s_r(\vartheta_r, \alpha)\right]}{d\alpha}\right) g(\boldsymbol{\vartheta}) d\boldsymbol{\vartheta} &= \frac{d\kappa_1(\alpha)}{d\alpha} \int_{\overline{\boldsymbol{\Theta}}(\alpha)} \psi_{rk}(\boldsymbol{\vartheta}, \alpha) g(\boldsymbol{\vartheta}) d\boldsymbol{\vartheta} \\ &\geq -\left(k_1' - k_1\right) \underline{\psi}_{1k} > 0 \end{aligned}$$

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.

Step 2: 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 $\theta_r \in [\tilde{\theta}_r(\alpha), \theta^u]$, $E[f_r(\boldsymbol{\vartheta}) \sum_{2=1}^{2} x_2(\boldsymbol{\vartheta}, \alpha) | \vartheta_r = \theta_r] \in (-\omega, \omega)$. Hence the conclusion of Lemma 5.2 applies.

Step 4: ¿From (17)

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

¿From Part (a) of Lemma 5.1, t'' is $((1+\gamma)^{1/3}-1)$-flat. By assumption, C_{tt} is also. Hence $\frac{\max[C_{tt}(\cdot,\cdot,\cdot)(t'(\cdot))^2]}{\min[C_{tt}(\cdot,\cdot,\cdot)(t'(\cdot))^2]} \leq (1+\gamma)$. Moreover, since $t''(\cdot) < (1+\gamma)\frac{\min(C_t)}{\max(C_t)}$, $\frac{\max[C_t(\cdot,\cdot,\cdot)t''(\cdot)]}{\min[C_t(\cdot,\cdot,\cdot)t''(\cdot)]} \leq (1+\gamma)$ 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 5.1:

(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 $\overline{x} > \underline{x}$.

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

(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 5.2:
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[16]

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

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 $\Delta x^m = \sum_{r=1}^2 \Delta x_r^m$. 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

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

[16] 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

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

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

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

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 $x_1^m(\theta_1^m) \geq E[x_2^m(\vartheta_1)|\vartheta_1 \geq \tilde\theta_1^m]$:

$$\begin{aligned}\lim_{m\to\infty} \frac{E[f_1^m(\boldsymbol{\vartheta})x_2^m(\boldsymbol{\vartheta})|\vartheta_1 \geq \tilde\theta_2^m]}{x_1^m(\theta_1^m)E(f_1^m(\theta_1^m,\vartheta_2))} &\geq \lim_{m\to\infty} \frac{\overline{f}_1 E[x_2^m(\boldsymbol{\vartheta})|\vartheta_1 \geq \tilde\theta_2^m]}{x_1^m(\theta_1^m)E(f_1^m(\theta_1^m,\vartheta_2))} \\ &\geq \lim_{m\to\infty} \frac{\overline{f}_1 E[x_2^m(\boldsymbol{\vartheta})|\vartheta_1 \geq \tilde\theta_2^m]}{\underline{f}_1 E[x_1^m(\boldsymbol{\vartheta})|\vartheta_1 \geq \tilde\theta_1^m]} \qquad (30)\end{aligned}$$

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

$$\lim_{m\to\infty} \frac{E[x_2^m(\vartheta_2)|\vartheta_2 \geq \tilde\theta_2^m]}{E[x_1^m(\vartheta_1)|\vartheta_1 \geq \tilde\theta_1^m]} \leq \frac{-1}{p_2(\tilde{\boldsymbol{\theta}}^m)(1+\epsilon)} \qquad (31)$$

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

$$\lim_{m\to\infty} \frac{E[x_1^m(\vartheta_2)|\vartheta_1 \geq \tilde\theta_1^m]}{E[x_2^m(\vartheta_1)|\vartheta_2 \geq \tilde\theta_2^m]} \leq \frac{-1}{p_1(\tilde{\boldsymbol{\theta}}^m)(1+\epsilon)} \qquad (32)$$

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 6: 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 $i \leq \overline{j}$ and all $\alpha \in [0,1]$, $\frac{d\kappa_i(\cdot)}{d\alpha} \geq 0$, while $\frac{d\kappa_i(\cdot)}{d\alpha} \leq 0$ otherwise. For each $\alpha \in [0,1]$, let $\boldsymbol{s}(\cdot;\alpha)$ be the PSNE corresponding to α. We will prove that $\frac{d}{d\alpha}[E_{\Theta^n}\ (t(\Sigma s(\boldsymbol{\theta},\alpha)))] > 0$. 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 $\kappa_i(\alpha)\frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}}(\tilde\theta_i(\alpha)+E_{\Theta^{n-1}}\Sigma\theta_{-i}) = \underline{\eta}+E_{\Theta^{n-1}}\Sigma s_{-i}(\boldsymbol{\theta}_{-i},\alpha)$. Observe from (8) that for i and $\theta_i > \tilde\theta_i(\alpha)$,

$$0 = \int_{\Theta^{n-1}} \left[s_i(\theta_i,\alpha) + \Sigma s_{-i}(\boldsymbol{\theta}_{-i},\alpha) - \kappa_i(\alpha)\frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}}(\theta_i + \Sigma\theta_{-i})\right]\boldsymbol{g}_{-i}(\boldsymbol{\theta}_{-i})d\boldsymbol{\theta}_{-i}$$

It follows that for all i and all α:

$$\frac{ds_i(\theta_i,\alpha)}{d\alpha} = \begin{cases} 0 & \text{if } \theta_i < \tilde\theta_i(\alpha) \\ \frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}}(\theta_i + E_{\Theta^{n-1}}\Sigma\theta_{-i})\frac{d\kappa_i(\alpha)}{d\alpha} - E_{\Theta^{n-1}}\frac{d\Sigma s_{-i}(\boldsymbol{\theta}_{-i},\alpha)}{d\alpha} & \text{if } \theta_i > \tilde\theta_i(\alpha) \end{cases} \qquad (33)$$

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

$$\frac{d}{d\alpha}[E_\Theta s_i(\cdot,\alpha)] = -s_i(\tilde\theta_i,\alpha)\frac{d\tilde\theta_i(\alpha)}{d\alpha} + \int_{\tilde\theta_i(\alpha)}^{\theta^u}\frac{ds_i(\theta_i,\alpha)}{d\alpha}g(\theta_i)d\theta_i = \int_{\tilde\theta_i(\alpha)}^{\theta^u}\frac{ds_i(\theta_i,\alpha)}{d\alpha}g(\theta_i)d\theta_i. \qquad (34)$$

For $\vartheta \in \Theta$, let $\Psi(\vartheta) = \frac{\gamma_{12}\beta_{11}}{\gamma_{11}\beta_{12}} \int_{\vartheta}^{\theta^u} (\vartheta + E_{\Theta^{n-1}} \Sigma\theta_{-i})\, g(\vartheta) d\vartheta$. Also let $p_i(\alpha) = \int_{\tilde{\theta}_i(\alpha)}^{\theta^u} g(\theta_i) d\theta_i$ denote the probability that $s_i(\cdot, \alpha) > \underline{\eta}$. Substituting (33) into (34), we obtain

$$\begin{aligned} \frac{d}{d\alpha}[E_\Theta s_i(\theta_i,\alpha)] &= \Psi(\tilde{\theta}_i(\alpha)) \frac{d\kappa_i(\alpha)}{d\alpha} - p_i(\alpha)\frac{d}{d\alpha}[E_{\Theta^n}\Sigma s(\boldsymbol{\theta},\alpha)] + p_i(\alpha)\frac{d}{d\alpha}[E_\Theta s_i(\theta_i,\alpha)] \\ &= \frac{\Psi(\tilde{\theta}_i(\alpha))}{1-p_i(\alpha)}\frac{d\kappa_i(\alpha)}{d\alpha} - \frac{p_i(\alpha)}{1-p_i(\alpha)}\frac{d}{d\alpha}[E_{\Theta^n}\Sigma s(\boldsymbol{\theta},\alpha)]. \end{aligned} \tag{35}$$

Now by assumption, there exists $\overline{j} \in \{1,...,n-1\}$ such that $\frac{d\kappa_i(\alpha)}{d\alpha} \geq 0$ iff $i \leq \overline{j}$ and $\frac{d\kappa_i(\alpha)}{d\alpha} \leq 0$ otherwise. Moreover, since $i \leq \overline{j}$ implies $k_i(\alpha) \geq k_{\overline{j}}(\alpha)$, it follows from Proposition 3 that $\tilde{\theta}_i(\alpha) \leq \tilde{\theta}_{\overline{j}}(\alpha)$ and $p_i(\alpha) \geq p_{\overline{j}}(\alpha)$ if $i \leq \overline{j}$, while $\tilde{\theta}_i(\alpha) \geq \tilde{\theta}_{\overline{j}}(\alpha)$ and $p_i(\alpha) \leq p_{\overline{j}}(\alpha)$ otherwise, with equality holding only if $\kappa_i(\alpha) = \kappa_{\overline{j}}(\alpha)$. It follows that $\frac{\Psi(\tilde{\theta}_i(\alpha))}{1-p_i(\alpha)} \geq \frac{\Psi(\tilde{\theta}_{\overline{j}}(\alpha))}{1-p_{\overline{j}}(\alpha)}$ if $i \leq \overline{j}$, while $\frac{\Psi(\tilde{\theta}_i(\alpha))}{1-p_i(\alpha)} \leq \frac{\Psi(\tilde{\theta}_{\overline{j}}(\alpha))}{1-p_{\overline{j}}(\alpha)}$, otherwise. Moreover, by assumption not all of the k_i's are equal, so that for at least one i, one of the above inequalities holds strictly. These relationships imply that for all i,

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

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

$$\begin{aligned} \frac{d}{d\alpha}[E_{\Theta^n}\Sigma s(\boldsymbol{\theta},\alpha)] &> \frac{\Psi(\tilde{\theta}_{\overline{j}}(\alpha))}{1-p_{\overline{j}}(\alpha)}\sum_{i=1}^{n}\frac{d\kappa_i(\alpha)}{d\alpha} - \sum_{i=1}^{n}\frac{p_i(\alpha)}{1-p_i(\alpha)}\frac{d}{d\alpha}[E_{\Theta^n}\Sigma s(\boldsymbol{\theta},\alpha)] \\ &= \left(1+\sum_{i=1}^{n}\frac{p_i(\alpha)}{1-p_i(\alpha)}\right)^{-1}\frac{\Psi(\tilde{\theta}_{\overline{j}}(\alpha))}{1-p_{\overline{j}}(\alpha)}\sum_{i=1}^{n}\frac{d\kappa_i(\alpha)}{d\alpha} \end{aligned} \tag{36}$$

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

Proof of Proposition 7: Let $\boldsymbol{k}'$ denote the burden profile after the buyout. That is, $k_i' = k_i\left(\sum_{\{i:k_i>\underline{k}\}} k_i\right)^{-1}$ if $k_i > \underline{k}$ and 0 otherwise. Clearly, the conditions of Proposition 6 are satisfied, except for the fact that each remaining RP's objective function (see expression (7)) is decremented by a constant number (the sum of *de minimis* RPs' buyout payments). Specifically, assume that burden shares are ordered as usual so that $i < j$ implies $k_i \geq k_j$. Let $\overline{j} = \min\{j \in \{1,...,n\} : k_j > \underline{k}\}$ and observe that $k_i' > k_i$ if $i \geq \overline{j}$ while $k_i' = 0$ 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 6. ■