PUBLIC CONTRACTING IN DELEGATED AGENCY GAMES\textsuperscript{1}

January 16, 2013

DAVID MARTIMORT AND LARS STOLE

We study games of public delegated common agency under asymmetric information. Using tools from non-smooth analysis and optimal control, we derive best responses and characterize equilibria (both continuous and discontinuous) using self-generating optimization programs of which any equilibrium allocation must be a solution. Special attention is given to common agency games in which each principal’s payoff is a linear function of the agent’s action. In such games the self-generating optimization program reduces to the maximization of the principals’ “aggregate” virtual surplus in which the agent’s marginal valuation is replaced by a confluence of “virtual” valuations that reflect common agency problems.

In all equilibria, we illustrate that there are two distinct sources of inefficiencies: inefficient contracting by a given coalition of active principals and inefficient activity by principals. One noteworthy subset of equilibrium allocations are maximal in the range of actions that arise. These allocations are straightforward to compute and are supported by continuously differentiable transfer functions. Furthermore, starting from a maximal allocation, it is possible to introduce discontinuities that are supported in a corresponding non-smooth, discontinuous equilibrium that has a smaller range of allocations (i.e., not maximal).

Our results are illustrated by means of two games: a public goods game in which each player simultaneously offers a menu contract to a common provider of the public good in order to induce greater supply, and a lobbying game between conflicting interest groups in which each group offers a menu of contributions to a common political decision-maker in an attempt to influence policymaking.

**KEYWORDS:** Common agency, asymmetric information, menu auctions, delegated contracting games, public goods, lobbying.

\textsuperscript{1}We thank seminar participants in Zürich, Toulouse, the 5th CSEF-IGIER Symposium in Economic Theory in Capri, ESEM-Barcelona 2009, University of British Columbia, Bocconi University of Milan, Princeton, University of Pennsylvania, Paris School of Economics, Bologne, University College of London, Ohio State University, Montpellier, the CEPR Applied Industrial Organization 2010, the Cowles Foundation Economic Theory Summer 2011 Conference, the Australasian Meeting of the Econometric Society 2012 in Melbourne, University British Columbia, London School of Economics, Mannheim and Munich for their helpful remarks. We are especially thankful to David Rahman for comments and to John Birge, Philippe Jehiel, Stephano Lovo and Aggey Semenov for many helpful discussions. Some results and applications have been taken from our previous paper, “Common Agency Games with Common Screening Devices,” (2005).

\textsuperscript{a}Paris School of Economics-EHESS, david.martimort@parisschoolofeconomics.eu

\textsuperscript{b}University of Chicago Booth School of Business, lars.stole@chicagobooth.edu
1. INTRODUCTION

This paper considers a large class of multi-principal, common-agent games defined by three properties: (i) the agent has private information at the time of contracting, (ii) the agent’s actions are publicly observable and contractable by any principal (i.e., contracting variables are public), and (iii) the agent is free to accept any subset of contract offers from the principals (i.e., the degree of common agency is endogenous and delegated to the agent). To be concrete, we will explore two motivating examples – public goods games and influence games – which differ in terms of the conflict between the principals. In the public goods example, the principals are a set of citizens, each of who may offer a contract to a common, privately-informed producer, which is a promise to make a payment to the agent as a function of the total public good produced. Because all citizens prefer more public good to less in these games, we say that the principals’ preferences are congruent, even if there are be differences in the willingness to pay for such goods. In influence games, we take the set of principals to be lobbyists who wish to influence the actions of a privately-informed decision maker (e.g., politician). A contract in this context is a credible promise by lobbyists to make contributions to the politician as a function of the latter’s publicly observed vote. In influence games such as these, it is natural to suppose that the principals may have opposite rankings over a decision maker’s choice. If this is the case, the principals’ preferences are not congruent, but are opposed. As we will see, whether the principals’ preferences are congruent (e.g., preferences for public goods) or in opposition (e.g., lobbyists on the left and right of a policy issue) will determine the character of the equilibrium distortions.

To obtain a characterization of the equilibria for this class of games, there are four conceptual and technical difficulties that must be overcome. First, even if the agent’s non-participation option yields a type-independent payoff, the agent’s participation constraint vis-à-vis principal \( i \) will typically be type dependent if the agent’s best alternative is to contract with a subset of the remaining principals. Moreover, the value of the partial-contracting option is an endogenous object. Second, the set of principals who are actively influencing a particular agent type (i.e., making positive payments to that agent type in equilibrium) is also an endogenous object that must be constructed as part of the equilibrium. As we will see, it is not uncommon for each principal’s activity set (i.e., the set of types to whom positive transfers are made) to be a proper subset of the agent’s type space and to differ fundamentally from each other. This subtlety is novel to delegated common agency games with adverse selection. Third, the appropriate strategy spaces in these games are large – in the present case, the set of upper semi-continuous functions from actions to transfers. Unfortunately, standard control-theoretic results require that the principals’ objective functions are continuous and, typically, also differentiable. Because principal \( i \)'s objective function depends upon the equilibrium transfer functions of the other principals, this smoothness restriction would require each principal to choose a differentiable transfer function. Even assuming that a continuously differentiable contract is a best response for a principal facing rivals who offer similarly-smooth tariffs, using standard control techniques in this context is implicitly making an equilibrium refinement. Such a refinement should be made explicit and its implications thoroughly understood. To this end we import and specialize the modern techniques of non-smooth analysis in optimal control allowing for discontinuous transfers. Fourth, most games in this class have a multiplicity of equilibria, so we would like theorems broadly applicable to the entire equilibrium set or, if this is not possible, theorems applicable to an important subset of equilibria where the equilibrium refinement has been appropriately applied.
This paper addresses each of these four difficulties and contributes in several dimensions. First, to address type-dependent optimization problems with non-smooth objective functions, we derive a novel set of necessary and sufficient conditions for the solution to control problems with objective functions that are linear in the state variable—an assumption satisfied in most screening contracting problems in which payoffs are quasi-linear in money. The general statement of this result is presented and proven in Appendix B.\(^1\) Theorem 1 specializes this non-smooth control-theory result to provide necessary and sufficient conditions for a principal’s best-response correspondence in common-agency games.

Using the non-smooth control theorem, we are able to establish broad necessary conditions for any equilibrium outcome. The key ingredient is to note that our class of common agency games is a special case of an aggregate game with infinite-dimensional strategy spaces. Using the principle of aggregate concurrence, a result from the aggregate games literature (Martimort and Stole (2012)), we are able to derive illuminating necessary conditions which apply to the set of all equilibrium allocations (i.e., all equilibrium actions and agent-rent profiles). Indeed, given the activity sets of the principals, these necessary conditions provide sharp predictions. This is the main contribution of Theorem 2.

The endogeneity of the principals’ activity sets introduces an additional set of technical issues, largely driven by potential nonlinearities in the principals’ virtual marginal valuations. To mitigate these difficulties while addressing the endogeneity of the activity sets, we place additional structure on the payoffs of players—in short, we require that each principal’s marginal return to the agent’s action is constant. Within this class of linear common-agency games, we are able to prove equilibrium existence and characterize the distortions caused by non-cooperative contracting in a broad sense. This class of equilibria includes discontinuous equilibria as well as a subclass of equilibria which we call maximal. For this subclass, we provide a complete characterization of equilibria. We find that these equilibria share interesting similarities with the Lindahl-Samuelson allocations from public finance. These equilibrium outcomes are extremely simple to calculate and comparative statics are immediately available. Moreover, as private information becomes less dispersed the limiting equilibrium allocation is the truthful equilibrium allocation of Bernheim and Whinston (1986a). In this sense, these maximal allocations can be thought of as generalizations to the complete-information concept of truthful equilibrium. We apply these results to our motivating examples of public goods games and lobbying games to explore the role of congruence and conflict in the principals’ preferences.

\textit{Literature review}

Broadly speaking, the various models of common agency contracting games in the literature have a similar strategic setting and timing. In the first stage, each principal non-cooperatively and simultaneously offers a common agent a contract. In the second stage, after receiving each principal’s contract offer, the agent makes an acceptance decision and then chooses optimal action(s) given the set of active contracts. This class of games, however, can differ on two important dimensions. The first is whether the contracting variables are publicly available to all principals (i.e., public contracting) or instead each principal has an exclusive domain of contracting (private contracting). A second dimension is whether the agent is free to accept any subset of the principals’ contractual offers (delegated agency) or instead the agent must instead choose to contract with all or

\(^1\)In its general form, this result may be of interest outside the context of common agency games. In a nutshell, it provides a non-smooth generalization of Jullien (2000).
none (intrinsic common agency). The current paper sits squarely in the domain of public contracting games with delegated agency.

There is an existing literature on common agency games with public contracting variables. Under complete information, Bernheim and Whinston (1986a) initiated the study of such games (referred to as “menu auctions” in that paper) and their results about the existence and properties of truthful equilibria have been successfully applied to many contexts, including international trade and lobbying. There is also a less developed literature that has looked at intrinsic common agency games with public contracting variables. In the context of moral hazard, Bernheim and Whinston (1986b) provided the first application in this framework. Later papers by Holmström and Milgrom (1988) and Dixit (1996) extended these insights to moral hazard games with linear contracts; more recently, Martimort and Stole (2012) extend the model of Innes (1990) to a common agency framework, proving equilibrium existence and characterizing the competitive distortions. Under adverse selection, models of intrinsic agency with public variables have been developed in specific applications to regulatory economics (Laffont and Tirole (1993), Olsen and Torsvik (1993), Martimort (1996)) or politics (Martimort and Semenov (2008)). Although, we share with this literature a focus on how distortions are compounded under non-cooperative contracting, delegated agency opens significant new difficulties related to how binding type-dependent participation constraints affect how distortions are compounded.

The literature on common agency games with adverse selection has largely been focused on games of private contracting variables. In a private contracting setting – i.e., a game in which each principal can contract on an exclusive domain of agent actions – Martimort and Stole (2009a) provide a general analysis of competition with nonlinear prices under both delegated and intrinsic common agency with endogenous activity sets. This is the appropriate class of models for games in which each principal controls a specific screening variable (e.g., the quantity he sells to the buyer) but does not observe the agent’s actions on other dimensions (e.g., the quantity purchased from rivals). The focus in the present paper, however, is public agency environments where each principal has access to the same screening variables. A second difference is that, in Martimort and Stole (2009a), manufacturers rank the agent’s types the same way, with the agent having the highest valuation for both goods being the most attractive for both manufacturers. Similarly, Biais, Martimort and Rochet (2000) analyze a model of competing market-makers on financial markets with traders privately informed on their willingness to buy or sell assets in a common value environment with private agency; because of symmetry, all market-makers have similar activity sets with a bid-ask spread such that traders having a mild preference for trading do not trade under asymmetric information. The present analysis is more general, allows for principals having conflicting preferences such as in our lobbying game and explores the implications of opposed preferences for the equilibrium activity sets of the principals.

---

3 A notable exception is Martimort (1996). Martimort (2007) introduces the distinction of private and public common agency and previews the analysis in the present paper.
4 Iviali and Martimort (1994) and Calzolaric and Scarpa (2008) are earlier studies of delegated common agency games with private contracting but focus a priori on cases where all types are served.
5 Mezetti (1997) provide a model with conflicting and differentiated principals but his focus was on an intrinsic common agency setting, putting aside the complete characterization of the activity sets.
Organization

Section 2 presents our model of delegated common agency under asymmetric information and some preliminary results characterizing each principal's set of incentive feasible allocations. Having dispensed with the preliminaries, in Section 3 we proceed to first take a detour and ask what would a continuous, smooth equilibrium look like, assuming that one exists. Here we develop the rough intuition for our equilibrium analysis. In particular, we derive a simple optimization program and argue that its solution set includes the smooth, continuous equilibrium allocations. This partial analysis, of course, leaves many unanswered questions: Do such smooth equilibria exist? Do other equilibria exist? Can we say something about the set of all equilibrium allocations? In the remainder of the paper we are able to answer each question in the affirmative. To this end, Section 4 constructs each principal's best response as a solution to a control problem using non-smooth analysis which can be used to understand best-responses to discontinuous tariffs. Necessary and sufficient conditions for equilibria are derived from these optimality conditions in Section 5. Section 6 specializes to the case in which the principals' preferences are linear in the agent's action. This allows us to obtain closed form solutions for the activity sets and to derive finer properties of equilibria. In Section 7, we apply our findings to public good provision and lobbying games to illustrate the nature of equilibrium distortions. In the former, we obtain an elegant recasting of the Lindahl-Samuelson condition for public goods. In the latter, we derive a simple condition characterizing which principals will exert exclusive influence over an agent type and whether or not the principals' spheres of influence will overlap. Section 8 concludes and highlights several avenues for further research. The proofs of the main results are collected in Appendix A. Appendix B contains a detailed presentation of necessary and sufficient conditions for non-smooth optimal control problems in which the objective is linear in the state variable.

2. THE MODEL

2.1. Public delegated agency games

We consider a common-agency game with n principals, indexed with the subscript \( i \in \{1, ..., n\} = N \). In the first stage of the game, each principal offers a contribution schedule, \( t_i \in \mathcal{T} \), (e.g., transfer, tariff, etc.) to a common agent. We take the set of such contracts to be the set of upper-semicontinuous functions that assign a monetary payment to the agent for each action, \( q \in \mathcal{Q} \equiv [q_{\min}, q_{\max}] \subset \mathbb{R}_+ \). In the second stage of the game, the agent may decide which subset of the principals' offers to accept and, conditional on these acceptances, the agent decides which action to take.

Given we are only interested in the equilibrium choices of the agent, \( \theta \mapsto q \), and the payoffs of the players, it is without loss of generality to restrict attention to equilibria in which principals choose nonnegative transfer functions and the agent excepts every offer. To understand this observation, consider an equilibrium in which some principal \( j \) offers a transfer function, \( \bar{t}_j \), such that \( \bar{t}_j(\hat{q}) < 0 \) for some \( \hat{q} \). If \( \hat{q} \) is chosen in equilibrium by some agent type, that agent must refuse principal \( j \)'s contract in the original equilibrium. Thus, there is no loss to principal \( j \) to offering instead \( \tilde{t}_j(\hat{q}) = 0 \). Indeed, it is optimal for principal

---

6Importantly, principal \( i \)'s contract cannot be a function of the details of principal \( j \)'s contract, or a function of the agent's acceptance decisions with respect to the other principals. “Contracts on contracts” are not allowed because, for instance, principals \( j \)'s offers \( (j \neq i) \) are non-observable by principal \( i \). On the impact of such self-referencing contracts, see Peters and Szentes (2012). Within our framework, and following the Delegation Principle in Peters (2001) and Martinot and Stole (2002), there is no loss of generality in considering nonlinear prices to describe equilibria of the common agency game.
\( j \) to offer \( \tilde{r}_j(q) = \max\{0, r_j(q)\} \) in place of the original offer \( r_j \). The agent’s choice mapping \( \theta \mapsto q \) is unaffected and payoffs are as before. We therefore proceed by restricting attention to games in which each principal must offer a nonnegative contribution schedule.\(^7\)

All players have quasi-linear preferences defined over outputs and monetary payments. We denote the agent’s direct benefit from choosing \( q \in \mathcal{Q} \) as

\[
S_0(q) = \theta q,
\]

where \( S_0: \mathcal{Q} \to \mathbb{R} \) is an upper-semicontinuous function and \( \theta \in \Theta \equiv \{\theta, \overline{\theta}\} \) is a privately-known preference parameter. Similarly, we denote principal \( i \)’s direct benefit from the choice of \( q \in \mathcal{Q} \) as \( S_i(q) \), where again \( S_i: \mathcal{Q} \to \mathbb{R} \). We assume these benefit functions are upper-semicontinuous and also strictly monotonic. We define the subset of principals \( \mathcal{A} \subseteq \mathcal{N} \) as those with increasing \( S_i \) benefits, and the subset of principals \( \mathcal{B} \) as those with decreasing \( S_i \) benefit functions.

The agent’s private preference information, \( \theta \in \Theta \), is referred to as the agent’s “type” which is distributed by the absolutely continuous cumulative distribution function \( F: \Theta \to [0, 1] \), with corresponding bounded density function, \( f: \Theta \to \mathbb{R}_{++} \). Throughout we make the standard assumption in the screening literature that the distribution of types satisfies a “generalized monotone hazard rate” property; i.e.,\(^8\)

\[
\frac{F(\theta) - \kappa}{f(\theta)} \text{ is strictly increasing over } \Theta \text{ for any } \kappa \in [0, 1].
\]

The delegated common agency game unfolds as follows. The agent learns his private information \( \theta \in \Theta \). The principals then non-cooperatively offer contracts to the agent. Because we have assumed without loss of generality that the principals make nonnegative offers, we may suppress the agent’s acceptance decision. Given the set of nonnegative payment offers, the agent chooses \( q \in \mathcal{Q} \) to maximize his utility. We study the set of pure-strategy Bayesian-Nash equilibria to this game. We will use \( t = (t_1, \ldots, t_n) \) to denote an arbitrary profile of transfers and \( t_{-i} \) to be a transfer profile absent \( t_i \)’s contribution schedule. To economize on notation, we denote \( T(q) \equiv \sum_{i \in \mathcal{N}} t_i(q) \) and \( T_{-i}(q) \equiv \sum_{j \neq i} t_j(q) \) as the aggregate tariffs associated with some profile. When referencing particular equilibrium components, we use an underline “\( \_ \)” to distinguish the variable from an arbitrary transfer function or agent action.

**Definition 1**  
A perfect Bayesian equilibrium is a profile of principals’ contribution schedules, \( \tilde{t} = \{\tilde{t}_1, \ldots, \tilde{t}_n\} \), and an agent’s output strategy, \( \overline{q}_0(\theta|t) \), such that the following properties hold.

1. Given any profile of contributions \( t \in T^n \), \( \overline{q}_0(\theta|t) \) maximizes the agent’s payoff:

\[
\overline{q}_0(\theta|t) \in \arg\max_{q \in \mathcal{Q}} S_0(q) - \theta q + \sum_{i \in \mathcal{N}} t_i(q).
\]

\(^7\)Note that this nonnegativity restriction would be with loss of generality for games of intrinsic common agency in which some agent types might accept negative contribution schedules from a subset of principals in order to participate in the game and simultaneously accept positive payments from other principals. For intrinsic common-agency games, the appropriate assumption is that transfer functions are unrestricted. Without a nonnegativity restriction, characterizing the set of equilibria is decidedly simpler. See Martimort and Stole (2012).

\(^8\)See Bagnoli and Berström (2005) and Julien (2000) among others. This condition always holds when both \( F(\theta) \) and \( 1 - F(\theta) \) are log-concave.
2. \( \bar{t}_i \) maximizes principal \( i \)'s expected payoff given \( \bar{t}_{-i} \) and \( \bar{q}_0 \): 
\[
\bar{t}_i \in \arg \max_{t_i \in T} \int_{\Theta} (S_i(q_0(\theta | t_i, \bar{t}_{-i})) - t_i(q_0(\theta | t_i, \bar{t}_{-i}))) f(\theta) d\theta.
\]

In the arguments that follow, it will be helpful to distinguish between those principals who are contributing positive surplus to a given type of agent from those who are "inactive". To this end, for any profile of transfers \( t \), we define an agent's outside option relative to principal \( i \in N \) as follows:
\[
U_{-i}(\theta) = \max_{q \in \mathcal{Q}} S_0(q) + \sum_{j \neq i} t_j(q).
\]

When considering equilibrium tariffs, we will denote this reservation utility as \( \bar{U}_{-i}(\theta) \). Correspondingly, we denote the type-\( \theta \) agent's optimal choice given equilibrium tariffs \( \bar{t}_{-i} \) and absent any offer from principal \( i \), as \( \bar{q}_{-i}(\theta) = \bar{q}_0(\theta | \Theta_0, \bar{t}_{-i}) \). We can now define the equilibrium activity sets as follows:

**Definition 2** Principal \( i \)'s **equilibrium activity set** is defined as
\[
\bar{\Theta}_i \equiv \{ \theta \in \Theta \mid U(\theta) > \bar{U}_{-i}(\theta) \}.
\]

The set of active principals for \( \theta \) is given by the **equilibrium activity map**
\[
\bar{\alpha}(\theta) \equiv \{ i \in N \mid U(\theta) > \bar{U}_{-i}(\theta) \}.
\]

Note that \( \bar{\Theta}_i \) is open in \( \Theta \) and principal \( i \)'s contribution is necessarily positive on his activity set. We define the complement of \( \bar{\Theta}_i \) as \( \bar{\Theta}_i^c \), i.e., the subset of types where the participation constraint (2.1) is binding. The activity mapping \( \bar{\alpha} \) contains a complete description of all equilibrium activity sets; in particular, \( |\bar{\alpha}(\theta)| \) denotes the number of active principals at a given \( \theta \).

With the notation \( \bar{U}_{-i} \) for an agent's outside option, relative to principal \( i \), we can define and characterize the set of incentive feasible allocations available to principal \( i \) in a familiar manner.

**Definition 3** A rent-output profile \((U, q), \Theta : \Theta \to \mathbb{R}_+ \) and \( q : \Theta \to \mathcal{Q} \), is **implementable by principal \( i \)** for a given \( T_{-i} \) if and only if there exists a contribution schedule \( t_i \in T \) such that for all \( \theta \in \Theta 
\[
U(\theta) = \max_{q \in \mathcal{Q}} S_0(q) - \theta q + t_i(q) + T_{-i}(q),
\]
\[
q(\theta) \in \arg \max_{q \in \mathcal{Q}} S_0(q) - \theta q + t_i(q) + T_{-i}(q),
\]
\[
U(\theta) \geq \bar{U}_{-i}(\theta).
\]

The set of allocations which principal \( i \) can implement, given \( T_{-i} \), can also be characterized in the dual space of indirect utilities:

**Lemma 1** A rent-output profile \((U, q) \) is implementable by principal \( i \) given an aggregate transfer profile, \( T_{-i} \), if and only if
\[
U(\theta) \geq \bar{U}_{-i}(\theta) \text{ for all } \theta \in \Theta,
\]
\begin{equation}
- q(\theta) \in \partial U(\theta) \text{ for all } \theta \in \Theta,
\end{equation}

\begin{equation}
U \text{ is convex.}
\end{equation}

Here, we have used the notation \( \partial \) in (2.2) as the subdifferential operator of a convex function. Because \( U \) is convex, it is differentiable almost everywhere; at all such smooth points the subdifferential is simply the gradient of \( U \) and we have \( U(\theta) = -q(\theta) \).\footnote{Condition (2.2) holds also for the equilibrium rent/output profile \( (\overline{U_{-i}}(\theta), \overline{q_{-i}}(\theta)) \) since it is itself implementable (when principal \( i \) offers a null contract) and for the status quo profile \( (U_0(\theta), q_0(\theta)) \) which arises if every principal offers the null contract. This property is called “homogeneity” in Jullien (2000).}

3. A HEURISTIC APPROACH

Rather than proceeding immediately to a characterization of the diverse set of equilibria to the common-agency game, it is useful to consider a more limited inquiry. Supposing that an equilibrium exists with the properties that each \( \overline{u}_i \) is twice-continuously differentiable and the endogenous equilibrium objectives \( S_0 + \overline{T}_{-i} \) are concave functions of \( q \), what can be said about the other properties of this equilibrium? We will see the results from this line of inquiry reappear in the more general analysis, with some additional modifications.

To be concrete, suppose that we restrict attention to sufficiently smooth equilibria, \( \overline{u}_i \in C^2(Q) \) (the set of twice continuously differentiable functions on \( Q \)), and we have reason to believe that they exist, and suppose further that for each principal \( i \), the following equilibrium residual payoff is concave in \( q \):

\[ S_0(q) - \theta q + \overline{T}_{-i}(q). \]

Then we can approach the program of principal \( i \) using standard techniques of optimal control with smooth objectives. The program to be solved is

\[ \max_{q, U} \int_{Q} \left( S_i(q(\theta)) + S_0(q(\theta)) - \theta q(\theta) + \overline{T}_{-i}(q(\theta)) - U(\theta) \right) dF(\theta), \]

subject to \( q(\theta) = -\overline{U}(\theta) \) at all points of differentiability, \( q \) nonincreasing (equivalently, \( U \) convex), and \( U(\theta) \geq \overline{U}_{-i}(\theta) \) where

\[ \overline{U}_{-i}(\theta) \equiv \max_{q \in Q} S_0(q) - \theta q + \overline{T}_{-i}(q). \]

This is, of course, a non-trivial program to solve given that the agent’s outside option with respect to principal \( i \) is type dependent and an equilibrium object.

Consider the simplest case in which \( S_i(q) = s_i q \), i.e., each principal’s preferences are linear in \( q \). This assumption, and the maintained assumptions that the objective is concave and twice continuously differentiable, afford us the use of a result from Jullien (2000, Theorem 1). A careful application of this theorem allows us to conclude that an optimal solution exists, the optimal allocation \( \overline{q} \) is continuous, and furthermore, for each \( \theta \in \Theta \), either

\[ \overline{q}(\theta) = \overline{q}_{-i}(\theta) = \arg \max_{q \in Q} S_0(q) - \theta q + \overline{T}_{-i}(q), \]

\[ U(-\overline{U}(\theta), \overline{q}_{-i}(\theta)) \]
or
\begin{equation}
\bar{q}(\theta) = \arg \max_{q \in \mathcal{Q}} S_i(q) + S_0(q) + T_{-i}(q) - \left( \theta + \frac{F(\theta)}{f(\theta)} - \gamma_i(\theta) \right) q,
\end{equation}

where \( \gamma_i : \Theta \to [0,1] \) is the adjoint multiplier associated with the outside option. The determination of the adjoint function, \( \gamma_i(\theta) \), is not obvious, but looking ahead to this paper’s result in Theorem 3, we are encouraged to guess that \( \gamma_i = 0 \) if \( i \in \mathcal{A} \) and \( \gamma_i = 1 \) if \( i \in \mathcal{B} \). This indeed results in a solution that satisfies complementary slackness and the other conditions required in Jullien’s (2000) result, allowing us to conclude that we have indeed found an optimal solution.

Consider, for example, the case in which \( i \in \mathcal{A} \), so principal \( i \) values higher activity, \( q_i \). In this case, for all types for which \( i \) is active (i.e., \( U(\theta) > U_{-i}(\theta) \)), the optimal allocation solves
\[
s_i + S'_0(\overline{q}(\theta)) + T'_{-i}(\overline{q}(\theta)) = \theta + \frac{F(\theta)}{f(\theta)}.
\]
When principal \( i \) is inactive, the characterization of the equilibrium output by means of first-order conditions for the agent’s problem yields
\[
S'_0(\overline{q}(\theta)) + T'_{-i}(\overline{q}(\theta)) = \theta.
\]
Putting together these observations, it follows that \( i \) is active and increases the agent’s output beyond what the agent would choose under the sole influence of all other principals precisely when
\[
s_i > \frac{F(\theta)}{f(\theta)}.
\]
On a boundary of \( i \)'s activity set, this condition becomes an equality; this is essentially a "smooth-pasting" requirement. Both total and marginal contributions are zero at that point.

We may thus characterize principal \( i \)'s optimal allocation in a more illuminating form as the solution to a novel optimization problem. Specifically, the results above, in tandem with the fact that the optimal allocation must be continuous, are equivalent to the requirement that for all \( \theta \), the optimal \( \overline{q} \) satisfies
\begin{equation}
\overline{q}(\theta) \in \arg \max_{q \in \mathcal{Q}} S_0(q) + T_{-i}(q) + (\beta_i(\theta) - \theta) q,
\end{equation}
where we have introduced the notation
\begin{equation}
\beta_i(\theta) = \begin{cases} 
\max \left\{ s_i - \frac{F(\theta)}{f(\theta)}, 0 \right\} & i \in \mathcal{A}, \\
\min \left\{ s_i + \frac{1 - F(\theta)}{f(\theta)}, 0 \right\} & i \in \mathcal{B}, 
\end{cases}
\end{equation}
to capture principal \( i \)'s virtual marginal return to the activity \( q_i \). To understand the above result, notice that when \( i \in \mathcal{A} \) and \( \beta_i = 0 \), principal \( i \) is inactive and the solution to (3.2) is \( \overline{q}_{-i} \). When \( \beta_i > 0 \), on the other hand, principal \( i \) is active and the solution to (3.2) is the solution to (3.1); here we used the concavity of the agent’s equilibrium program. The allocation which solves (3.2) is continuous, as required, and therefore we conclude at this stage that this is the optimal allocation for principal \( i \) given that twice continuously differentiable transfers emerge in equilibrium and that \( S_0 + T_{-i} \), is concave.
The transfers that principal $i$ would use to implement $\bar{q}$, given $T_{-i}$, are given by a simple formula that makes use of $\bar{\theta}(q)$ as the inverse of $\bar{q}(\theta)$.\footnote{Formally, the inverse is a correspondence, but because $\bar{q}$ is nonincreasing, the inverse is monotonic and almost everywhere single valued. Hence, with some abuse of notation, we may introduce the correspondence $\bar{\theta}(q)$ into an integrand.} Suppose that there exists a type, $\theta_i \in \text{int } \Theta$ such that

$$s_i = \frac{F(\hat{\theta}_i)}{f(\theta_i)}.$$

Then principal $i$ may implement $\bar{q}$ using the transfer function

$$\bar{t}_i(q) = \int_{\bar{\theta}_i}^{q} \max \left\{ s_i - \frac{F(\bar{\theta}(x))}{f(\bar{\theta}(x))}, 0 \right\} dx.$$

It is worth noting at this stage that the generalized monotone hazard rate property, together with the fact that $\bar{q}$ is nonincreasing, implies that $\bar{t}_i$ is convex.

Consider the equilibrium implications of (3.2). We may now use the principle of aggregate congruence, developed in Martimort and Stole (2012), to draw conclusions about the equilibrium aggregate. In particular, because each principal’s choice must satisfy (3.2), we may form an aggregate of these programs across the principals by summation to obtain the condition that

$$(3.4) \quad \bar{q}(\theta) \in \arg \max_{q \in \mathcal{Q}} S_0(q) + \left( \left( \sum_{i \in N} \beta_i(\theta) \right) - \theta \right) q + (n - 1) \left( S_0(q) - \theta q + T(q) \right).$$

Because this program is assumed to be smooth and concave (an equilibrium assumption), incentive compatibility implies that the last term can be ignored. Using the notation $\beta(\theta) = \sum_{i \in N} \beta_i(\theta)$ to denote the aggregate virtual margin, we conclude that \textit{if a smooth, concave equilibrium exists}, then the equilibrium allocation must satisfy

$$(3.5) \quad \bar{q}(\theta) \in \arg \max_{q \in \mathcal{Q}} S_0(q) + (\beta(\theta) - \theta) q.$$ 

Equation (3.5) is a remarkable condition which we will discuss in detail below. For now, however, we should entertain the possibility that it is vacuous. We do not know whether a twice-differentiable, continuous equilibrium with concave objective functions exists. Under general conditions, the tariffs that implement the conjectured equilibrium allocation are not everywhere twice differentiable, although we will see that they are continuous and differentiable over the range of equilibrium actions, $\bar{q}(\Theta) \subset \mathcal{Q}$, and that a generalization of Jullien’s (2000) result is still available. The conjectured tariffs are also necessarily convex, so $\bar{T}_{-i}$ is a convex function. This leaves open the possibility that $S_0 + \bar{T}_{-i}$ is not concave as required. We will demonstrate that with some additional arguments, one can establish the requisite concavity. Theorem 3, in part, establishes that the allocation described by (3.5) is an equilibrium allocation. We should also consider that there may be other equilibria that should be entertained. As we will see, in the delegated common agency game studied in this paper, the allocation which is characterized by (3.5) is only one of an infinite number of continuous and discontinuous equilibrium allocations. Fortunately, Theorem 3 also establishes that every equilibrium allocation satisfies a weaker version of (3.5); this similarity allows us to make broader economic statements about the nature of distortions in \textit{any} equilibrium. Indeed, Theorem 4 shows how to construct discontinuous equilibrium allocations that would not be predicted by this heuristic approach.
4. OPTIMALITY CONDITIONS FOR BEST RESPONSES

We return now to a general analysis of the game without any a priori restrictions on the nature of the equilibrium. We begin with an analysis of principal $i$'s best response to $t_{-i}$. Principal $i$ looks for an allocation that maximizes the joint payoff of the bilateral coalition he forms with the agent,

$$S_i(q) + S_0(q) + T_{-i}(q) - \theta q,$$

minus the information rent, $U(\theta)$, that the type-$\theta$ agent retains. Of course, the ability of principal $i$ to extract this rent is limited by the agent’s option to contract with other principals. Formally, for any principal $i$ and any arbitrary aggregate transfer offered by the rival principals, $T_{-i}$, principal $i$ best response can be cast as the solution to an optimal-control problem with a possibly discontinuous objective function:

$$\max \; \int_{\Theta} \left( S_i(q(\theta)) + S_0(q(\theta)) + T_{-i}(q(\theta)) - \theta q(\theta) - U(\theta) \right) f(\theta) \, d\theta$$

subject to $(2.1)$, $(2.2)$, $(2.3)$,

where $AC(\Theta, \mathbb{R})$ denotes the set of all absolutely continuous, real functions with domain $\Theta$ and $\mathcal{L}(\Theta, \mathcal{Q})$ is the space of measurable functions from $\Theta$ to $\mathcal{Q}$.

We rely on a technical contribution which characterizes the solutions to optimal control problems with objectives which are upper-semicontinuous in the control variable and linear in the state variable. The necessary and sufficient conditions for optimality are presented and proven in Appendix B which can be skipped in a first reading.\footnote{These conditions generalize Jullien’s (2000) result that only applies for twice-continuously differentiable objectives although they take similar forms.} We specialize these conditions to program $(P_i)$ above in the following theorem.

**Theorem 1** Given the profile of transfers offered by rival principals, $t_{-i}$ and the agent’s corresponding outside option $U_{-i}(\theta)$, the rent-output profile $(\bar{U}, \bar{q})$ is a solution to $(P_i)$ if and only if $(\bar{U}, \bar{q})$ satisfies $(2.1)$, $(2.2)$ and $(2.3)$, and there exists a measure $\mu_i$ (possibly with mass points) defined over the Borel subsets of $\Theta$ with an associated adjoint function, $\bar{M}_i : \Theta \to [0, 1]$ defined by $\bar{M}_i(\theta) = 0$ and for $\theta > \theta$,

$$\bar{M}_i(\theta) = \int_{[\theta, \theta]} \mu_i(d\theta),$$

such that the following two conditions are satisfied:

(4.1) $\text{supp} \{\mu_i\} \subseteq \Theta_i^c \equiv \{\theta \mid U(\theta) = U_{-i}(\theta)\}$,

(4.2) $\bar{q}(\theta) \in \arg\max_{\theta \in \mathcal{Q}} S_i(q) + S_0(q) + T_{-i}(q) + \left( \frac{\bar{M}_i(\theta) - F(\theta)}{f(\theta)} - \theta \right) q$, a.e. $\theta \in \Theta$.

In single-principal contracting models with differentiable preferences and an outside option independent of an agent’s type, these conditions reduce to familiar requirements. For example, if $n = 1$ and $i \in \mathcal{A}$, we have a setting similar to Baron and Myerson’s...\footnote{$\bar{M}_i$ is piecewise absolutely continuous, but might have jumps wherever $\mu_i$ has mass points.}
(1982) model of regulation for a monopolist with unknown marginal cost and no outside option. Immediately, one can show that the participation constraint binds only for the worst type $\theta$ so that $\overline{M}_i(\theta) = 0$ for all types and the optimal allocation satisfies

$$S'_i(q) + S'_0(q) = \theta + \frac{F(\theta)}{f(\theta)}.$$ 

More generally, these optimality conditions illustrate how the standard rent-efficiency trade-off of the screening literature is modified in a competitive screening environment. When an agent with type $\theta$ behaves like a less (resp. more) efficient type $\theta + d\theta$ (resp. $\theta - d\theta$), he produces the same amount at a lower (resp. higher) cost. Mitigating these incentives forces a principal to distort output downwards for all types except the most (resp. least) efficient one. Those “incentive” distortions are captured by the usual hazard rate term $F'(\theta)/f(\theta)$ (resp. $F(\theta) - 1/f(\theta)$) that discounts each principal’s marginal contribution below (resp. above) his marginal valuation.

Because the type-dependent participation constraints that are specific to the delegated agency game limit the ability of any principal to extract rent and reduce outputs, other “participation” distortions must now be taken into account. Those distortions increase marginal contributions on any activity sets. They are captured by the new non-negative term $\overline{M}_i(\theta)/f(\theta)$.

When the measure $\mu_i$ does not put too much mass on types less than $\theta$, we have $\overline{M}_i(\theta) < F(\theta)$. This corresponds to the case where principal $i$ finds it relatively cheap to influence types below $\theta$. The driving force behind output distortions are then the incentive distortions which tend to reduce output and principal $i$’s marginal contribution. When instead $\mu_i$ puts significant mass on types lower than $\theta$, i.e., principal $i$ finds it too costly to induce participation from those types less than $\theta$ and prefers to stay inactive. We have $\overline{M}_i(\theta) > F(\theta)$ and output distortions are mitigated to ensure participation.

Some properties of $\overline{M}_i$ are immediate from the above result. First, $\overline{M}_i$ must be constant over any connected interval of $\overline{S}_i$. This follows from the fact that over such a set principal $i$ is active, hence the participation constraint is slack (i.e., $\mu_i = 0$), and it follows that $\overline{M}_i$ must be constant. Second, we know that if $S_i$ is differentiable and concave, then over any open interval of $\overline{S}_i$ for which $q_{-i}$ is strictly decreasing,

$$\overline{M}_i(\theta) = F(\theta) - S'_i(q_{-i}(\theta))f(\theta).$$

This conclusion follows because over any interval for which the participation constraint binds, it must be that $\overline{q}(\theta) = q_{-i}(\theta)$. This can only be the solution to (4.2), however, if (4.3) holds.\footnote{Indeed, one can establish the continuity of the adjoint function $\overline{M}_i$ by manipulating (4.1) and (4.2) and the fact that the agent’s reservation utility, $\overline{U}_{-i}$, is a convex function of $\theta$. This more general statement was proven in a previous version of this paper; in the present version, we prove adjoint continuity for the case in which principals’ preferences are linear.}

Notice that if the equilibrium allocation is strictly decreasing, (4.3) implies that each $\overline{M}_i$ is completely characterized for a given configuration of equilibrium activity, $\overline{a}$. For example, if it were known that each principal’s activity set is a single interval of $\Theta$, then constructing such an equilibrium is as simple as searching for the endpoints of each principal’s activity set and checking that the resulting $\overline{M}_i$ constructions satisfy the optimality conditions with respect to $(\overline{q}, \overline{U})$ in Theorem 1. In the sequel, we will consider the case where principals have linear surplus functions, $S_i(q) = q$. Proving that the
inactivity set $\overline{\gamma}_i$ is a connected interval of the form $[\hat{\theta}_i, \bar{\theta}_i]$ amounts then to checking that $\overline{M}_i(\theta) = F(\theta) - s_i f(\theta)$ is monotonically increasing on such interval; then the Lagrange multiplier $\mu_i(\theta) = \overline{M}_i(\theta) \geq 0$ is immediately retrieved. More generally, beyond this linear case, activity sets are connected whenever $\overline{M}_i(\theta) = F(\theta) - s_i'(\overline{\gamma}_i(\theta)) f(\theta)$ is monotonically increasing.

Of course, the conditions above for a best response do not yet characterize an equilibrium. To do so, we will need to derive more precise implications from these conditions. Indeed, one would like to describe both the equilibrium allocation and the corresponding adjoint functions with only minimal reference to the $n$-tuple of equilibrium transfers which implements them. Section 5 makes progress towards the first objective with necessary conditions for equilibrium allocations being derived with respect to the aggregate transfers and not their exact distribution among principals. Section 6 specializes the principals’ preferences to derive even sharper predictions on adjoint functions and equilibrium allocations.

5. EQUILIBRIUM ANALYSIS

This section unveils a set of necessary conditions that must hold at any equilibrium. Later we will use these conditions to construct a large class of equilibria including more than the "smooth" equilibrium allocation that was conjectured in Section 3.

5.1. General structure of equilibria

If $\{\pi_1, \ldots, \pi_n\}$ is an equilibrium profile of contracts which induces the allocation $q : \Theta \rightarrow \mathcal{Q}$, then it must each principal finds it weakly optimal to implement the allocation $\overline{q}$, given the transfers offered by the other principals. Formally, Theorem 1 implies that $\overline{q}$ must be a solution for each $i \in N$ to the program

$$\max_{q \in \mathcal{Q}} S_i(q) + S_0(q) + \sum_{j \neq i} t_j(q) + \left( \frac{\overline{M}_i(\theta) - F(\theta)}{f(\theta)} - \theta \right) q, \ a.e. \theta \in \Theta.$$ 

If $\overline{q}$ maximizes each of these programs, then it must also maximize any aggregation of these objectives. This immediately implies that an important set of necessary conditions hold for any equilibrium.

**Theorem 2** In any equilibrium of the delegated agency game with activity sets represented by $\pi$ and the principals’ adjoint profile given by $\{\overline{M}_1, \ldots, \overline{M}_n\}$, the allocation satisfies two conditions

\begin{align}
(5.1) \quad \overline{q}(\theta) & \in \arg \max_{q \in \mathcal{Q}} S_0(q) - \theta q + \sum_{i=1}^{n} S_i(q) + \sum_{i=1}^{n} \left( \frac{\overline{M}_i(\theta) - F(\theta)}{f(\theta)} \right) q \\
& \quad + (n - 1) \left( S_0(q) - \theta q + \mathcal{T}(q) \right), \ a.e.
\end{align}

\begin{align}
(5.2) \quad \overline{q}(\theta) & \in \arg \max_{q \in \mathcal{Q}} S_0(q) - \theta q + \sum_{i \in \pi(\theta)} S_i(q) + \sum_{i \in \pi(\theta)} \left( \frac{\overline{M}_i(\theta) - F(\theta)}{f(\theta)} \right) q \\
& \quad + (|\pi(\theta)| - 1) \left( S_0(q) - \theta q + \mathcal{T}(q) \right), \ a.e.
\end{align}
The two ingredients giving rise to these necessary conditions are Theorem 1 and the insight that because the allocation can be controlled by individual principals, in equilibrium all principals must be satisfied with its choice. This latter idea has been called “the aggregate-concurrence principle for aggregate games” (Martimort and Stole (2012)).

The objective function in (5.1) is obtained by summing over all principals the optimality conditions (4.2). Proceeding similarly but over active principals only, we obtain a more precise requirement in (5.2). The difference between the two is that inactive principals do not change the agent’s decision; in other words, their objectives are aligned with that of the agent given the contracts of active principals. This congruence is captured by the joint requirement that an equilibrium allocation satisfies (5.1) and (5.2). Turning to this latter condition, the last term

$$(|\alpha(\theta)| - 1) \left( S_0(q) - \theta q + \bar{T}(q) \right),$$

is a multiple of the agent’s equilibrium maximand, which of course is maximized for any incentive compatible allocation. If the transfers are continuously differentiable around the equilibrium allocation, then the presence of this term in the maximand is immaterial by the agent’s first-order condition. We consider more subtle arguments that allow for non-smooth equilibria below; in such equilibria this term plays a direct role.

For the present, consider the first terms in (5.2)

$$S_0(q) - \theta q + \sum_{i \in \alpha(\theta)} S_i(q) + \sum_{i \in \alpha(\theta)} \left( \frac{\overline{M}_i(\theta) - F(\theta)}{f(\theta)} \right) q.$$ 

This simplified objective provides a simple intuition for the equilibrium allocation: it is as if the allocation $\bar{q}(\theta)$ is chosen to maximize the collective surplus of the agent and the principals in the active coalition $\alpha(\theta) \subseteq N$, less an information rent term. This term, moreover, is a pointwise sum of the agent’s bilateral information rents vis-à-vis each active principal which captures the idea that agency problems are compounded under delegated agency. If we define the average adjoint function across the coalition as

$$\overline{M}_S(\theta) \equiv \frac{1}{|S|} \sum_{i \in S} \overline{M}_i(\theta),$$

then the aggregate information rent term is $|\alpha(\theta)|$ times larger than the rent term for the average principal:

$$\sum_{i \in \alpha(\theta)} \left( \frac{\overline{M}_i(\theta) - F(\theta)}{f(\theta)} \right) q = |\alpha(\theta)| \left( \frac{\overline{M}_{\alpha(\theta)}(\theta) - F(\theta)}{f(\theta)} \right) q.$$ 

In particular, if each principal’s adjoint function $\overline{M}_i$ is identical within the coalition, which is the case if all principals do agree on the identity of the “worst” agent’s type from their point of view, then the information rent term of the coalition is exactly $|\alpha(\theta)|$ times

\footnote{The genesis of the “aggregate concurrence principle” can be found in Cournot’s output games. In such a game, each firm has the ability to increase the aggregate output (and thereby reduce the equilibrium price), so each active firm must be satisfied with the equilibrium price, given the supplies of its rival. Selten (1970) was one of the earliest game theory papers to make explicit use of this property. Bernheim and Whinston (1986b) employ a similar idea in their analysis of intrinsic common agency games with moral hazard.}
larger than the individual terms. Hence, all else equal, an $|\bar{\alpha}(\theta)|$-fold larger distortion is introduced.

While the preceding remarks are economically interesting, we must emphasize that $\bar{\alpha}$ is an endogenous object, so the necessary condition (5.2) is more insightful than practical.\footnote{Under intrinsic common agency, $\bar{\alpha}(\theta) \equiv N$, and the necessary condition above can be shown under mild conditions to also be a sufficient condition for equilibrium. (See Martimort and Stole (2012).)} Nevertheless, Section 6 makes significant progress toward an understanding of (5.2) by restricting attention to the case where principals have linear preferences.

6. COMMON-AGENCY GAMES WITH LINEAR PAYOFFS

Characterizing equilibrium activity sets is difficult without more structure on preferences. In the general case, difficulty arises because each principal’s marginal virtual surplus may depend upon the level of agent activity. Recall, for example, that in our heuristic derivation of smooth equilibrium allocations under the assumption of linear preferences, we were able to define a principal’s marginal virtual surplus as

$$\beta_i(\theta) \equiv \begin{cases} \max \left\{ s_i - \frac{F(\theta)}{f(\theta)}, 0 \right\} & i \in \mathcal{A}, \\ \min \left\{ s_i + \frac{1-F(\theta)}{f(\theta)}, 0 \right\} & i \in \mathcal{B}, \end{cases}$$

and thus the sum of the principals’ marginal virtual utilities is

$$\beta(\theta) \equiv \sum_{i \in \mathcal{N}} \beta_i(\theta) = \sum_{i \in \mathcal{A}} \max \left\{ s_i - \frac{F(\theta)}{f(\theta)}, 0 \right\} + \sum_{i \in \mathcal{B}} \min \left\{ s_i + \frac{1-F(\theta)}{f(\theta)}, 0 \right\}.$$ 

Without linear preferences, we would need to replace $s_i$ with $S_i'(\bar{q}(\theta))$, which in turn depends upon the allocation $\bar{q}$; the activity sets and the equilibrium allocation would need to be jointly determined. Providing that the principal’s preferences do not display too much curvature, this dependence does not pose too great of a technical difficulty.\footnote{Consider for instance the case of two principals with similar preferences, $S_1(q) = S_2(q) = S(q)$ and define $k = \sup_{\theta} x'(q) + s''(q)$. When the type distribution is uniform, and $k \geq 2$, one can show that activity sets are again of the form $[\hat{\theta}_i, \check{\theta}_i]$. (Proof available upon request.)}

To bypass these technical difficulties altogether, however, throughout the remainder of the paper we focus attention on common-agency games in which the preferences of the principals (but not necessarily the agent) are linear in $q$. To be precise, we offer the following specialization to linear games.

**Definition 4** A delegated common-agency game is linear if for all $i \in \mathcal{N}$,

$$S_i(q) = s_i q,$$

and additionally $s_i f(\tilde{\theta}) < 1$ for $i \in \mathcal{A}$, $|s_i| f(\tilde{\theta}) < 1$ for $i \in \mathcal{B}$, and $S_0$ is concave.

The main import of the linearity assumption is that each principal’s marginal virtual utility is independent of output. This is the fundamental ingredient in results which follow. The additional requirement regarding the magnitude of each $s_i$ relative to the density of types is for technical convenience; it guarantees that for each principal $i$ there is a unique root $\tilde{\theta}_i \in \Theta$ such that, for $i \in \mathcal{A}$, $s_i - F(\tilde{\theta})/f(\tilde{\theta}) = 0$, and for $i \in \mathcal{B}$,
\[ s_i + (1 - F(\theta_i))/f(\theta_i) = 0. \] The existence of such a root avoids indeterminacies in the construction of tariffs when multiple principals are active for all types. The concavity requirement implies that the total surplus function, \( S_0(q) + \sum_{i \in N} S_i(q) - \theta q \) is concave (and therefore almost everywhere differentiable), which simplifies the analysis.

With our additional assumption of linearity, we are able to greatly simplify the equilibrium analysis by proving that each principal’s adjoint equation from Theorem 1 can be replaced with a simpler expression that is independent of the equilibrium allocation.

**Lemma 2** In the linear common-agency game, if \((\bar{\nu}, \bar{U})\) is an equilibrium allocation, then for each principal \(i\), \((\bar{\nu}, \bar{U})\) satisfies the conditions in Theorem 1 using the adjoint function

\[
(6.1) \quad M_i(\theta) = \begin{cases} 
\max\{F(\theta) - s_i f(\theta), 0\}, & i \in A, \forall \theta \in (\theta, \bar{\theta}), \\
\min\{F(\theta) - s_i f(\theta), 1\}, & i \in B, \forall \theta \in (\theta, \bar{\theta}),
\end{cases}
\]

such that \( M_i(\theta) = 0 \).

Linearity provides a closed-form solution of the adjoint equations that is independent of the equilibrium allocation. We are thus able to construct the principals’ activity sets independently of the choice of equilibrium allocation, greatly simplifying the characterization of equilibria, and present our main characterization theorem as a pair of intuitive, straightforward optimization programs:

**Theorem 3** Suppose that the common-agency game is linear. If \( \bar{\nu} \) is an equilibrium allocation, then

\[
(6.2) \quad \bar{\nu}(\theta) \in \arg \max_{q \in \bar{\nu}(\Theta)} S_0(q) + (\beta(\theta) - \theta)q, \text{ for all } \theta \in \Theta.
\]

Moreover, if an allocation \( \bar{\nu}^Q(\theta) \) satisfies

\[
(6.3) \quad \bar{\nu}^Q(\theta) \in \arg \max_{q \in Q} S_0(q) + (\beta(\theta) - \theta)q, \text{ for all } \theta \in \Theta,
\]

then \( \bar{\nu}^Q \) is an equilibrium allocation. Such an equilibrium allocation always exists.

The difference between conditions (6.3) and (6.2) is subtle, but significant. Condition in (6.3) is stronger than (6.2) since it requires optimality over the whole set of possible actions, \( Q \). Because we will see there are some equilibrium allocations which satisfy (6.2) but do not satisfy (6.3), the latter condition implicitly refines the equilibrium set. When, for instance, \( S_0 \) is strictly concave, this restriction implies that \( \bar{\nu} \) is in fact continuous. Instead, condition (6.2) is a priori compatible with the existence of discontinuity gaps in the output profile. We will see below that the equilibrium set is large – an infinite number of both continuous and discontinuous allocations are equilibrium outcomes. Much of the remaining analysis in this section aims at characterizing the range of outcomes, \( \bar{Q} = \bar{\nu}(\Theta) \), that may arise in an equilibrium.

Theorem 3 provides some justification for the heuristic approach which began the analysis in Section 3. Such “smooth” equilibria do, in fact, exist when the principals’ preferences are linear, and the associated equilibrium allocations take the simple form of (6.3). Hence, the sufficiency portion of Theorem 3 lays to rest many of the concerns that were
initially articulated with the approach. There are an infinity of other equilibrium allocations, of course, but the necessity portion of Theorem 3 gives a measure of insight in that all such equilibria have a similar information-rent distortion, captured by the difference between $\sum s_i$ and $\beta(\theta)$. Indeed, an immediate corollary of Theorem 3 is that over any open interval of any equilibrium range, $(q_1, q_2) \subseteq \overline{Q} = \overline{q}(\theta)$, the equilibrium allocation must coincide with the allocation, $\overline{q}^2$, defined by (6.3). All of the differences between the various equilibrium allocations can be entirely catalogued by the differences in their equilibrium ranges.

6.1. $Q$-Maximal equilibrium allocations

The equilibrium allocation defined by (6.3) has many interesting characteristics in addition to the property of existence. First, as demonstrated in the proof, it can be implemented with continuous, continuously differentiable tariffs. And while each principal’s equilibrium tariff is shown to be convex, the agent’s aggregate objective, $S_0(q) - \theta q + \overline{T}(q)$ is concave in $q$. Second, one can see in comparison of (6.2) and (6.3), that the latter has the largest possible domain of optimization. For this reason, we refer to any allocation which satisfies (6.3) as a $Q$-maximal equilibrium allocation and will denote it by $\overline{q}^Q$ to distinguish it from other non-maximal equilibrium allocations. Given that $S_0$ is concave, such an allocation is uniquely defined for almost every $\theta$; with strict concavity, this allocation is unique everywhere.

The strength of Theorem 3 comes from the fact that (6.3) provides a clear characterization of $Q$-maximal allocations as solutions of a maximization problem over the full domain $Q$. In comparison with condition (5.1), this is a tremendous simplification: any endogenous variables related to activity sets and equilibrium transfers have now disappeared.

Equilibria which satisfy (6.3) have an appealing property that each active principal “shades” his marginal valuation for the agent’s marginal action. Indeed, the marginal equilibrium transfer that principal $i$ offers is given by

$$\overline{\tau}_i(q) = \beta_i(\overline{\mathcal{V}}(q)) = \begin{cases} 
\max \left\{ s_i - \frac{\overline{F}(\overline{\mathcal{V}}(q))}{f(\overline{\mathcal{V}}(q))}, 0 \right\}, & i \in \mathcal{A}, \\
\min \left\{ s_i + \frac{1 - \overline{F}(\overline{\mathcal{V}}(q))}{f(\overline{\mathcal{V}}(q))}, 0 \right\}, & i \in \mathcal{B}, 
\end{cases}$$

where as before, $\overline{\mathcal{V}}(q)$ is the appropriate inverse of $\overline{q}(\theta)$. Immediately we see that principals who like the agent’s action, $i \in \mathcal{A}$, reduce their marginal contributions, whereas those who dislike $q$, $i \in \mathcal{B}$, exaggerate these marginal contributions. Indeed, as the agent’s private information shrinks, each principal’s marginal virtual valuation converges to his true marginal valuation and equilibrium marginal transfers $\overline{\tau}_i(q)$ converge to $s_i$. Thus, the equilibria satisfying (6.3) are a natural generalization of Bernheim and Whinston’s (1986a) truthful equilibria in games of complete information.17

17 The precise argument is that as the agent’s private information becomes less heterogeneous, the equilibrium marginal transfers which support the allocation in (6.3) converge in a strong sense to $s_i$ for each principal, which in turn corresponds to the truthful equilibrium of the complete-information menu auction game in Bernheim and Whinston (1986). Consider a class of absolutely-continuous distributions of $\theta$ in which each distribution has the same mean and each satisfies the generalized monotone-hazard rate property. Suppose that this class of distributions is second order stochastically ordered, and let $\kappa \in [0, \infty)$ be an index representing a mean-preserving contraction of $F$ around the common mean. An
6.2. Other (non-maximal) equilibrium allocations

We now consider the set of allocations which satisfy (6.2) but not (6.3). Theorem 3 implies that some characteristics of maximal allocations are shared by all equilibrium allocations. As noted above, we have the following immediate implication.

**Corollary 1** If \( \bar{q} \) is an equilibrium allocation in a linear common-agency game, then

\[
\bar{q}(\theta) \in \text{int} \, \bar{q}(\Theta) \implies \bar{q}(\theta) = \bar{q}^Q(\theta).
\]

Because all equilibrium allocations within the interior of the equilibrium range are equal to the maximal allocation, we have another rationale for the "maximal" designation. The differences between all equilibrium allocations can be completely summarized by the gaps in \( \bar{Q} \) that do not arise in the \( Q \)-maximal allocation.

When \( S_0 \) is strictly concave and the maximal allocation \( \bar{q}^Q \) satisfying (6.3) is necessarily continuous, an equilibrium allocation which satisfies only (6.2) may exhibit discontinuities. That said, such discontinuities are still restricted by the underlying program. In particular, there is an indifference condition which must be satisfied across any pair of discontinuity points.

**Corollary 2** If \( \bar{q} \) is an equilibrium allocation with a discontinuity at an interior \( \theta_0 \), then the right-hand limit, \( q_1 = \bar{q}(\theta_0^+) = \lim_{\theta \to \theta_0^+} \bar{q}(\theta) \), and the left-hand limit, \( q_2 = \bar{q}(\theta_0^-) = \lim_{\theta \to \theta_0^-} \bar{q}(\theta) \), exist and must satisfy

\[
(6.4) \quad S_0(q_1) + (\beta(\theta_0) - \theta_0)q_1 = S_0(q_2) + (\beta(\theta_0) - \theta_0)q_2.
\]

This requirement places additional structure on the character of equilibrium discontinuities. At any discontinuity point, \( \theta_0 \), the corresponding discontinuity points in the range of the equilibrium allocation must be equally optimal with respect to \( S_0(q) + (\beta(\theta_0) - \theta_0)q \). Furthermore, providing the discontinuity is sufficiently small and satisfies these indifference, then we can always introduce a discontinuity in \( \bar{q}^Q \) and construct another equilibrium allocation.

**Theorem 4** Let \( S_0 \) be strictly concave and \( \bar{q} \in \text{int} \, \bar{q}^Q(\Theta) \) be an interior action at which all principals are active in the maximal equilibrium. If all principals’ preferences are congruent, then there exists a sufficiently small neighborhood \( (q_1, q_2) \) of \( \bar{q} \) such that the allocation

\[
(6.5) \quad \bar{q}(\theta) \in \arg \max_{q \in \bar{q}^Q(\Theta) \setminus (q_1, q_2)} S_0(q) + (\beta(\theta) - \theta)q,
\]

is also an equilibrium allocation. If the agent’s type is uniformly distributed, then the conclusion of (6.5) holds regardless of whether preferences are congruent or opposed.

In addition, every agent type weakly prefers the maximal equilibrium \( \bar{q}^Q \) to \( \bar{q} \), i.e.,

\[
\bar{q}^Q(\theta) \geq \bar{U}(\theta), \text{ with strict preference for some positive measure of types.}
\]

arbitrary distribution in this class is denoted \( F(\theta, \kappa) \) and \( \lim_{\kappa \to -\infty} F(\theta, \kappa) \) is a Dirac distribution centered on the common mean. Using our previous construction of marginal transfers, we may thus write the marginal transfer of principal \( i \in \mathcal{A} \), for a given distribution \( \kappa \), as \( \tau_i(\theta, \kappa) \equiv \max \left\{ s_i - \frac{F(\theta, \kappa)}{f(\theta, \kappa)}, 0 \right\} \). It is straightforward to check that the expected marginal transfer of principal \( i \) in the complete-information limit game is principal \( i \)'s marginal benefit: \( \lim_{\kappa \to -\infty} \int_{\Theta} \tau_i(\theta, \kappa) f(\theta, \kappa) d\theta = s_i \) as in a truthful equilibrium. In Martimort and Stole (2009b), we develop other arguments to select among the equilibrium set.
Applying Theorem 3 and the arguments used in the proof of Theorem 4, we can also conclude that wherever a discontinuity exists in an equilibrium allocation, the allocation is equal to the maximal allocation at the discontinuity endpoints in the range, with pooling of types at these endpoints.

**Corollary 3** Suppose that $S_0$ is strictly concave and differentiable and that $q$ is an equilibrium allocation such that $q^2(0) < q^2(0^-)$. If $q$ is discontinuous at $0_0$, then for the discontinuity points $q_1 = q^2(0_0^+)$ and $q_2 = q^2(0_0^-)$, there exist $0_1 > 0_2$ such that $q_1 = q^2(0_1)$ and $q_2 = q^2(0_2)$, and

$$q(\theta) = \begin{cases} q_1 = q^2(0_1), & \text{for } \theta \in (0_0, 0_1] \\ q_2 = q^2(0_2), & \text{for } \theta \in [0_2, 0_0). \end{cases}$$

The above characterization bears strong similarities with the literature on mechanism design without transfers in monopolistic screening environments. In that literature, much effort has been devoted to characterize possible actions implementable in a context with asymmetric information when the preferences of the principal and the agent differ. The basic lesson of this literature is that any implementable action may either be flat over some range and not responsive to the agent’s private information or correspond to the latter’s ideal point. In light of this literature, everything happens as if, over a range $Q$, the equilibrium output $q$ were chosen by a surrogate principal who aggregates the behavior of all principals and maximizes their aggregate virtual surplus as defined in (6.2). Of course, the objective of this surrogate principal differs from what would be optimal had principals merged; the difference being related to the fact that in a non-cooperative context, each principal introduces output distortions for rent extraction reasons.

The discontinuous equilibria so constructed have the property that, on both sides of a discontinuity, the rent profile $U(\theta)$ has a kink at $\theta_0$ with adjacent linear segments having slopes $q_1 = q^2(0_0^-) = \lim_{\theta \to \theta_0^-} q(\theta)$ and $q_2 = q^2(0_0^+) = \lim_{\theta \to \theta_0^+} q(\theta) > q_1$.

7. APPLICATIONS

We now turn to our two original motivating applications so as to provide a better understanding of the economics of delegated common agency games. In the public goods games below, all principals/contributors have congruent objectives: principals value the public good, but each with possibly different intensity. In the lobbying example, two principals/lobbying groups have conflicting objectives and want to push policies in opposite directions. The analysis of those games in a context with asymmetric information unveils new economic effects that enrich our knowledge of those economic settings. Our focus is mostly on the maximal equilibrium allocations in each game, but we further explore non-maximal equilibria in the context of the lobbying game.

7.1. Voluntary provision of a public good

7.1.1. Public goods game in which $S_0$ is strictly concave.

Consider a public goods game in which $S_0$ is strictly concave.

---

to scale. We also assume that \( Q = [0, q_{\text{max}}] \) with \( q_{\text{max}} \) sufficiently large (e.g., \( q_{\text{max}} \geq -\theta + \sum_{i \in N} s_i \)) and \( \Theta = [0, \theta] \).

**Proposition 1** In the public goods game with decreasing returns to scale \( S_0(q) = -\frac{1}{2}q^2 \), the maximal equilibrium allocation is

\[
\bar{q}^Q(\theta) = \max \left\{ \sum_{i \in N} \max \left\{ s_i - \frac{F(\theta)}{f(\theta)}, 0 \right\} - \theta, 0 \right\}.
\]

**Example:** For \( N = \{1, 2\} \), the efficient output is \( q^{FB}(\theta) = \max\{s_1 + s_2 - \theta, 0\} \), the cooperative solution had principals cooperated \( q^{C}(\theta) = \max\{s_1 + s_2 - \theta, 0\} \) entails only one screening distortion whereas the optimal output that principal 1 would implement if he was alone is \( q^*_1(\theta) = \max\{s_1 - \theta - \frac{F(\theta)}{f(\theta)}, 0\} \). It again entails only one screening distortion as it can be seen on Figure 1 below. The equilibrium output of the delegated agency game exhibits two such distortions when principal 2 finds it also worth to intervene:

\[
\bar{q}^Q(\theta) = \max \left\{ s_1 + s_2 - \theta - 2\frac{F(\theta)}{f(\theta)}, s_1 - \theta - \frac{F(\theta)}{f(\theta)}, 0 \right\}.
\]

**Figure 1:** Voluntary provision of a public good.

This example nicely illustrates how distortions in delegated common agency games with congruent principals manifest themselves in two dimensions. First, because each active principal contributes less than his marginal valuation, inefficient provision arises at the intensive margin arises. The equilibrium output is lower than the cooperative solution and features the same two-fold distortion that is present in intrinsic common-agency games. A second distortion, novel to delegated agency games, emerges from limited participation by the weaker principal; output is also distorted at the extensive margin. In this example, there exists an non-empty interval of types, \( s_2 \leq \frac{F(\theta)}{f(\theta)} \leq s_1 \), such that only principal 1 is active under asymmetric information while both principals would be active under complete information.

The fact that output is inefficiently low may suggest that the familiar free-riding problem in classic public-goods contributions games is also present in public goods games with more complex strategy spaces. This interpretation is potentially misleading. First, if
information were complete (which is tantamount to eliminating the inverse-hazard terms from the equation), the maximal equilibrium leads to full efficiency: each principal offers the marginal tariff \( t_i'(q) = s_i \). This is the same efficient equilibrium outcome that arises in Bernheim and Whinston’s (1986b) “truthful equilibrium.” Thus, free riding need not arise in complete-information public-goods games if the principals have the ability to offer nonlinear tariffs rather than one-dimensional contributions. When incomplete information is present, however, each principal has a private incentive to distort the agent’s output choice to extract additional information rent. Because each principal ignores the negative externality that this imposes on others, from a collective viewpoint, the principals inefficiently extract too much rent. Thus, it is probably more accurate to say that the distortion is akin to a “tragedy of the commons” rather than “free riding.” Each principal over harvests the common resource – the agent’s information rent.

It is worth observing that the equilibrium output is not necessarily invariant with respect to redistributions of the principals’ surplus, keeping the aggregate \( \sum_{i=1}^{n} s_i \) constant. Thus, a unit tax on principal 1’s use of the public good that is exactly offset by a unit subsidy on principal 2’s use would have a real impact on the equilibrium allocation of public goods by changing the set of active principals. This is not the case in models of intrinsic common agency, as shown in Martimort and Stole (2012), because in such games all principals are active on the same type set. This insight is reminiscent of the public finance literature on voluntary contributions (see, e.g., Bergstrom, Blume and Varian (1986), et al.) although this literature obtains that non-neutrality result by making a restriction on instruments (fixed contributions instead of schedules) and assumes complete information while the very same result is obtained here with asymmetric information and no such restriction on contracts.

Another interpretation of the limited participation that may arise under asymmetric information is that some form of exclusive contracting emerges endogenously even if exclusivity clauses cannot be enforced at the outset. This is so even if both principals would otherwise have contracted with the agent under complete information. This finding is reminiscent of an important insight developed by Bernheim and Whinston (1998) in their study of vertical relationships between manufacturers and retailers. They showed that exclusive dealing in marketing practices arises when the agency costs of a common representation are too large compared with those under exclusive dealing. There is, however, an important difference between their result and ours. They assume that the possibility of exclusive representation arises ex ante, i.e., before the realization of uncertainty. Although their general contracting model is thus consistent with hidden actions or hidden information, it cannot account with the possibility of exclusivity arising for some realization of shocks and not for others. In this regard, our model, where contracting takes place once the agent is already informed, generates such richer patterns of behaviors.

7.1.2. Public goods game with linear payoffs \( q \in [0, 1] \).

Suppose now that \( Q = [0, 1] \) and \( S_0(q) = 0 \), i.e., the agent has no value for the good and \( q \) can be interpreted as a probability of realizing an indivisible public good and all preferences are now linear. For simplicity, assume that \( \sum_{i \in N} s_i \geq \theta \) (i.e., the first best has some public good provision) and that \( s_1 \geq \cdots \geq s_n \). Our assumption that for each principal \( i \) there exists an interior type \( \theta_i \) such that \( s_i f(\theta_i) = F(\theta_i) \) implies that
\[ \bar{\theta} \geq \hat{\theta}_1 \geq \cdots \geq \hat{\theta}_n \geq \theta \] and that there is a unique \( \hat{\theta} \) such that
\[ \sum_{i \in N} \max \left\{ s_i - \frac{F(\hat{\theta})}{f(\hat{\theta})}, 0 \right\} = \hat{\theta}. \]

A direct application of (6.3) in Theorem 3 generates the following conclusion.

**Proposition 2** In the public goods game with linear preferences, the maximal equilibrium allocation satisfies
\[ (7.2) \quad \bar{q}^Q(\theta) = \begin{cases} 
1 & \theta < \hat{\theta} \\
0 & \theta > \hat{\theta}.
\end{cases} \]

The equilibrium transfers are linear in \( q \)
\[ (7.3) \quad \bar{t}^Q_i(q) = \max \left\{ s_i - \frac{F(\hat{\theta})}{f(\hat{\theta})}, 0 \right\} q, \quad q \in [0, 1], \ i \in N. \]

This linear example illustrates properties which must be satisfied in the construction of discontinuous equilibria proposed in Theorem 4 when \( S_0 \) is not strictly concave. Corollary 2 implies that necessarily any discontinuity is at \( \hat{\theta} \) in any equilibrium. This fact taken, together with payoff linearity and the fact that each principal wants to push the decision on extreme points of \( Q = [0, 1] \) implies that the maximal equilibrium is essentially unique, up to an arbitrary specification of \( \bar{q}^Q(\hat{\theta}) \).

In a related context, Lebreton and Salanié (2003) model competing lobbying groups in a setting with a binary choice of policy and a decision-maker having private information on the weight he gives to social welfare in his objective. A given interest group is restricted to offer a single positive payment paid if and only if the decision-maker implements the decision that group prefers. Proposition 2 highlights conditions under which Lebreton and Salanié’s (2003) result holds more broadly when random policy choices are possible and no restriction is a priori imposed on contribution schedules.\(^{19}\)

### 7.2. Influence and Lobbying Games

Consider now two competing interest groups having instead conflicting preferences \( s_1 > 0 > s_2 \). For instance, principal 1 prefers a higher tax rate, \( q \), whereas principal 2 prefers a lower tax rate. The decision-maker (agent) has some ideal policy he would like to pursue in the absence of any influence by lobbying groups. To model these preferences, assume that \( S_0(q) = -\frac{q^2}{2} \) where \( q \in Q = [-q_{\text{max}}, q_{\text{max}}] \) with \( q_{\text{max}} \) being large enough to ensure interior solutions to (6.3). Assume also that the agent’s ideal point \( q_0(\theta) = -\theta \) is symmetrically distributed over \([-\delta, \delta]\) with \( \delta < 1.\)\(^{20}\) Choosing this bliss point gives a status quo payoff \( U_0(\theta) = \frac{\theta^2}{2} \) to the agent.\(^{21}\) Our maintained assumption that for each \( i \in N \) there exists an interior type \( \hat{\theta}_i \) such that \( \beta_i(\hat{\theta}_i) = 0 \) requires that \( f(\delta) < 1.\)

\(^{19}\) Analyzing a model where competing principals have private information on their marginal valuations for a discrete public good but the agent has no private information, Laussel and Palfrey (2003) use techniques from the auction literature to derive equilibrium conditions in such contexts. Virtual surpluses there reflect the principals’ private information and not that of the agent as in our context.

\(^{20}\) Since principals are symmetrically biased in opposite directions, they would just agree on letting the agent choose his status quo policy had they cooperated.

\(^{21}\) We could mirror the analysis in Section 7.1.2 and entertain the possibility that competing interest groups are opposed on whether to undertake a discrete policy (allowing free trade or not, allowing some drugs or not, etc.). We leave to the reader to develop such straightforward extension.
7.2.1. Segmented markets for influence in maximal equilibria

Applying the general methodology developed in Theorem 3, we obtain

**Proposition 3** The maximal equilibrium allocation of the lobbying game with \( s_1 > 0 > s_2 \) is

\[
\bar{q}^Q(\theta) = \max \left\{ s_1 - \frac{F(\theta)}{f(\theta)}, 0 \right\} + \min \left\{ s_2 + \frac{1 - F(\theta)}{f(\theta)}, 0 \right\} - \theta.
\]

If \( \theta \) is uniformly distributed, the activity sets of the principals are

\[
\Theta_1 = [-\delta, \min\{s_1 - \delta, \delta\}] \quad \text{and} \quad \Theta_2 = (\max\{\delta + s_2, -\delta\}, \delta].
\]

If type heterogeneity is small relative to the strength of the principals’ preferences, \( \delta < \frac{s_1 + |s_2|}{2} \),

then the principals commonly influence a positive measure of intermediate-type agents; otherwise, each principal has a separate domain of influence.

The lobbying model shows that only decision-makers with mild preferences receive contributions from both interest groups. Unchallenged influence only arises endogenously for the decision-makers who are the most “ideologically” oriented.\(^{22}\) This is, of course, a much richer pattern of influence and contributions than what is predicted by complete information lobbying games as in Grossman and Helpman (1994) or Dixit and al. (1997). In those complete information models, group \( i \) enjoys exclusive influence on policy only when other potential interest groups are just indifferent between that policy induced by group \( i \) and other policies that they may induce with positive contributions. Moreover, the absence of heterogeneity in the decision-maker’s preferences in those models makes it impossible to generate different patterns of contributions and thus it remains a puzzle in that literature as to why some groups target some legislators and not others.

\(^{22}\)Martimort and Semenov (2008) derive further results on the patterns of contributions in a lobbying game with a different objective function for the agent.
7.2.2. A competitive nonlinear pricing reinterpretation

Interestingly, the lobbying model can be transposed mutatis mutandis to an industrial organization setting to study how a consumer having private information on his most preferred bundle mixes between two goods marketed by two competing sellers. Suppose that this consumer wants to acquire one unit of an homogenous good and is located at a point \( \theta \in [0, 1] \) on a unit line, one seller being located at each extreme. The consumer has a valuation \( v \) for the good and incurs a quadratic transportation cost \( -\frac{1}{2}(q - \theta)^2 \) when moving away from his “ideal mix” where he consumes \( q_0(\theta) = \theta \) from principal 1 and \( 1 - q_0(\theta) \) from principal 2. Up to some normalizations, the consumer and the sellers’ profits are similar to those of the lobbying model above when the marginal cost is the same for both sellers and equal to one. Our previous results can be reinterpreted as giving conditions under which a share of the market is always covered by both sellers. When type-heterogeneity is sufficiently, mixed bundling arises and global exclusivity cannot be an equilibrium. Hoernig and Valletti (2010) have independently derived a similar insight but, at the outset, restricted their analysis to smooth tariffs. As we will see below when studying discontinuous equilibria in the (similar) lobbying game, this restriction may indeed be justified because such smooth equilibrium may have attractive welfare properties among a much larger class of equilibria allowing for discontinuities.

7.2.3. Discontinuous equilibria in the lobbying game.

In that specific lobbying context, we establish the existence of discontinuous equilibria by extending the local arguments used in the proofs Theorem 4 and Corollary 3 to construct arbitrary discontinuous allocations and then verifying that they are supported by
equilibrium transfer functions provided that the size of the discontinuities are bounded. To this end, we assume a uniform and symmetric distribution around zero, and we introduce a single discontinuity to the maximal (continuous) allocation at $\theta = 0$. The following proposition provides an exact upper bound on the size of the equilibrium discontinuity; this bound makes clear that such discontinuity gaps may be significant.

**Proposition 4** Suppose that $s_1 = -s_2 = 1 < 2\delta$, $S_0(q) = -\frac{q^2}{2}$ and that $\theta$ is uniformly distributed on $\Theta = [-\delta, \delta]$. For any $q_0 \in (0, (1 - \delta)\sqrt{3})$, there exists an equilibrium with a discontinuity at $\theta_0 = 0$ and such that $\overline{q}(0^-) = -\overline{q}(0^+) = q_0$. Both the agent’s rent and the principals’ expected payoffs in such discontinuous equilibria are lower than at the maximal equilibrium.

In the proof of Proposition 4, we provide a construction of the tariffs supporting the discontinuous allocation and show that indeed the tariffs comprise an equilibrium to the common-agency game. The tariffs have a very natural structure. If $\overline{\ell}_i^q$ is principal $i$’s equilibrium tariff in the maximal equilibrium, and if the hypotheses of Proposition 4 are satisfied, then the modified tariffs

$$\tilde{\ell}_i(q) = \begin{cases} 0 & \text{for } q \in (-q_0, q_0), \\ \overline{\ell}_i^q & \text{otherwise}, \end{cases}$$

support the discontinuous equilibrium.

To sustain those equilibria, principals design their contracts with “non-serious” out-of-equilibrium offers. For instance, principal 2 stipulates zero payments for outputs within the discontinuity gap $[\overline{q}(\theta_0^+), \overline{q}(\theta_0^-)]$ which are such that principal 1 is just indifferent to inducing the agent with type $\theta_0$ to produce any output within that range. This construction makes it possible to sustain the discontinuity in the agent’s choice.\(^{23}\) Importantly, we demonstrate in the Appendix that a discontinuity can only be sustained if the equilibrium schedules lie below the maximal ones on the discontinuity gap. On the range of equilibrium outputs corresponding to those discontinuous equilibria, principals offer schedules which have the same margin as the maximal equilibria. So doing ensures that the agent still chooses the maximal output on any connected set in that range.

The comparison of the players’ payoffs across equilibria in the lobbying context shows that the maximal equilibrium Pareto dominates, making it of focal interest. Not only the agent but also principals lose from coordinating on a discontinuous equilibrium. From Theorem 4, this result is clear for the agent since aggregate payments in those discontinuous equilibria are lower than at the maximal one. To explain the principals’ preferences, observe that not paying the agent for policy choice within the discontinuity gap has two effects. First, it increases polarization since types nearby the discontinuity now pool at the boundaries of that discontinuity gap. This corresponds to more extreme policies than under the maximal equilibrium. Because principals have opposite preferences, this reallocation effect has no impact on their aggregate gross surplus. Second, those types who pool on the “wrong” decisions end up being paid excessively compared with the maximal equilibrium. This is costly for the principals.

\(^{23}\)By the same token, such construction could be replicated to sustain equilibria with multiple discontinuities.
8. CONCLUSION

This paper has developed a methodology for solving public delegated common agency games under asymmetric information. In a nutshell, the basic economic insights of this research is that the well-known rent-efficiency trade-off must now be significantly modified to account for the impact of competition among principals. In particular, the compounded output distortions that arise at equilibrium reflect whether principals have conflicting or congruent objectives. Remarkably, these distortions can be understood by observing that an equilibrium allocation must maximize the “virtual aggregate” surplus of principals which leads to Lindahl-Samuelson conditions over a restricted range of allocations. A rich pattern of activity sets that reflects the congruence or conflict between principals may emerge at equilibrium and we illustrated those facts by means of examples.

Other specific settings of economic interest (trade, regulation, multi-unit auctions, common representation on retailer markets, etc.) that have already been deeply studied by means of the common agency methodology in a world of complete information would benefit from a serious consideration of agency problems using our methodology. Inefficient representations of interest groups in trade negotiations, endogenous limited entry in auction contexts, exclusive dealing agreements on retailing markets are, among others, interesting and important economic issues that can be explained by a careful study of the activity sets of some principals in specific contexts.

Beyond, our paper also suggests a few alleys for more theoretical works. First, the techniques developed in this paper might also be useful in private common agency games as well, i.e., when principals rely on different screening variables to control the agent. The ultimate objectives of such investigation should be to describe patterns of market coverage where either exclusive or multiple purchases endogenously arise on different subset of the type space. Such investigation is particularly important since students of market competition have generally found hard to reconcile market data with existing models of competition in nonlinear prices.\(^{24}\)

Second, some applied settings may require to develop a framework where principals share some common screening devices but keep others private. For instance, one may think of specific games between competing manufacturers dealing with the same retailers and contracting on some commonly observed price downstream but keeping their sales of intermediary goods secret. These settings lie somewhere in between the case of public delegated agency games and the case of private agency games. Extending our methodology to semi-public environments may be important.

We plan to investigate such extensions in future research.

REFERENCES


\(^{24}\)Such models either impose exclusivity clauses (Stole, 1995, Rochet and Stole, 2002) or restrict the analysis to cases where a buyer jointly buys from different sellers (Ivaldi and Martimort (1994), Martimort and Stole (2009a), Calzolari and De Nico (2010)). Stole (2007) surveys this literature.


APPENDIX A: PROOFS OF MAIN RESULTS

**Proof of Lemma 1:**

An allocation is implementable if and only if it is incentive compatible and individually rational.

Because the agent’s preferences are bilinear in and , it follows from Rochet (1987) and Milgrom and Segal (2002, Theorem 2) that incentive compatibility is equivalent to the requirement that is convex, absolutely continuous and its subdifferential contains for every . Thus, (2.3)- (2.2). Individual rationality with respect to principal i’s contract offer is equivalent to the conditions that and , which is equivalent to (2.1). Q.E.D.

**Proof of Theorem 1:**

**Step 1: The relaxed program.** Consider the relaxed program, , that ignores the convexity constraint (2.3):

\[
\text{(P}_i^r\text{)}: \max_{(U \in AC,q \in \mathcal{Q}^n)} \int \left( S_i(q(\theta)) + S_0(q(\theta)) + \sum_{j \neq i} t_j(q(\theta)) - \theta q(\theta) - U(\theta) \right) f(\theta) d\theta
\]

subject to (2.2), (2.1).

First, we rewrite this program using a change of variables in order to get it into a more useful format for applying a result from non-smooth control. Specifically, define the net utility that principal i’s contract provides to the agent: . It follows that, a.e., . We will use as the state variable in the restated optimal control problem, and as the control variable. Because is data to this given program, is effectively the control variable of principal i. Now we can state principal i’s relaxed program in net payoffs as

\[
\max_{(\Delta_i \in AC,q \in \mathcal{Q}^n)} \int \left( S_i(q(\theta)) - S_i(q_{-i}(\theta)) + S_0(q(\theta)) + \sum_{j \neq i} t_j(q(\theta)) - \theta q(\theta) - U_{-i}(\theta) - \Delta_i(\theta) \right) f(\theta) d\theta
\]
subject to \( q(\theta) - q_{-i}(\theta) \in -\partial \Delta_i(\theta), \Delta_i(\theta) \geq 0 \).

Second, we apply Theorem B.1 (which is stated and proved in Appendix B) and conclude that for any given profile of transfers \( t_{-i} \) offered by rival principals the rent-output profile \((\bar{U}, \bar{q})\) is a solution to \((P_i^\tau)\) if and only if \((\bar{U}, \bar{q})\) satisfies (2.2), (2.1) and there exists a probability measure \( \mu_i \) defined over the Borel subsets of \( \Theta \) with an associated adjoint function, \( \bar{M}_i : \Theta \rightarrow [0,1] \), defined by \( \bar{M}_i(\theta) = 0 \) and for \( \theta > \theta \),

\[
\bar{M}_i(\theta) = \int_{[\theta, \theta']} \mu_i(d\theta),
\]

such that the following two conditions are satisfied:

(A1) \( \text{supp} \{\mu_i\} \subseteq \{\theta : U(\theta) = U_{-i}(\theta)\} \),

(A2) \( \bar{q}(\theta) \in \arg \max_{q \in \mathcal{Q}} S_i(q) + S_0(q) + \sum_{j \neq i} t_j(q) + \left( \frac{\bar{M}_i(\theta) - F(\theta)}{f(\theta)} - \theta \right) q \), a.e.

**Step 2: Convexity of the “relaxed” solution:** To demonstrate that the solution \( \bar{U} \) to the relaxed program is convex it is sufficient to show that there is a non-increasing selection within the best-response correspondence defined by (A2). If so, the conditions of the “relaxed” Theorem apply to the more constrained program obtained by appending (2.3) to (2.2) and (2.1).

For any open interval in \( \Theta_i \), the adjoint function is constant, and therefore \( \bar{q} \) is nonincreasing given our generalized monotone hazard-rate condition. For any open interval in the complement, \( \Theta_i^c \), \( \bar{U}(\theta) = U_{-i}(\theta) \), and therefore \( \bar{q}(\theta) = q_{-i}(\theta) \). Because the latter is a selection from a nonincreasing correspondence, \( -\partial U_{-i}(\theta) \), it follows that \( \bar{q} \) is nonincreasing as well.

What remains is to establish that, for any boundary point, \( \bar{q} \) does not jump upward. This can only arise if \( \bar{M}_i \) jumps upward. The construction of \( \bar{M}_i \), however, only allows for upward jumps at boundary points when moving from an unconstrained region to a constrained region. Suppose an upward jump arises at such a point, \( \hat{\theta} \). Then there is a neighborhood \( (\hat{\theta}, \hat{\theta} + \epsilon) \) for which \( \bar{U}(\theta) = U_{-i}(\theta) \) and an adjacent neighborhood \( (\hat{\theta} - \epsilon, \hat{\theta}) \) such that \( \bar{U}(\theta) > U_{-i}(\theta) \). Because \( U_{-i} \) is convex and weakly below \( \bar{U} \), it must minorize \( \bar{U} \) at \( \hat{\theta} \). This implies that \( \bar{U} \) is locally convex at \( \hat{\theta} \), which in turn implies that \( \bar{q} \) cannot jump upwards. It follows that any solution that is non-increasing on \( \Theta_i \) and int(\( \Theta_i^c \)) must be non-increasing over all \( \Theta \). Hence, \( \bar{q} \) is a solution to the relaxed program that is non-increasing.

*Q.E.D.*

**Proof of Theorem 2:** The condition in (4.2) holds for almost every \( \theta \in \Theta \) and for any \( i \in N \). Thus, fixing \( \theta \), we may add up all maximands in (4.2) for \( i \in N \). From this aggregation, it follows that \( \bar{q}(\theta) \) also maximizes the maximand in (5.1), almost everywhere.

We obtain (5.2) by summing only over active principals, \( i \in \bar{\alpha}(\theta) \), expanding terms and noting that \( \sum_{i \in \bar{\alpha}(\theta)} \sum_{j \not\in \bar{\alpha}(\theta)} \bar{t}_i(q) = (|\bar{\alpha}(\theta)| - 1) T(q) + \sum_{j \not\in \bar{\alpha}(\theta)} \bar{t}_j(q) \), we get:

\[
\bar{q}(\theta) \in \arg \max_{q \in \mathcal{Q}} S_0(q) - \theta q + \sum_{i \in \bar{\alpha}(\theta)} S_i(q) + \sum_{i \not\in \bar{\alpha}(\theta)} \bar{t}_i(q) + \sum_{i \in \bar{\alpha}(\theta)} \left( \frac{\bar{M}_i(\theta) - F(\theta)}{f(\theta)} - \theta \right) q
\]

\[
+ (|\bar{\alpha}(\theta)| - 1) \left( S_0(q) - \theta q + T(q) \right), \text{ a.e.}
\]

Because \( \sum_{i \not\in \bar{\alpha}(\theta)} \bar{t}_i(\bar{q}(\theta)) = 0 \) by definition and \( \sum_{i \not\in \bar{\alpha}(\theta)} \bar{t}_i(q) \geq 0 \), we get the result. *Q.E.D.*
Proof of Lemma 2:
We show that any adjoint distribution profile which supports \((\bar{q}, \bar{U})\) can be replaced by the distributions defined in (6.1) without affecting the requirements in Theorem 1, conditions (4.1)-(4.2). We present the proof for \(i \in A\); the case for \(i \in B\) proceeds accordingly.

- Step 1. Define \(\hat{\theta}_i\) as the unique solution to \(s_i f(\hat{\theta}_i) = F(\hat{\theta}_i)\). \(\hat{\theta}_i\) is well defined given the monotone hazard rate condition. Two properties are immediately implied for the region \((\hat{\theta}_i, \bar{\theta})\). First,

\[
(\hat{\theta}_i, \bar{\theta}) = \{ \theta \mid F(\theta) - s_i f(\theta) > 0 \}.
\]

The slope of \(F(\theta) - s_i f(\theta)\) is positive if \(f(\theta) > s_i f'(\theta)\); because the monotone hazard rate condition requires \(f'(\theta)/f(\theta) \leq f(\theta)/F(\theta)\), it follows that \(F - s_i f\) is increasing if \(F(\theta)/f(\theta) > s_i\). We conclude a second property of \((\hat{\theta}_i, \bar{\theta})\) is that \(F(\theta) - s_i f(\theta)\) is strictly increasing on this interval.

- Step 2: \(\bar{\theta}_i = [\theta_0, \theta_0]\) where \(\theta_0 \leq \hat{\theta}_i\).

Suppose otherwise that on \([\theta_0, \theta_1] \subseteq \text{int} \Theta\) we have \(\bar{U}(\theta) = \bar{U}_{-i}(\theta)\), but for \(\varepsilon > 0\) sufficiently small we \(\bar{U}(\theta) > \bar{U}_{-i}(\theta)\) on the adjacent neighborhoods, \(\theta \in (\theta_0 - \varepsilon, \theta_0) \cup (\theta_1, \theta_1 + \varepsilon)\). If there are multiple regions satisfying these properties, then let \([\theta_0, \theta_1]\) be the leftmost such region.

Because \(\bar{U}(\theta) > \bar{U}_{-i}(\theta)\) on \((\theta_1, \theta_1 + \varepsilon)\) and both rent functions are continuous and non-increasing, it must be that \(\bar{q}(\theta) < \bar{q}_{-i}(\theta)\) on this region for \(\varepsilon\) sufficiently small. For this inequality to be satisfied, (4.2) requires that \(\bar{M}_i(\theta) < F(\theta) - s_i f(\theta)\) for all \(\theta \in (\theta_1, \theta_1 + \varepsilon)\). Because Theorem 1 implies that \(\bar{M}_i(\theta)\) is constant on \((\theta_1, \theta_1 + \varepsilon)\) and equal to \(\bar{M}_i(\theta_1)\), we have also \(\bar{M}_i(\theta) < F(\theta) - s_i f(\theta)\). Because \(\bar{M}_i(\theta) \geq 0\), it follows that \(F(\theta) - s_i f(\theta) > 0\) on this interval which implies \(\theta_1 \geq \hat{\theta}_i\). Because \(F - s_i f\) is increasing for all \(\theta > \hat{\theta}_i\), we can also conclude that \(s_i f(\theta) + \bar{M}_i(\theta_1) - F(\theta) \leq s_i f(\theta_1) + \bar{M}_i(\theta_1) - F(\theta_1) < 0\) for all \(\theta \in (\theta_1, \bar{\theta})\).

Suppose now that the participation constraint is binding on a second interval \([\theta_2, \theta_3]\) (possibly reduced to a point) with \(\varepsilon\) small enough so that \(\theta_1 + \varepsilon < \theta_2 - \varepsilon\). On the interval \((\theta_1, \theta_2)\), the fact that the participation constraint remains slack means \(\bar{M}_i(\theta) = \bar{M}_i(\theta_1)\) on that interval. Because the participation constraint binds at \(\theta_2\), it must be that \(\bar{q}(\theta) > \bar{q}_{-i}(\theta)\) on \((\theta_2 - \varepsilon, \theta_2)\) which would mean \(s_i f(\theta) + \bar{M}_i(\theta_1) - F(\theta) > 0\) on that interval. A contradiction. Thus, there is at most one region of binding participation, \([\theta_0, \theta_1]\).

Suppose that \(\bar{\theta}_i^c = [\theta_0, \theta_1]\) is the optimal region of binding participation. Note that principal \(i\) is implementing \(\bar{q}(\theta) < \bar{q}_{-i}(\theta)\) in the region \((\hat{\theta}_i, \bar{\theta})\). The principal could, instead, choose an alternative tariff \(\tilde{\ell}_i\) such that

\[
\tilde{\ell}_i(q) = \begin{cases} \ell_i(q) & \text{if } q \in \bar{q}([\theta_0, \theta_0]), \\ 0 & \text{if } q \notin \bar{q}([\theta_0, \bar{\theta}]). \end{cases}
\]

This would implement \(\bar{q}(\theta)\) on \([\theta, \theta_1]\), as before, but it would implement higher activity, \(\bar{q}_{-i}(\theta) > \bar{q}(\theta)\), at a lower cost on \((\theta_1, \bar{\theta})\) because \(s_i > 0\). Hence, we conclude that \(\theta_1 = \bar{\theta}\) and therefore \(\bar{\theta}_i^c = [\theta_0, \bar{\theta}]\).

Lastly, note that \(\bar{U}(\theta) \geq \bar{U}_{-i}(\theta)\) on \((\theta_0 - \varepsilon, \theta_0)\) implies (using (4.2)) that \(\bar{M}_i(\theta) > F(\theta) - s_i f(\theta)\). Because of Step 2, there is only one region with a binding participation.
constraint, \( \overline{M}_i(\theta) = 0 \) on \([\hat{\theta}_i, \theta_0]\), it follows that \( \theta_0 \leq \hat{\theta}_i \).

- Step 3. Characterizing \( \overline{M}_i \) over the (possibly empty) interval \((\theta_0, \hat{\theta}_i)\). Suppose that \( \theta_0 < \hat{\theta}_i \).

Because \( \overline{M}_i \geq 0 \) and \( \theta < \hat{\theta}_i \), over the interval, we have

\[
s_i + \frac{\overline{M}_i(\theta) - F(\theta)}{f(\theta)} > 0.
\]

An implication of (4.2) is thus that \( \overline{q}(\theta) \geq \overline{q}_{-i}(\theta) \). By supposition, however, over the interval \((\theta_0, \hat{\theta}_i)\) we have \( \overline{q}(\theta) = \overline{q}_{-i}(\theta) \). The following sequence of inequalities thus holds:

\[
\overline{q}_{-i}(\theta) = \overline{q}(\theta) \leq \arg \max_{q \in Q} S_0(q) - \theta q + \sum_{j \neq i} l_j(q) + \left( s_i + \frac{\overline{M}_i(\theta) - F(\theta)}{f(\theta)}\right) q
\]

\[
\geq \arg \max_{q \in Q} S_0(q) - \theta q + \sum_{j \neq i} l_j(q) + \left( s_i - \frac{F(\theta)}{f(\theta)}\right) q
\]

\[
\geq \arg \max_{q \in Q} S_0(q) - \theta q + \sum_{j \neq i} l_j(q) = \overline{q}_{-i}(\theta),
\]

where the first inequality follows from \( \overline{M}_i(\theta) \geq 0 \), the second from the fact that \( s_i \geq \frac{F(\theta)}{f(\theta)} \) on \((\theta_0, \hat{\theta}_i)\) and the last equality is by definition. Thus, the inequalities must be equalities and

\[
\overline{q}_{-i}(\theta) \in \arg \max_{q \in Q} S_0(q) - \theta q + \sum_{j \neq i} l_j(q) + s_i q - \frac{F(\theta)}{f(\theta)} q.
\]

Hence, setting \( \overline{M}_i(\theta) = 0 \) for \((\theta_0, \hat{\theta}_i)\) will suffice to implement \( \overline{q}_{-i}(\theta) \) as required.

- Step 4. We can now given a complete construction of \( \overline{M}_i \) over \( \Theta \) that implements the original equilibrium allocation for principal \( i \) in accord with (4.2). Set \( \overline{M}_i(\theta) = 0 \) for all \( \theta \leq \hat{\theta}_i \). For the region for \( \theta > \hat{\theta}_i \), it suffices to set \( \overline{M}_i(\theta) = F(\theta) - s_i f(\theta) \) in order to implement \( \overline{q}_{-i}(\theta) \). (Note, if \( \overline{q}_{-i} \) exhibits pooling, there is a family of adjoint functions, including the one above, that will satisfy the requirements of (4.2).)

- Step 5. What remains to check is (4.1). Note that the above construction of \( \overline{M}_i \) results in a continuous, nondecreasing function with range contained in \([0, 1]\); and \( \overline{M}_i \) is strictly increasing over \((\hat{\theta}_i, \overline{\theta})\). Hence, there exits a probability measure \( \overline{\mu}_i \) which generates \( \overline{M}_i \) and which has its support contained in \( \overline{\Theta}^c \), i.e., \( [\hat{\theta}_i, \overline{\theta}] \subseteq \overline{\Theta}_i \), as required. (Note that if \( \overline{M}_i(\overline{\theta}) < 1 \), then we construct \( \overline{\mu}_i \) with an atom of mass at \( \overline{\theta} \) equal to \( 1 - \overline{M}_i(\overline{\theta}) \) in order to satisfy the requirement that \( \overline{\mu}_i \) is a probability measure.)

\[ Q.E.D. \]

**Proof of Theorem 3:**

**Necessity.** Lemma 2 in tandem with Theorem 2, implies that, for almost every \( \theta \), the allocation satisfies

\[
(A3) \quad \overline{q}(\theta) \in \arg \max_{q \in Q} S_0(q) + (\beta(\theta) - \theta)q + (n - 1)(S_0(q) - \theta q + T(q)),
\]
where $\mathcal{T}$ implements $(\overline{q}, \mathcal{U})$. Given that the above program seeks to maximize an upper-semicontinuous function over a compact set, it has a solution. Define the value function of this program by

$$V(\theta) = \max_{q \in \mathcal{Q}} S_0(q) + (\beta(\theta) - \theta)q + (n - 1)(S_0(q) - \theta q + \overline{T}(q)).$$

Again, from Milfrom and Segal (2002, Theorem 2) and the fact that the maximand above is absolutely continuous as a function of $\theta$ (the derivative of $\beta(\theta)$ exists almost everywhere), it follows that $V(\theta)$ is absolutely continuous. Moreover, given $(\overline{q}, \mathcal{U})$ is an incentive-compatible allocation which solves this program,

$$V(\theta) = S_0(\overline{q}(\theta)) + (\beta(\theta) - \theta)\overline{q}(\theta) + (n - 1)\overline{U}(\theta).$$

Because $V$ is absolutely continuous, it is almost everywhere differentiable and for any pair $(\theta, \theta')$,

$$V(\theta) - V(\theta') = \int_{\theta'}^{\theta} (\beta'(x) - n)q(x)dx.$$

From the proof of Lemma 1, we know that $\mathcal{U}$ is also absolutely continuous and therefore for any pair $(\theta, \theta')$ we have

$$\overline{U}(\theta) - \overline{U}(\theta') = -\int_{\theta'}^{\theta} q(x)dx.$$

It follows that $S_0(\overline{q}(\theta)) + (\beta(\theta) - \theta)\overline{q}(\theta)$ is itself absolutely continuous in $\theta$ and

$$S_0(\overline{q}(\theta)) + (\beta(\theta) - \theta)\overline{q}(\theta) - [S_0(\overline{q}(\theta')) + (\beta(\theta') - \theta')\overline{q}(\theta')] = V(\theta) - V(\theta') - (n - 1) \left[ \overline{U}(\theta) - \overline{U}(\theta') \right]$$

or more simply

$$S_0(\overline{q}(\theta)) + (\beta(\theta) - \theta)\overline{q}(\theta) - [S_0(\overline{q}(\theta')) + (\beta(\theta') - \theta')\overline{q}(\theta')] = \int_{\theta'}^{\theta} (\beta'(x) - 1)q(x)dx.$$  \hfill (A4)

Using the relationship

$$((\beta(\theta) - \theta) - (\beta(\theta') - \theta'))\overline{q}(\theta') = \int_{\theta'}^{\theta} (\beta'(x) - 1)\overline{q}(\theta')dx,$$

and the fact that $\beta$ and $\overline{q}$ are both weakly non-increasing, we obtain:

$$S_0(\overline{q}(\theta)) + (\beta(\theta) - \theta)\overline{q}(\theta) - [S_0(\overline{q}(\theta')) + (\beta(\theta') - \theta)\overline{q}(\theta')]$$

$$= \int_{\theta'}^{\theta} (\beta'(x) - 1)(q(x) - q(\theta'))dx \geq 0.$$}

Because any $q' \in \overline{q}(\Theta)$ can be identified with some $\theta' \in \Theta$, the inequality implies $\overline{q}(\theta)$ satisfies (6.2) pointwise in $\theta$.

**SUFFICIENCY.** Suppose that $\overline{q}^2$ satisfies (6.3). Such profile exists by continuity of $S_0$ and compactness of $\mathcal{Q}$. Because $\beta(\theta) - \theta$ is decreasing, $\overline{q}^2$ is nonincreasing. Define the inverse of $\overline{q}^2$, as the correspondence $\overline{\mathcal{F}}(q) = \min\{\theta | q = \overline{q}^2(\theta)\}, \max\{\theta | q = \overline{q}^2(\theta)\}$. Because $\overline{q}^2$ is nonincreasing, this correspondence is monotone and almost everywhere single valued. Abusing notations, we will use $\overline{\mathcal{F}}(q)$ as an arbitrary nonincreasing selection from this correspondence when integrating.

We construct the individual tariffs of each principal $i \in N$ as follows:

$$\overline{T}_i^2(q) = \int_{\overline{\mathcal{F}}(q_i)} \beta_i(\overline{\mathcal{F}}(x))dx.$$
Note that $t_i^\mathcal{Q}$ is nonnegative by construction and $\bar{t}_i^\mathcal{Q}(q) = 0$ for $q = \bar{q}(\hat{\theta}_i)$. Also, since $\beta_i$ is a nonincreasing function and $\bar{q}(q)$ is nonincreasing, each constructed tariff is convex by construction (because $\beta_i$ and $\bar{q}(q)$ are nonincreasing) Denote the aggregates by $T_i^\mathcal{Q} = \sum_{i \in N} t_i^\mathcal{Q}(q)$ and $\mathcal{T}_{-i}^\mathcal{Q} = \sum_{j \neq i} t_j^\mathcal{Q}(q)$. It follows that the aggregates are also convex. What remains to be shown is (i) the aggregate transfer $\mathcal{T}$ induces the agent to choose $\bar{q}$, and (ii) each principal $i$, facing the rivals’ aggregate $\mathcal{T}_{-i}$, finds it optimal to implement $\bar{q}$.

**Incentive compatibility.** Consider the agent’s problem when facing aggregate payment, $\mathcal{T}^\mathcal{Q}$. For any pair $(\theta, q)$, the following conditions hold:

\[
S_0(\bar{q}(\theta)) + \mathcal{T}^\mathcal{Q}(\bar{q}(\theta)) + (\beta(\theta) - \theta)\bar{q}(\theta) \geq S_0(q) + \mathcal{T}^\mathcal{Q}(\bar{q}(\theta)) + (\beta(\theta) - \theta)q
\]

\[
\geq S_0(q) + \mathcal{T}^\mathcal{Q}(q) + \beta(\bar{q}(q))(\bar{q}(\theta) - q) + (\beta(\theta) - \theta)q
\]

where the first inequality follows from the definition of $\bar{q}(\theta)$ and the second uses the convexity of $\mathcal{T}^\mathcal{Q}$. Simplifying further, we obtain

\[
S_0(\bar{q}(\theta)) + \mathcal{T}^\mathcal{Q}(\bar{q}(\theta)) - \theta\bar{q}(\theta) \geq S_0(q) + \mathcal{T}^\mathcal{Q}(q) - \theta q + \left[(\beta(\bar{q}(q)) - \beta(\theta))(\bar{q}(\theta) - q)\right].
\]

Because $\beta(\bar{q}(q))$ is non-increasing in $q$, the bracketed difference is always non-negative. Incentive compatibility is implied, as desired.

**Principals’ optimality.** Consider principal $i$’s program in light of Theorem 1 and Lemma 2. $\bar{q}^\mathcal{Q}$ is an optimal allocation for principal $i$ if and only if

\[
(A5) \quad \bar{q}^\mathcal{Q}(\theta) \in \arg \max_{\bar{q} \in \mathcal{Q}} S_0(q) + \mathcal{T}_{-i}^\mathcal{Q}(q) - \bar{t}_i^\mathcal{Q}(q) + (\beta_i(\theta) - \theta)q, \quad \text{a.e.}
\]

Remember that each tariff $\bar{t}_i^\mathcal{Q}$ is convex and therefore $\mathcal{T}_{-i}^\mathcal{Q}$ is convex. Now observe that for all pairs $(\theta, q)$, the following sequence of relationships holds:

\[
S_0(\bar{q}(\theta)) + \mathcal{T}_{-i}^\mathcal{Q}(\bar{q}(\theta)) + (\beta_i(\theta) - \theta)\bar{q}(\theta)
\]

\[
= S_0(\bar{q}(\theta)) + \mathcal{T}_{-i}^\mathcal{Q}(\bar{q}(\theta)) + (\beta(\theta) - \beta_i(\theta) - \theta)\bar{q}(\theta)
\]

\[
\geq S_0(q) + (\beta(\theta) - \theta)q + \mathcal{T}_{-i}^\mathcal{Q}(\bar{q}(\theta)) - \beta_i(\theta)\bar{q}(\theta)
\]

\[
\geq S_0(q) + (\beta(\theta) - \beta_i(\theta) - \theta)q + \mathcal{T}_{-i}^\mathcal{Q}(q) + [((\beta_i(\bar{q}(q)) - \beta_i(\theta))(\bar{q}(\theta) - q)]
\]

\[
= S_0(q) + (\beta_i(\theta) - \theta)q + \mathcal{T}_{-i}^\mathcal{Q}(q) + [((\beta_i(\bar{q}(q)) - \beta_i(\theta))(\bar{q}(\theta) - q)]
\]

\[
\geq S_0(q) + (\beta_i(\theta) - \theta)q + \mathcal{T}_{-i}^\mathcal{Q}(q).
\]

Both of the equalities above follow from the definition of $\beta_i$. The first inequality uses the fact that $\bar{q}^\mathcal{Q}(\theta)$ solves (6.3), while the second inequality follows from the convexity of $\mathcal{T}_{-i}^\mathcal{Q}$. The final inequality follows from the fact that, $\beta(\bar{q}(q))$ is nonincreasing in $q$, and therefore the bracketed difference is always non-negative. This proves that (A5) holds and that principal $i$ desires to implement $\bar{q}^\mathcal{Q}$ when facing a rival aggregate of $\mathcal{T}_{-i}^\mathcal{Q}$. Because $\bar{t}_i^\mathcal{Q}$ is zero for $x_i^\mathcal{Q}(\hat{\theta}_i)$, the constructed tariff $\bar{t}_i^\mathcal{Q}$ is the least-cost (nonnegative) transfer that accomplishes this end.  

*Q.E.D.*
Proof of Corollary 1: Observe first that the Corollary is trivially true when int $\overline{q}(\Theta) = \emptyset$. Suppose thus int $\overline{q}(\Theta) \neq \emptyset$. If $\overline{q}$ is an equilibrium allocation in a linear common-agency game, then at any $\theta$ such that $\overline{q}(\theta) \in$ int $\overline{q}(\Theta)$, (6.2) implies that: 

$$0 \in \partial^* S_0(\overline{q}(\theta)) + \beta(\theta) - \theta$$

(where $\partial^* f = - \partial (-f)$ for a concave function $f$) Hence, $\overline{q}(\theta)$ is a local maximizer of $S_0(q) + (\beta(\theta) - \theta)q$ and thus a global maximizer over the broader domain $Q$ since $S_0$ is concave. Q.E.D.

Proof of Corollary 2: Because $\overline{q}$ is weakly decreasing, it is almost everywhere differentiable and any point of discontinuity $\theta_0$ which is interior is isolated. Moreover, $\overline{q}$ being weakly decreasing, the righthand side limit as $\theta \to \theta_0^-$, say $q_2 = \overline{q}(\theta_0^-) = \lim_{\theta \to \theta_0^-} \overline{q}(\theta) >$ does exist. Similarly, the lefthand side limit $q_1 = \overline{q}(\theta_0^+) = \lim_{\theta \to \theta_0^+} \overline{q}(\theta)$ also exists. Moreover, at such a discontinuity, $q_2 > q_1$. From the proof of Theorem 3, we know that $S_0(\overline{q}(\theta)) + (\beta(\theta) - \theta)\overline{q}(\theta)$ is (absolutely) continuous. Therefore, right- and left-hand side limits of that expression at $\theta_0$ are equal which yields (6.4).

Q.E.D.

Proof of Theorem 4: Because $\hat{q} \in$ int $\overline{q}^Q(\Theta)$, there exists $\theta_0$ such that $\hat{q}$

\[ \hat{q} = \arg \max_{q \in \mathcal{Q}} S_0(q) + (\beta(\theta_0) - \theta_0)q \]

where the above equality follows from strict concavity of $S_0$. By continuity of $S_0$, one can choose $q_1$ and $q_2$ sufficiently close to $\hat{q}$ such that $\hat{q} \in (q_1, q_2) \subset$ int $\overline{q}^Q(\Theta)$ and simultaneously satisfy

\[ S_0(q_1) + (\beta(\theta_0) - \theta_0)q_1 = S_0(q_2) + (\beta(\theta_0) - \theta_0)q_2. \]

Let $\theta_1$ and $\theta_2$ be types such that $\overline{q}^Q(\theta_1) = q_1$ and $\overline{q}^Q(\theta_1) = q_2$. The solution to program (6.5) then becomes:

\[ \overline{q}(\theta) = \begin{cases} 
\overline{q}^Q(\theta) & \theta \in [\theta_1, \theta_2) \\
q_2 & \theta \in [\theta_2, \theta_0) \\
q_1 & \theta \in (\theta_0, \theta_1] \\
\overline{q}^Q(\theta) & \theta \in (\theta_1, \theta_2]. 
\end{cases} \]

The proof now proceeds by constructing equilibrium tariffs which implement this output profile. Let $(\overline{t}_{i1}^Q, \ldots, \overline{t}_{in}^Q)$ be the maximal equilibrium transfers which implement the allocation $(\overline{q}^Q, \overline{T}^Q)$. By assumption, all principals are active at $\hat{q}$ in the maximal equilibrium, and therefore $\overline{t}_{i}^Q(\hat{q}) > 0$ for all $i$ and $\overline{T}^Q(\hat{q}) > 0$. We now construct tariffs in the new equilibrium as follows. For $i \in \mathcal{A}$,

\[ \overline{t}_i(q) = \begin{cases} 
\overline{t}_i^Q(q) & \text{for } q \leq q_1 \\
0 & \text{for } q \in (q_1, q_2) \\
\overline{t}_i^Q(q) + \overline{t}_i^b & \text{for } q \geq q_2, 
\end{cases} \]

and for $i \in \mathcal{B}$,

\[ \overline{t}_i(q) = \begin{cases} 
\overline{t}_i^Q(q) + \overline{t}_i^b & \text{for } q \leq q_1 \\
0 & \text{for } q \in (q_1, q_2) \\
\overline{t}_i^Q(q) & \text{for } q \geq q_2. 
\end{cases} \]
The constants \( t^a_i \) and \( t^b_i \) remain to be determined. For now, we only require that \( \sum_{i \in A} t^a_i = 0 \) and \( \sum_{i \in B} t^b_i = 0 \). These conditions imply that the aggregate payment is unchanged when moving from the maximal equilibrium to the new one except for the gap \((q_1, q_2)\):

(A11) \[
T(q) = \begin{cases} 
T^Q(q) & q \in \mathcal{Q} \setminus (q_1, q_2) \\
0 & q \in (q_1, q_2).
\end{cases}
\]

**Incentive compatibility.** We first establish that the constructed aggregate tariff \( \hat{T} \) will implement \( \overline{q} \) defined in (A8). Note that

\[
\bar{T}(q_2) - \bar{T}(q_1) = T^Q(q_2) - T^Q(q_1) = \int_{q_1}^{q_2} \beta(q) dq,
\]

where \( \beta(q) \) is the inverse correspondence of \( \beta(q) \), which is single-valued for almost every \( q \in \mathcal{Q} \). Given that \( T(q) \leq T^Q(q) \) with equality for \( q \leq q_1 \) and \( q \geq q_2 \), and given that it was incentive compatible under \( T^Q \) for types \( \theta \in \Theta \setminus (\theta_2, \theta_1) \) to choose \( \beta(q) \), it remains incentive compatible for them to do so under \( T \). What remains is to establish incentive compatibility within the gap \((\theta_2, \theta_1)\).

First, we establish that no agent type \( \theta \in (\theta_2, \theta_1) \) will prefer a point \( q \in (q_1, q_2) \) to any other choice given the aggregate tariff, \( \bar{T} \). This requires that

\[
\max_{q \in \mathcal{T}(\mathcal{Q}) \setminus (q_1, q_2)} S_0(q) - \theta_0 + \bar{T}(q) \geq \sup_{q \in (q_1, q_2)} S_0(q) - \theta q \quad \forall \theta \in (\theta_2, \theta_1).
\]

Taking into account (A11), the continuity of the maximand in \( q \) on the right-hand side, and the fact that this condition certainly holds at the boundaries \( \theta_1 \) and \( \theta_2 \) from the argument above, this latter condition becomes:

\[
\max_{q \in \mathcal{T}(\mathcal{Q}) \setminus (q_1, q_2)} S_0(q) - \theta_0 + \bar{T}^Q(q) \geq \max_{q \in [q_1, q_2]} S_0(q) - \theta q \quad \forall \theta \in [\theta_2, \theta_1].
\]

This condition holds if

(A12) \[
\min_{\theta \in [\theta_2, \theta_1]} \max_{q \in \mathcal{T}(\mathcal{Q}) \setminus (q_1, q_2)} S_0(q) - \theta_0 + \bar{T}^Q(q) \geq \max_{q \in [q_1, q_2], \theta \in [\theta_2, \theta_1]} S_0(q) - \theta q.
\]

Fixing \( \theta \) and thus \( \theta_0 \) from (A6), define the size of the gap by \( \varepsilon = q_2 - q_1 \). Because \( S_0 \) is strictly concave, (6.5) uniquely defines \( q_1, q_2, \theta_1 \) and \( \theta_2 \) as continuous functions of \( \varepsilon \) for \( \varepsilon \) small enough. Denote each side of (A12) as a continuous function of \( \varepsilon \), say \( L(\varepsilon) \) and \( R(\varepsilon) \) respectively. In the limit for \( \varepsilon = 0 \) we have necessarily \( \hat{q} = q_1 = q_2 \) and \( \theta_0 = \theta_1 = \theta_2 \). Because all principals are active at \( \hat{q} \) for the maximal equilibrium, we have:

\[
L(0) = U^Q(\hat{\theta}_0) = S_0(\hat{q}) - \theta_0 \hat{q} + T^Q(\hat{q}) > S_0(q) - \theta_0 \hat{q} = U_0(\theta_0) = R(0).
\]

By continuity, \( L(\varepsilon) \geq R(\varepsilon) \) also holds for \( \varepsilon \) small enough in which case (A12) holds.

We must also show that the marginal agent type that is indifferent between these choices corresponds to \( \theta_0 \). To this end, notice that from our aggregate tariff construction, \( \bar{U}(\theta_1) = U^Q(\theta_1) \) and \( \bar{U}(\theta_2) = \bar{U}^Q(\theta_2) \). From these boundary conditions, we deduce

\[
\bar{U}^Q(\theta_2) - \bar{U}^Q(\theta_1) = \int_{q_1}^{\theta_2} \bar{U}(q) dq = \int_{\theta_2}^{\theta_1} \bar{U}(q) dq = \bar{U}(\theta_2) - \bar{U}(\theta_1) = (\theta_0 - \theta_2)q_2 + (\theta_1 - \theta_0)q_1.
\]

From incentive compatibility at \( \theta_1 \) and \( \theta_2 \), we therefore have

\[
S_0(q_2) + \bar{T}(q_2) - \theta_2 q_2 - S_0(q_1) - \bar{T}(q_1) + \theta_1 q_1 = \bar{U}(\theta_2) - \bar{U}(\theta_1) = (\theta_0 - \theta_2)q_2 + (\theta_1 - \theta_0)q_1,
\]

\[
= (\theta_0 - \theta_2)q_2 + (\theta_1 - \theta_0)q_1.
\]
or more simply

$$S_0(q_2) + \overline{T}(q_2) - \theta_0 q_2 = S_0(q_1) + \overline{T}(q_1) - \theta_0 q_1.$$  

Therefore type $\theta_0$ is indifferent between $q_1$ and $q_2$ under the aggregate tariff $\overline{T}$ as desired.

**Principals optimality.** We lastly establish that for appropriately chosen constants, $t_i^0$ and $t_i^1$, the constructed transfers induce each principal to optimally implement $\overline{q}$ defined in (A8). Using Theorem 1 and Lemma 2, it is sufficient to show for each $i \in N$ and for almost all $\theta$ that

$$\overline{q}(\theta) \in \arg \max_{q \in \mathcal{Q}} S_0(q) + T_{-i}(q) + (\beta_i(\theta) - \theta) q.$$  

For types $\theta \notin (\theta_2, \theta_1)$, the fact that $\overline{T}_{-i}(q) \leq \overline{T}_{-i}(\overline{q})$ with equality for $q \in \mathcal{Q} \setminus (q_1, q_2)$ implies that $\overline{q}(\theta) = \overline{q}^\mathcal{Q}(\theta)$, as required. Consider $\theta \in (\theta_2, \theta_1)$. First, we establish that principal $i$ chooses to implement either $q_1$ or $q_2$ for these types when $\varepsilon$ is sufficiently small; i.e., principal $i$ does not want to implement an allocation within the gap, $q \in (q_1, q_2)$, for any type $\theta \in (\theta_2, \theta_1)$. Formally, this will be true if and only if

$$\max_{q \in \mathcal{Q}(\theta) \setminus (q_1, q_2)} S_0(q) + (\beta_i(\theta) - \theta) q + \overline{T}_{-i}(q) \geq \sup_{q \in (q_1, q_2)} S_0(q) + (\beta_i(\theta) - \theta) q \quad \forall \theta \in (\theta_2, \theta_1).$$

Proceeding analogously as in the case of agent optimality above, this condition holds for $\varepsilon$ small enough if

$$S_0(\hat{q}) + (\beta_i(\theta_0) - \theta_0) \hat{q} + \overline{T}_{-i}(\hat{q}) > S_0(\hat{q}) + (\beta_i(\theta_0) - \theta_0) \hat{q}.$$  

But this latter condition holds trivially when all principals are active at $\hat{q}$ in the maximal equilibrium as assumed.

What remains to be shown is that if principal $i$ can only choose between $q_1$ and $q_2$ for types $\theta \in (\theta_2, \theta_1)$, then the marginal type for which principal $i$ is indifferent coincides with $\theta_0$. We establish this by choosing the transfers constants, $t_i^0$ and $t_i^1$, while maintaining the nonnegativity of each $\overline{t}_i$ transfer function. We first consider two special cases separately, from which we later are able to generalize: (1) congruent preferences (i.e., $A = N$ or $B = N$) for arbitrary $n$, and (2) opposed preferences for $n = 2$.

**(Case 1.) Congruent preferences.** Without loss of generality, consider the case in which $A = N$. We need to determine $t_i^0$ for each principal’s tariff. Let $\bar{\theta}_i$ be the type which makes principal $i$ indifferent between implementing $q_1$ and $q_2$. It is defined by

$$S_0(q_1) + (\beta_i(\bar{\theta}_i) - \bar{\theta}_i) q_1 + \overline{T}_{-i}(q_1) = S_0(q_2) + (\beta_i(\bar{\theta}_i) - \bar{\theta}_i) q_2 + \overline{T}_{-i}(q_2),$$

which can be written

$$\frac{S_0(q_2) - S_0(q_1)}{q_2 - q_1} + \frac{T(q_2) - T(q_1)}{q_2 - q_1} - \frac{\overline{t}_i(q_2) - \overline{t}_i(q_1)}{q_2 - q_1} = \bar{\theta}_i - \beta_i(\bar{\theta}_i).$$

The incentive compatibility condition (A13), combined with $T^\mathcal{Q}(q_1) = T(q_1)$ and $T^\mathcal{Q}(q_2) = \overline{T}(q_2)$, implies that

$$\beta_i(\bar{\theta}_i) = \bar{\theta}_i - \theta_0 + \frac{\overline{t}_i(q_2) - \overline{t}_i(q_1)}{q_2 - q_1}.$$

To have $\bar{\theta}_i = \theta_0$, we need

$$\frac{\overline{t}_i(q_2) - \overline{t}_i(q_1)}{q_2 - q_1} = \beta_i(\theta_0).$$
For \( i \in \mathcal{A} \), this requires that \( t_i^a \) satisfies

\[
\frac{\bar{T}_i^Q(q_2) - \bar{T}_i^Q(q_1)}{q_2 - q_1} + \frac{t_i^a}{q_2 - q_1} = \beta_i(\theta_0)
\]

or

(A16) \( t_i^a = \beta_i(\theta_0)(q_2 - q_1) - \int_{q_1}^{q_2} \beta_i(\bar{T}_i^Q(q))dq \).

Conditions (A7) and (A13) imply that

\[
\frac{T_i^Q(q_2) - T_i^Q(q_1)}{q_2 - q_1} = \beta(\theta_0).
\]

But by definition, we have also:

\[
\frac{T_i^Q(q_2) - T_i^Q(q_1)}{q_2 - q_1} = \frac{1}{q_2 - q_1} \int_{q_1}^{q_2} \beta(\bar{T}_i^Q(q))dq.
\]

Hence, we are assured that the constants \( (t_1^a, \ldots, t_n^a) \) chosen in (A16) will also satisfy \( \sum_{i \in \mathcal{A}} t_i^a = 0 \) and \( \sum_{i \in \mathcal{B}} t_i^b = 0 \), as required.

Lastly, we need to establish that each tariff \( \bar{t}_i \) remains nonnegative. When \( \epsilon \) is small enough, all constants \( t_i^a \) as defined by (A16) are small enough, so that \( \bar{t}_i \) defined by (A9) remains nonnegative.

(Case 2.) Conflicting preferences, \( n = 2 \). We now consider \( n = 2 \) with principal \( 1 \in \mathcal{A} \) and \( 2 \in \mathcal{B} \). We also additionally assume that the agent’s type is distributed uniformly. This implies that each \( \beta_i(\theta) \) is linear in a sufficiently small neighborhood \((\theta_1, \theta_2)\) corresponding to \( q_1 = \bar{T}_i^Q(\theta_1) \) and \( q_2 = \bar{T}_i^Q(\theta_2) \).

As in the case of congruent principals above, principal \( i \) will implement \( \bar{q} \) if and only if, again (A15) holds. Because of the requirements that \( \sum_{i \in \mathcal{A}} t_i^a = \sum_{i \in \mathcal{B}} t_i^b = 0 \), with \( n = 2 \) we have immediately that \( t_1^a = t_2^b = 0 \). Thus, (A15) requires that

\[
\frac{\bar{T}_i^Q(q_2) - \bar{T}_i^Q(q_1)}{q_2 - q_1} = \beta_i(\theta_0),
\]

which using the definition of \( \bar{T}_i^Q \) can be further simplified to

\[
\int_{q_1}^{q_2} \beta_i(\bar{T}_i^Q(q))dq = \beta_i(\bar{\theta}_0)
\]

for each principal \( i = 1, 2 \). By assumption, each \( \beta_i \) is linear in \( \theta \) around \( \theta_0 \), and these requirements reduce to a single equation relating \( q_1, q_2 \) and \( \theta_0 \):

\[
\int_{q_1}^{q_2} \bar{T}_i^Q(q)dq = \theta_0.
\]

We know, however, that

\[
\frac{T_i^Q(q_2) - T_i^Q(q_1)}{q_2 - q_1} = \int_{q_1}^{q_2} \beta(\bar{T}_i^Q(q))dq = \beta(\theta_0),
\]

which implies the requirement, again by linearity. Thus, we are assured that each principal’s indifferent type \( \tilde{\theta}_i = \theta_0 \) for any \( \epsilon \) small enough so that the gap \((q_1, q_2)\) satisfies (A7).
General preference profiles. Given our results for congruent preferences with arbitrary \( n \) and for general preferences with \( n = 2 \), we are able to extend our analysis to the general setting in two steps. Suppose that there \( |A| = n_a \) and \( |B| = n - n_a = n_b \). Define the aggregate transfer by principals on both sides by respectively \( \overline{T}_A = \sum_{i \in A} t_i \) and \( \overline{T}_B = \sum_{i \in B} t_i \). Similar definitions apply to \( \beta_A = \sum_{i \in A} \beta_i \) and \( \beta_B = \sum_{i \in B} \beta_i \). As in the case of congruent principals above, principals \( i \in A \) will implement \( \overline{q} \) if and only if
\[
\frac{\overline{T}_A(q_2) - \overline{T}_A(q_1)}{q_2 - q_1} = \beta_A(\theta_0).
\]
The same applies for principals \( i \in B \):
\[
\frac{\overline{T}_B(q_2) - \overline{T}_B(q_1)}{q_2 - q_1} = \beta_B(\theta_0).
\]
By construction,
\[
\frac{\overline{T}_A(q_2) - \overline{T}_A(q_1)}{q_2 - q_1} = \overline{T}_A(q_2) - \overline{T}_A(q_1) = \frac{1}{q_2 - q_1} \int_{q_1}^{q_2} \beta_A(\overline{q}^Q(q))dq
\]
and
\[
\frac{\overline{T}_B(q_2) - \overline{T}_B(q_1)}{q_2 - q_1} = \overline{T}_B(q_2) - \overline{T}_B(q_1) = \frac{1}{q_2 - q_1} \int_{q_1}^{q_2} \beta_B(\overline{q}^Q(q))dq.
\]
But again the linearity of the \( \beta_i \) over a domain where all principals are active at the maximal equilibrium (which is implied by the uniform distribution) implies that \( \beta_A(\theta_0) = \frac{1}{q_2 - q_1} \int_{q_1}^{q_2} \beta_A(\overline{q}^Q(q))dq \) and \( \beta_B(\theta_0) = \frac{1}{q_2 - q_1} \int_{q_1}^{q_2} \beta_B(\overline{q}^Q(q))dq \) both hold when \( \beta(\theta_0) = \frac{1}{q_2 - q_1} \int_{q_1}^{q_2} \beta(\overline{q}^Q(q))dq \) as requested by our construction. The requirements that \( \sum_{i \in A} t_i^a = \sum_{i \in B} t_i^b = 0 \) are finally satisfied with the constructions:
\[
(A17) \quad t_i^a = t_i^b = \beta_i(\theta_0)(q_2 - q_1) - \int_{q_1}^{q_2} \beta_i(\overline{q}^Q(q))dq.
\]
When \( \varepsilon \) is small enough, all constants \( t_i^a \) and \( t_i^b \) as defined by \( (A17) \) are small enough, so that the \( t_i \)s defined by \( (A9) \) and \( (A10) \) remain nonnegative.

Agent’s preference for equilibria. Lastly, note that \( \overline{T}^Q(q) \geq \overline{T}(q) \) for all \( q \in \Omega \), and so \( \overline{U}^Q(\theta) \geq \overline{U}(\theta) \) for all \( \theta \in \Theta \).

**Proof of Corollary 3:** Let \( \theta_0 \) be a point of discontinuity of \( \overline{q} \). Such point is isolated because \( \overline{q} \) is non-increasing and thus almost everywhere differentiable. Moreover \( \overline{q} \) admits right- and left-hand side limits at \( \theta_0 \), denoted respectively by \( \overline{q}(\theta_0^-) \) and \( \overline{q}(\theta_0^+) \) with \( \overline{q} \) being continuous and differentiable both on a right- and a left-neighborhoods of \( \theta_0 \). We also deduce from monotonicity that \( \overline{q}(\theta_0^-) > \overline{q}(\theta_0^+) \) by incentive compatibility. The optimality conditions (6.2) at \( \theta_0 \) imply that (6.4) must hold. Because \( S_0 \) is strictly concave, \( S_0(q) + (\beta(\theta_0) - \theta_0)q \) has a unique maximum at \( \overline{q}^Q(\theta_0) \), and we thus have \( \overline{q}(\theta_0^-) > \overline{q}^Q(\theta_0) > \overline{q}(\theta_0^+) \).

The argument in the necessity proof of Theorem 3 implies that
\[
S_0(\overline{q}(\theta)) + (\beta(\theta) - \theta)\overline{q}(\theta) = V(\theta) + (n - 1)\overline{U}(\theta)
\]
is the sum of two absolutely continuous functions, thus absolutely continuous itself and a.e. differentiable. Using \( (A4) \), the differentiability of \( S_0 \), the following condition holds at any point of differentiability of \( \overline{q} \):
\[
(A18) \quad \overline{q}(\theta)\left(S_0'(\overline{q}(\theta)) + \beta(\theta) - \theta\right) = 0.
\]
From this, it follows that \( \tilde{q}(\theta) = 0 \) whenever \( \overline{q}(\theta) \neq \overline{Q}(\theta) \) at a point of differentiability.

Using (A18) on the right- and a left-neighborhoods of \( \theta_0 \), we deduce that \( \tilde{q}(\theta) = 0 \) on such neighborhoods. By assumption, \( \overline{q}(\Theta) \subset \overline{Q}(\Theta) \). Therefore, there exist \( \theta_1 \) and \( \theta_2 \) such that \( \theta_2 < \theta_0 < \theta_1 \) and \( \overline{q}(\theta_0) = \overline{Q}(\theta_2) \) and \( \overline{q}(\theta_0) = \overline{Q}(\theta_1) \). Because the allocation \( \overline{q} \) must be non-decreasing, it can only be constant on the whole intervals \( [\theta_2, \theta_0) \) and \( (\theta_0, \theta_1] \). \( Q.E.D. \)

**Proof of Proposition 4:** We first remind the expressions of rents and payments in the maximal equilibrium when \( s_1 = -s_2 = 1 < 2\delta \) (which ensures that both \( \theta_1 \) and \( \theta_2 \) are interior) and the distribution is uniform on \( \Theta = [-\delta, \delta] \) with \( Q = [-1 - \delta, 1 + \delta] \). From Proposition 3, we know that, on the interval \([-\min(1 - \delta, \delta), \min(1 - \delta, \delta)]\) that contains \( \theta_0 = 0 \), the maximal equilibrium policy is given by \( \overline{Q}(\theta) = -3\theta \) for \( \theta \in [-1 + \delta, 1 - \delta] \). The individual equilibrium schedules and the aggregate payment are respectively

\[
\tilde{t}_1^Q(q) = \tilde{t}_2^Q(-q) = \begin{cases} 
0 & \text{for } q \leq -3(1 - \delta), \\
\frac{9}{4}(1 - \delta)^2 + \frac{1}{2}(1 - \delta - 2\delta)^2 & \text{for } q \in [-3(1 - \delta), 3(1 - \delta)] \\
\frac{9}{4}(1 - \delta)^2 + \frac{1}{2}(1 - \delta - 2\delta)^2 & \text{for } q \in [3(1 - \delta), 1 + \delta]
\end{cases}
\]

and \( \overline{T}^Q(q) = \overline{t}_1^Q(q) + \overline{t}_2^Q(q) \), while the agent’s rent writes as

\[
U^Q(\theta) = \begin{cases} 
\frac{9}{4}(1 - \delta)^2 + \frac{1}{2}(1 - \delta - 2\delta)^2 & \text{for } \theta \in [-\delta, -1 + \delta] \\
3(1 - \delta)^2 + \frac{3}{2}\theta^2 & \text{for } \theta \in [-1 + \delta, 1 - \delta] \\
\frac{9}{4}(1 - \delta)^2 + \frac{1}{2}(1 - \delta + 2\delta)^2 & \text{for } \theta \in [1 - \delta, \delta].
\end{cases}
\]

We now construct an equilibrium with a discontinuity at \( \theta_0 = 0 \) so that the discontinuity gap \( [-q_0, q_0] \) remains in \( \overline{Q}([-1 + \delta, 1 - \delta]) \), i.e., on an area where principals’ activity sets overlap in the maximal equilibrium which implies \( q_0 \leq 3(1 - \delta) \). In particular, we have \( \overline{T}^Q(q_0) = \overline{T}^Q(-q_0) = \frac{1}{6}(q_0 + 3(1 - \delta))^2 + \frac{1}{6}(-q_0 + 3(1 - \delta))^2 = \frac{q^2_0}{3} + 3(1 - \delta)^2 \). Following the proof of Theorem 4 (case 2), the so-constructed discontinuous equilibrium preserves aggregate and individual payments beyond the discontinuity gap:

\[
\overline{T}(q) = \begin{cases} 
0 & \text{for } q \in (-q_0, q_0) \\
\overline{T}^Q(q) & \text{for } q \geq q_0 \text{ and } q \leq q_0.
\end{cases}
\]

This yields the following expression of the agent’s rent in the discontinuous equilibrium:

\[
(A19) \quad U(\theta) = \min \left\{ U^Q(\theta), -q_0 - \frac{q_0^2}{2} + \overline{T}^Q(q_0), \theta q_0 - \frac{q_0^2}{2} + \overline{T}^Q(-q_0) \right\}.
\]

Following notations in the proof of Corollary 3, we denote \( \theta_2 = -\theta_1 = -\frac{q_0}{2} \). To find out the maximal value of the \( q_0 \) that can be sustained, we again closely follow the proof of Theorem 4.

The first condition to be checked is (A12) that rewrites in this specific context as:

\[
(A20) \quad U(\theta) = \max_{q \in \overline{Q}(\theta) \backslash (-\theta_0, \theta_0)} -\theta q - \frac{q^2}{2} + \overline{T}^Q(q) \geq \sup_{q \in (-\theta_0, \theta_0)} -\theta q - \frac{q^2}{2} + \frac{\theta^2}{2} \quad \forall \theta \in [\theta_2, \theta_1].
\]

Using (A19) to express the lefthand side and symmetry of the rent profile in \( \theta \) around the origin, this condition holds when \( U(\theta) = \theta q_0 - \frac{q_0^2}{2} + 3(1 - \delta)^2 \geq \frac{q_0^2}{2} \) for all \( \theta \in [0, \theta_1] \) which is always true if it holds at \( \theta = 0 \), i.e., \( U(0) = -\frac{q_0^2}{6} + 3(1 - \delta)^2 > 0 \) but this latter inequality is always true for all \( q_0 \leq 3(1 - \delta) \).
The second condition to be checked is \((A14)\) for each principal. Taking into account symmetry, it suffices to verify that this condition holds for principal 1 which gives:

\[
(A21) \quad \max_{q \in \mathcal{Q}(\cdot)(1-q_0, q_0)} -\frac{q^2}{2} + (1 - \delta - 2\theta) q + \mathcal{I}_Q^\infty(q) \geq \sup_{q \in (-q_0, q_0)} -\frac{q^2}{2} + (1 - \delta - 2\theta) q \quad \forall \theta \in [\theta_2, \theta_1].
\]

When \(q_0 \leq 3(1 - \delta)\), the max on the lefthand side is achieved either at \(-q_0\) (for \(\theta \in [0, \theta_1]\)) or at \(q_0\) (for \(\theta \in [\theta_2, 0]\)). Again using symmetry, we focus on the case \(\theta \in [0, \theta_1]\) and note that the sup on the righthand side can be rewritten so that \((A21)\) becomes:

\[
(A22) \quad \frac{3}{2} (1 - \delta)^2 - \frac{q_0^2}{3} \geq \mathcal{R}(\theta) = -2\theta q_0 + \max_{q \in (-q_0, q_0)} -\frac{q^2}{2} + (1 - \delta - 2\theta) q \quad \forall \theta \in [0, \theta_1].
\]

Because the maximum of linear functions of \(\theta\) is convex, \(\mathcal{R}\) is also convex in \(\theta\). Using the envelope theorem to evaluate the derivative of this max, it is immediate that \(\mathcal{R}\) is also decreasing. Hence, the condition always holds when it holds at \(\theta = 0\). We compute

\[
\mathcal{R}(0) = \begin{cases} 
\frac{(1-\delta)^2}{3} & \text{if } q_0 \in [1 - \delta, 3(1 - \delta)], \\
-\frac{q_0^2}{2} + (1 - \delta) q_0 & \text{if } q_0 \in [0, 1 - \delta].
\end{cases}
\]

Hence, \((A22)\) holds when \(\frac{3}{2} (1 - \delta)^2 - \frac{q_0^2}{3} \geq \mathcal{R}(0)\) which is true when \(q_0 \leq \sqrt{3}(1 - \delta)\).

\textit{Welfare comparison.} Fix \(q_0 \in [0, \sqrt{3}(1 - \delta)]\) (the case \(q_0 = 0\) corresponding to the maximal equilibrium). We know from Theorem 4 that the agent always prefers the maximal equilibrium to any discontinuous equilibrium keeping aggregate payments the same outside the discontinuity gap. Turning now overall expected payoff of the principals in a discontinuous equilibrium, we observe that, because of opposite interests, this expected payoff is the opposite of their overall expected payment. This expected payment writes as:

\[
\mathcal{T}(q_0) = \frac{1}{2\delta} \left( \int_{-\delta}^{q_0/3} \mathcal{I}_Q^\infty(q \mathcal{Q}(\theta)) d\theta + \int_{-\delta}^{q_0/3} \mathcal{I}_Q^\infty(q_0) d\theta + \int_{0}^{q_0/3} \mathcal{I}_Q^\infty(-q_0) d\theta + \int_{q_0/3}^{\delta} \mathcal{I}_Q^\infty(q_0 \mathcal{Q}(\theta)) d\theta \right).
\]

Observe that:

\[
\frac{dT}{dq_0}(q_0) = \frac{1}{2\delta} \left( \int_{0}^{q_0/3} \frac{d}{dq_0}(\mathcal{I}_Q^\infty(q_0)) d\theta + \int_{0}^{q_0/3} \frac{d}{dq_0}(\mathcal{I}_Q^\infty(-q_0)) d\theta \right) = \frac{2q_0}{9\delta}.
\]

Henceforth, \(\mathcal{T}(q_0)\) is convex for \(q_0 \geq 0\) and minimized at \(q_0 = 0\), i.e., the maximal equilibrium is also preferred by the principals. Since both the principals and the agent prefers the maximal equilibrium, welfare is higher at that equilibrium.

\section*{APPENDIX B: NECESSARY AND SUFFICIENT CONDITIONS FOR NON-SMOOTH OPTIMAL CONTROL WITH A LINEAR STATE VARIABLE}

We are interested in the following pure-state control program \((P)\):

\[
\text{Maximize } \Lambda(x) \equiv \int_{0}^{1} (S(\theta, u(\theta)) - x(\theta)f(\theta)) d\theta
\]

subject to \(x \in AC(\Theta, \mathbb{R}), \ x(\theta) = u(\theta), \ x(\theta) \geq 0 \text{ for all } \theta \in \Theta \equiv [0, 1].\)

The constraints require that the state variable \(x\) is a non-negative, absolutely continuous function, \(x \in AC(\Theta, \mathbb{R})\). \(x\) is said admissible if it satisfies these constraints. Note that the integrand \(L(\theta, x, u) = (S(\theta, u) - x)f(\theta)\) is linear in \(x\) and that the state constraint, \(x \geq 0\),
is independent of $\theta$. These two restrictions within the class of state-constrained, non-smooth optimal control problems are the source of many sharp results in the analysis that follows. Note also that we allow the principal’s surplus to depend directly on $\theta$ which is more general than our assumption in the text but this generality is only imposed here to make this appendix of independent interest for readers willing to apply those techniques in other specific contexts.

We assume that $S(\theta, \cdot)$ is an upper-semi continuous function bounded from above and that $f(\theta)$ is a positive and bounded from above function so that $F(\theta) \equiv \int_{[\theta, \theta]} f(\theta)$ is absolutely continuous. Without loss of generality, we normalize $f$ such that $F(1) = 1$ and interpret $F$ as a continuous probability distribution. Lastly, we assume that $S(\cdot, \cdot)$ is $L \times B$-measurable, where $L$ denotes the set of Lebesgue measurable subsets of $\Theta$ and $B$ is the set of Borel measurable subsets of $\mathbb{R}$. Importantly, we do not assume a priori that $S(\theta, \cdot)$ is a continuous function. We present our main result for this class of problems.

**Theorem B.1** $\pi$ is a solution to program $(P)$ if and only if $\pi$ is admissible and there exists a probability measure $\mu$ defined over the Borel subsets of $\Theta$ with an associated adjoint function, $\bar{M}: \Theta \to [0, 1]$, defined by $\bar{M}(\theta) = 0$ and for $\theta > \theta$,

$$\bar{M}(\theta) = \int_{[\theta, \theta]} \mu(ds),$$

such that the following two conditions are satisfied:

(B1) $\text{supp}\{\mu\} \subseteq \{\theta | \pi(\theta) = 0\}$,

(B2) $\bar{x}(\theta) \in \arg\max_{v \in \mathbb{R}} S(\theta, v) + (F(\theta) - \bar{M}(\theta))v$, for a.e. $\theta \in \Theta$.

Furthermore, if $y(\theta, \sigma) \equiv \arg\max_{v \in \mathbb{R}} S(\theta, v) + (F(\theta) - \sigma)v$ is single-valued and continuous over the domain $(\theta, \sigma) \in \Theta \times [0, 1]$, then the solution $\pi$ to $(P)$ is continuously differentiable.

**Remarks:**
- Theorem 1 is very similar to Theorem 1 in Julien (2000). In both theorems, necessary and sufficient conditions are stated in terms of a probability measure which serves to express a “complementary slackness condition” (B1) and an optimality condition (B2). Moreover, both theorems use a similar condition to establish the continuity of $\bar{x}(\theta)$ in the solution to $(P)$. Julien’s Theorem, however, uses the stronger hypothesis that $S(\cdot, \cdot)$ is twice continuously differentiable. Our technical contribution is to weaken these hypotheses to requirements of upper semi-continuity. This generalization allows us to apply the necessary and sufficient conditions above to our class of agency games with upper-semi continuous contract menus.
- The condition that $y(\theta, \sigma)$ is single-valued and continuous is implied by the strict concavity of $S(\theta, \cdot)$. It is also implied by the weaker condition in Julien (2000, Assumption 2) that $S(\theta, v) - (\sigma - F(\theta))v$ is strictly quasi-concave in $v$ for any $\sigma \in [0, 1]$.
- The adjoint function $\bar{M}(\theta)$. Note in particular that the function $\bar{M}$ is constructed to be left-continuous rather than right-continuous.

**Proof of Theorem B.1:**

**Overview.** We prove necessity by specializing Theorem 3 from Vinter and Zheng (1998), exploiting fact that our integrand in $\Lambda$ is a linear function of $x$ and that the state constraint
$x(\theta) \geq 0$ is linear and independent of $\theta$. Sufficiency is proven by generalizing Arrow’s Sufficiency Theorem to non-smooth optimal control problems and specializing the theorem to the case in which the objective integrand is a linear function of $x$. The regularity of the optimal solution follows from arguments involving the necessary conditions. While the proof seems straightforward when viewed from this broad vantage point, a considerable investment in concepts and notation from non-smooth, non-convex analysis is required along the way.

Preliminaries for Non-Smooth Analysis. We first introduce some additional notation. We draw heavily from Vinter and Zheng (1998) in the following presentation.\footnote{A complete treatment can be found in the monograph of Vinter (2000). Theorem 3 from Vinter and Zheng (1998) appears as Theorem 10.2.1 in Vinter (2000).}

Take a closed set $A \subseteq \mathbb{R}^k$ and a point $x \in A$. A vector $p \in \mathbb{R}^k$ is a limiting normal to $A$ at $x$ if there exists a sequence $(x_i, p_i) \rightarrow (x, p)$ and a $K \geq 0$ such that for each $i$ in the sequence $p_i \cdot |x_i - x| \leq K |x_i - x|^2$. The cone of limiting normal vectors to $A$ at $x$ is denoted $N_A(x)$. Given a lower semi-continuous function $g : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ and a point $x \in \mathbb{R}^k$ such that $g(x) < +\infty$, the limiting subdifferential of $g$ at $x$ is defined as

$$\partial g(x) \equiv \{ \xi \mid (\xi, -1) \in N_{\text{epi}(g)}(x, g(x)) \},$$

where $\text{epi}(g)$ is the epigraph of the function $g$ defined as

$$\text{epi}(g) \equiv \{ (x, \alpha) \in \mathbb{R}^k \times \mathbb{R} \mid \alpha \geq g(x) \}.$$

The asymptotic limiting subdifferential of $g$ at $x$, written $\partial^\infty g(x)$, is defined as

$$\partial^\infty g(x) \equiv \{ \xi \mid (\xi, 0) \in N_{\text{epi}(g)}(x, g(x)) \}.$$

Two results from nonsmooth analysis (e.g., Vinter (2000), Propositions 4.3.3 and 4.3.4) that we use are (1) $\partial^\infty g(x) = \{0\}$ if $g$ is Lipschitz continuous and (2) for any $x$ such that $g(x)$ is finite,

$$N_{\text{epi}(g)}(x, g(x)) = \{ (\xi d, -\xi) \mid \xi > 0, d \in \partial g(x) \} \cup \{ \partial^\infty g(x) \times \{0\} \}.$$

We denote the Euclidean norm in $\mathbb{R}^k$ by $|\cdot|$, and denote the norm on the space of absolutely continuous functions by

$$||x|| \equiv |x(\theta)| + \int_{\Theta} |\dot{x}(\theta)|d\theta.$$

A local maximizer of $\Lambda(x)$ is a feasible arc, $\overline{x}$, which maximizes $\Lambda(x)$ over all feasible arcs $x \in AC(\Theta, \mathbb{R}_+)$ within an $\varepsilon$ neighborhood of $\overline{x}$,

$$||\overline{x} - x|| \leq \varepsilon.$$

A local minimizer is defined analogously.

Necessity. For completeness, we state Theorem 3 of Vinter and Zheng (1998) which provides necessary conditions for solutions to the following minimization program:

$$(P') : \quad \text{Minimize } J(x) = \int_\Theta g(\theta, x(\theta), \dot{x}(\theta))d\theta$$

subject to $x \in AC(\Theta, \mathbb{R})$ and $h(\theta, x(\theta)) \leq 0$ for all $\theta \in \Theta \equiv [\theta, \bar{\theta}]$.\footnote{We specialize their theorem to our present problem in which the range of $x(\theta)$ is one-dimensional and there is no endpoint cost function.}

**Theorem B.2** (Vinter and Zheng (1998), Theorem 3) Let $\overline{x}$ be a AC local minimizer for $(P')$ such that $J(\overline{x}) < +\infty$. Assume that the following hypotheses are satisfied:
H1. \( L(\cdot, x, \cdot) \) is \( \mathcal{L} \times \mathcal{B} \) measurable for each \( x \) and \( L(\theta, \cdot, \cdot) \) is lower semi-continuous for a.e. \( \theta \in \Theta \).

H2. For every \( K > 0 \) there exists \( \delta > 0 \) and \( k \in L^1 \) such that
\[
|L(\theta, x', v) - L(\theta, x, v)| \leq k(\theta)|x' - x|, \quad L(\theta, \varpi(\theta), v) \geq -k(\theta)
\]
for a.e. \( \theta \in \Theta \), for all \( x, x' \in \varpi(\theta) + \delta B \) and \( v \in \varpi(\theta) + KB \), where \( B \) is a unit Euclidean ball.

H3. \( h \) is upper semi-continuous near \( (\theta, \varpi(\theta)) \) for all \( \theta \in \Theta \), and there exists a constant \( k_h \) such that
\[
|h(\theta, x') - h(\theta, x)| \leq k_h|x' - x|
\]
for all \( \theta \in \Theta \) and all \( x', x \in \varpi(\theta) + \delta B \).

Then there exist an arc \( p \in AC \), a constant \( \lambda \geq 0 \), a non-negative measure \( \mu \) on the Borel subsets of \( \Theta \) and a \( \mu \)-integrable function \( \gamma : \Theta \to \mathbb{R} \), such that

(i). \( \lambda + \max_{\theta \in \Theta} |p(\theta)| + \int_{\Theta} \mu(ds) = K > 0 \) (where \( K \) is an arbitrary normalization constant),\(^{27}\)

(ii).
\[
p(\theta) \in \text{co}\left\{ \eta \mid \eta(p(\theta)) + \int_{\Theta} \gamma(s)\mu(ds), -\lambda \right\} 
\]
\[
\in N_{\text{epi}(L(\theta, \cdot, \cdot))}(\varpi(\theta), \varpi(\theta), L(\theta, \varpi(\theta), \varpi(\theta)))\text{ a.e.,}
\]

(iii).
\[
p(\theta) = p(\varpi) - \int_{\Theta} \gamma(s)\mu(ds) = 0,
\]

(iv).
\[
\left(p(\theta) + \int_{\Theta} \gamma(s)\mu(ds)\right) \cdot \varpi(\theta) - \lambda L(\theta, \varpi(\theta), \varpi(\theta)) 
\]
\[
\geq \left(p(\theta) + \int_{\Theta} \gamma(s)\mu(ds)\right) \cdot v - \lambda L(\theta, \varpi(\theta), v)
\]
for all \( v \in \mathbb{R} \) a.e.,

(v). \( \gamma(\theta) \in \partial^\infty h(\theta, \varpi(\theta)) \) \( \mu \)-a.e. and \( \text{supp}\{\mu\} \subseteq \{ t \mid h(\theta, \varpi(\theta)) = 0 \} \), where
\[
\partial^\infty h(\theta, x) \equiv \text{co}\{\lim_i \xi_i \mid \exists t_i \to t, x_i \to x \text{ such that} \}
\]
\[
h(\theta, t_i, x_i) > 0 \text{ and } \xi_i \in \partial h(t_i, x_i) \text{ for all } i\}.
\]

\(^{27}\)We choose to state the Theorem using \( K > 0 \) as an arbitrary normalization rather than \( K = 1 \), which is the normalization chosen in Vinter and Zheng (1998). Later, by setting \( K = 3 \), we will succeed in normalizing \( \mu \) to a probability measure which is a more familiar object.
We apply this result to our setting by substituting \( x f(\theta) - S(\theta, v) \) in program (P) in place of \( L(\theta, x, v) \) and thereby converting the maximization functional \( \Lambda \) in program (P) to the minimization functional \( J \) in program (P'). We complete the transformation by requiring that \( h(\theta, x) = -x \), and that \( L(\theta, x, v) \) is a linear function of \( x \) for any \((\theta, v)\).

First, we verify that hypotheses III-III are satisfied for our program (P). Because \( S(\theta, \cdot) \) is upper semi-continuous and \( B \)-measurable, and because \( L(\theta, x, v) \) is linear in \( x \), H1 is satisfied. H2 requires that \( L(\theta, \cdot, v) \) is Lipschitz continuous, which is trivial given that \( L \) is linear in \( x \) with coefficient \( f(\theta) \). Because the transformed program has \( h(\theta, x) = -x \), \( h \) is a continuous linear functional of \( x \) and thus H3 is also satisfied.

Next, we specialize the conclusions of Vinter and Zheng (1998) my making use of the additional restrictions on \( L(\cdot) \) and \( h(\cdot) \). We present this in the following Lemma.

**Lemma 3** Suppose that \( L(\theta, x, v) \) is a linear function of \( x \) and that \( h(\theta, x) = -x \). Then the conclusions (i)-(v) of Theorem B.2 imply

(a). \( \lambda + \max_{\theta \in \Theta} |p(\theta)| + \int_{\Theta} \mu(\mu) = K > 0 \),

(b). \( \hat{p}(\theta) = \lambda f(\theta) \) a.e.,

(c). \( p(\theta) = p(\theta) + \int_{\Theta} \gamma(s) \mu(\mu) = 0 \)

(d). \( \tilde{\pi}(\theta) \in \arg \max_{v \in \mathbb{R}} \left( p(\theta) + \int_{[0,1]} \gamma(s) \mu(\mu) \right) \cdot v + \lambda S(\theta, v), \) a.e.,

(e). \( \gamma(\theta) = -1 \mu - a.e. \) and \( \text{supp} \{\mu\} \subseteq \{t \mid h(\theta, \pi(\theta)) = 0\} \).

**Proof of Lemma 3:** Implications (i) and (a) are identical. Implication (ii) requires almost everywhere that

\[
\hat{p}(\theta) \in \text{co} \left\{ \eta \mid \left( \eta, p(\theta) + \int_{[S,t]} \gamma(s) \mu(\mu), -\lambda \right) \in N_{\text{epi}(L(\theta, \cdot))} (\pi, \tilde{\pi}, L(\theta, \pi, \tilde{\pi})) \right\}.
\]

Because \( L(\theta, \pi(\theta), \tilde{\pi}(\theta)) = f(\theta) \pi(\theta) - S(\theta, \tilde{\pi}(\theta)) \) is finite, the limiting normal cone in the above expression can be written as

\[
N_{\text{epi}(L(\theta, \cdot))} (\pi, \tilde{\pi}, L) = \left\{ (\xi d_1, \xi d_2, -\xi) \mid \xi > 0, (d_1, d_2) \in \partial \left( f(\theta) \cdot \pi(\theta) - S(\theta, \tilde{\pi}(\theta)) \right) \right\} \cup \left\{ \partial^\infty \left( f(\theta) \cdot \pi(\theta) - S(\theta, \tilde{\pi}(\theta)) \right) \times \{0\} \right\}.
\]

Using the fact that \( L(\cdot) \) is additively separable in \( x \) and \( \tilde{x} \), a basic chain rule for lower semi-continuous functions (RW, Proposition 10.5) yields

\[
\partial \left( f(\theta) \pi(\theta) - S(\theta, \tilde{\pi}(\theta)) \right) = \partial \left( f(\theta) \pi(\theta) \right) \times \partial \left( -S(\theta, \tilde{\pi}(\theta)) \right) = \left\{ f(\theta) \times \partial \left( -S(\theta, \tilde{\pi}(\theta)) \right) \right\},
\]

and

\[
\partial^\infty \left( f(\theta) \pi(\theta) - S(\theta, \tilde{\pi}(\theta)) \right) \subseteq \partial^\infty \left( f(\theta) \pi(\theta) \right) \times \partial^\infty \left( -S(\theta, \tilde{\pi}(\theta)) \right) = \left\{ \{0\} \times \partial^\infty \left( -S(\theta, \tilde{\pi}(\theta)) \right) \right\},
\]


where the last equality uses the fact that a linear function is Lipschitz continuous and hence \( \partial^{\infty}(f(\theta)\bar{\pi}(\theta)) = \{0\} \). Substituting these subdifferentials into the expression for the limiting normal cone, we have a simple inclusion:

\[
N_{\text{epi}(L(\theta, \cdot, \cdot))}(x, \dot{x}, \bar{\pi}) \subseteq \{(\xi f(\theta), \xi d_2, -\xi) | \xi > 0, \; d_2 \in \partial \left(-S(\theta, \dot{x}(\theta)) \right)\} \cup \{(0) \times \partial^{\infty}(\bar{\pi}(\theta)) \times \{0\} \}.
\]

This simplifies yet again to the inclusion

\[
N_{\text{epi}(L(\theta, \cdot, \cdot))}(x, \dot{x}, \bar{\pi}) \subseteq \{(\xi f(\theta), \xi d_2, -\xi) | \xi \geq 0, \; d_2 \in \partial \left(-S(\theta, \dot{x}(\theta)) \right) \cup \partial^{\infty}(\bar{\pi}(\theta)) \}.
\]

The key point to note is that any vector in the limiting normal cone must point in the same direction in the \((x, \bar{\pi})\) plane, regardless of \(d_2\). Returning to implication (ii), we see that any point \(\eta\) in the given convex hull must satisfy \(\eta \cdot -\lambda = (\xi f(\theta), -\xi)\) for some \(\xi \geq 0\), and hence the convex hull reduces to \(\{\lambda f(\theta)\}\). We conclude that implication (ii) simplifies to implication (b) given that \(L(\cdot)\) is both additively separable and linear in \(x\).

Implication (iii) is identical to implication (c).

Using the transformation \(L(\theta, x, v) = xf(\theta) - S(\theta, v)\), implication (iv) simplifies to implication (d). Lastly, the fact that \(h(\theta, x) = -x\) yields \(\partial_x h(\theta, \bar{\pi}(\theta)) = \partial^2_x h(\theta, \bar{\pi}(\theta)) = \{-1\}\). Thus, implication (v) simplifies to \(\gamma(\theta) = -1\) \(\mu\)-a.e. and \(\text{supp}(\mu) \subseteq \{t | \bar{\pi}(\theta) = 0\}\). This is implication (e).

An immediate inspection of conditions (a)-(e) suggest further simplifications by combining these conditions. Conditions (b) and (c) jointly yield

\[
p(\theta) = \lambda F(\theta).
\]

Because \(p(\bar{\theta}) = \lambda\) and \(\gamma(\theta) = -1\) a.e. with respect to \(\mu\), condition (c) also implies

\[
\int_{\Theta} \mu(ds) = \lambda.
\]

Because we also have \(\max_{\theta \in \Theta} |p(\theta)| = \lambda\), condition (a) implies \(\lambda > 0\) and in particular \(\lambda = \frac{K}{\bar{x}}\). Because the choice of \(K\) is arbitrary, we choose \(K = 3\) as a normalization, yielding \(\lambda = 1\) and \(\int_{\Theta} \mu(ds) = 1\). Thus, the normalization makes \(\mu\) a probability measure on \(\Theta\). Defining \(\bar{M}(\theta) = \int_{[\theta, \cdot]} \mu(ds)\), the implication in (d) is therefore

\[
\bar{\pi}(\theta) \in \text{arg max}_{v \in \mathbb{R}} S(\theta, v) + (F(\theta) - \bar{M}(\theta)) v, \text{ a.e.},
\]

which is condition (B2) of Theorem B.1. Lastly, the implication of (e) delivers the complementary slackness condition (B1). We have therefore proven the necessity of the conditions in Theorem B.1.

**Sufficiency:** We adapt the argument of Arrow’s Sufficiency Theorem using the basic approach of Seierstad and Sydsaeter (1987) but relaxing their continuity and smoothness assumptions.

Let \(x\) be any admissible arc: \(x \in AC(\Theta, \mathbb{R})\) and \(x(\theta) \geq 0\) for all \(\theta \in \Theta\). Define

\[
\Delta = \int_{\Theta} \{(S(\theta, \dot{x}(\theta)) - \bar{x}(\theta)f(\theta)) - (S(\theta, \dot{x}(\theta)) - x(\theta)f(\theta))\} d\theta.
\]

We will demonstrate that, under conditions (B1) and (B2) of Theorem B.1, \(\Delta \geq 0\).
To this end, it is useful to define the Hamiltonian for program (P) using $M(\theta) - F(\theta)$ as the adjoint equation which satisfies conditions (B1) and (B2):

$$H(\theta, x, v) = S(\theta, v) - x \cdot f(\theta) - (\bar{M}(\theta) - F(\theta)) \cdot v.$$ 

Note that $\bar{M}(\theta)$ is defined for $\theta \in [\underline{\theta}, \bar{\theta}]$ and thus $H(\cdot)$ inherits the same domain. Nonetheless, because $\mu$ is not part of expression of $\Delta$ and $F$ is absolutely continuous, we can ignore the point $\underline{\theta}$ in the integral and conclude that

$$\Delta = \int_{[\underline{\theta}, \bar{\theta}]} (H(\theta, \bar{x}(\theta), \dot{x}(\theta)) - H(\theta, x(\theta), \dot{x}(\theta))) d\theta + \int_{\Theta} (F(\theta) - \bar{M}(\theta)) (\dot{x}(\theta) - \dot{\bar{x}}(\theta)) d\theta.$$ 

Define the optimized Hamiltonian as

$$\hat{H}(\theta, x) = \sup_{v \in \mathbb{R}} H(\theta, x, v).$$ 

Because $\bar{M}(\theta) - F(\theta)$ is bounded on $[\underline{\theta}, \bar{\theta}]$ and $S(\theta, \cdot)$ is bounded from above by assumption, we note that $\hat{H}(\cdot)$ must be finite. Condition (B2) implies that

$$\hat{H}(\theta, \bar{x}(\theta)) = H(\theta, \bar{x}(\theta), \dot{\bar{x}}(\theta))$$

and for any admissible $x \in AC(\Theta; \mathbb{R}_+),$

$$\hat{H}(\theta, x(\theta)) \geq H(\theta, x(\theta), \dot{x}(\theta)).$$

Combining these facts, we obtain

$$H(\theta, \bar{x}(\theta), \dot{\bar{x}}(\theta)) - H(\theta, x(\theta), \dot{x}(\theta)) \geq \hat{H}(\theta, \bar{x}(\theta)) - \hat{H}(\theta, x(\theta))$$

$$= f(\theta)(x(\theta) - \bar{x}(\theta)).$$

The last statement relies fundamentally on the linearity of $H(\cdot)$ in $x$. Substituting into the previous statement for $\Delta$, we have

$$\Delta \geq \int_{[\underline{\theta}, \bar{\theta}]} f(\theta)(x(\theta) - \bar{x}(\theta)) d\theta + \int_{\Theta} (F(\theta) - \bar{M}(\theta)) (\dot{x}(\theta) - \dot{\bar{x}}(\theta)) d\theta$$

$$= \int_{\Theta} f(\theta)(x(\theta) - \bar{x}(\theta)) + F(\theta) (\dot{x}(\theta) - \dot{\bar{x}}(\theta)) d\theta - \int_{[\underline{\theta}, \bar{\theta}]} \bar{M}(\theta) (\dot{x}(\theta) - \dot{\bar{x}}(\theta)) d\theta$$

$$= \int_{\Theta} \frac{d}{d\theta}[F(\theta)(x(\theta) - \bar{x}(\theta))] d\theta - \int_{[\underline{\theta}, \bar{\theta}]} \bar{M}(\theta) (\dot{x}(\theta) - \dot{\bar{x}}(\theta)) d\theta$$

$$= (x(1) - \bar{x}(1)) - \int_{[\underline{\theta}, \bar{\theta}]} \bar{M}(\theta) (\dot{x}(\theta) - \dot{\bar{x}}(\theta)) d\theta.$$

It follows that $\Delta \geq 0$ if

$$(x(1) - \bar{x}(1)) - \int_{[\underline{\theta}, \bar{\theta}]} \bar{M}(\theta) (\dot{x}(\theta) - \dot{\bar{x}}(\theta)) d\theta \geq 0.$$ 

If $\bar{M}$ were absolutely continuous, we would be able to integrate the second term by parts and reach such a conclusion. Because $\bar{M}$ is possibly discontinuous, we must proceed more carefully. Note that $\bar{M}$ is non-decreasing on $[\underline{\theta}, \bar{\theta}]$ with at most a countable number of upward jump discontinuities. Furthermore, $\bar{M}$ is absolutely continuous elsewhere, allowing us to integrate by parts between any pair of discontinuities. Also note that at any such upward jump point, $\tau$, $\bar{M}$ is left and right continuous with $\bar{M}(\tau) < \bar{M}(\tau^+)$ and (by condition (B1)) we have $\bar{x}(\tau^+) = 0.$
Denote the set of jump discontinuities by \{τ_1, τ_2, \ldots\}, a possibly infinite set. Let \( \mathcal{I} \) be the index set of \( τ_i \). Between any two points \( τ_i \) and \( τ_{i+1} \), we know

\[
\int_{(τ_i, τ_{i+1})} M(θ) (\dot{x}(θ) - \dot{\bar{x}}(θ)) \, dθ \\
= \left. M(θ)(x(θ) - \bar{x}(θ)) \right|_{τ_i}^{τ_{i+1}} - \int_{(τ_i, τ_{i+1})} (x(θ) - \bar{x}(θ)) \mu(θ) \, dθ \\
= M(τ_{i+1})(x(τ_{i+1}) - \bar{x}(τ_{i+1})) - M(τ_i)(x(τ_i) - \bar{x}(τ_i)) \\
- \int_{(τ_i, τ_{i+1})} (x(θ) - \bar{x}(θ)) \mu(θ) \, dθ.
\]

The second equality above uses the fact that \( x \) and \( \bar{x} \) are continuous on \( Θ \).

Define the size of the jump discontinuity at \( τ \) by \( d(τ) = M(τ^+) - M(τ) > 0 \). Then we may write

\[
\int_{[θ, θ]} M(θ) (\dot{x}(θ) - \dot{\bar{x}}(θ)) \, dθ \\
= \sum_{i \in \mathcal{I}} M(τ_{i+1})(x(τ_{i+1}) - \bar{x}(τ_{i+1})) - (d(τ_i) + M(τ_i))(x(τ_i) - \bar{x}(τ_i)) \\
- \int_{(τ_i, τ_{i+1})} (x(θ) - \bar{x}(θ)) \mu(θ) \, dθ \\
= (x(1) - \bar{x}(1)) - \sum_{i \in \mathcal{I}} d(τ_i)(x(τ_i) - \bar{x}(τ_i)) - \int_{(τ_i, τ_{i+1})} (x(θ) - \bar{x}(θ)) \mu(θ) \, dθ.
\]

By complementary slackness in condition (B1), we know \( \bar{x}(θ) \mu(θ) = 0 \) and at any jump point \( τ \) we must have \( \bar{x}(τ) = 0 \). Thus,

\[
\int_{[θ, θ]} M(θ) (\dot{x}(θ) - \dot{\bar{x}}(θ)) \, dθ = (x(1) - \bar{x}(1)) - \sum_{i \in \mathcal{I}} d(τ_i)x(τ_i) + \int_{(τ_i, τ_{i+1})} x(θ) \mu(θ) \, dθ.
\]

We deduce

\[
\Delta \geq (x(1) - \bar{x}(1)) - \int_{[θ, θ]} M(θ) (\dot{x}(θ) - \dot{\bar{x}}(θ)) \, dθ \\
= \sum_{i \in \mathcal{I}} d(τ_i)x(τ_i) + \int_{(τ_i, τ_{i+1})} x(θ) \mu(θ) \, dθ.
\]

Because \( x(θ) \geq 0, \mu \) is a non-negative measure, and jump discontinuities \( d(τ_i) \) are positive, we conclude \( \Delta \geq 0 \) as claimed. We have proven that conditions (B1) and (B2) are sufficient for a solution.

**Smoothness of the solution, \( \bar{x} \):** We add the hypothesis that

\[
y(θ, σ) = \arg \max_{v \in \mathbb{R}} S(θ, v) + (F(θ) - σ)v
\]

is single-valued and continuous for \( (θ, σ) \in Θ \times [0, 1] \). It follows that \( y(θ, σ) \) is non-increasing in \( σ \) and from condition (B2), that \( \dot{\bar{x}}(θ) = q(θ, M(θ)) \) a.e.

Suppose to the contrary that \( \dot{\bar{x}} \) is discontinuous at some point \( τ \in Θ \). Initially, suppose that Condition (B2) is extended to hold for all \( θ \in Θ \) rather than for a.e. \( θ \in [θ, θ_0] \); call this Condition (B2'). Condition (B2') and the additional hypothesis that \( y(θ, σ) \) is continuous in \( (θ, σ) \) jointly imply that \( \dot{\bar{x}}(θ) \) is discontinuous at \( τ \) only if \( M \) is also discontinuous at \( τ \). Any discontinuity in \( M \), however, must be an upward jump, \( d(τ) = M(τ^+) - M(τ) > 0 \), implying that \( \dot{\bar{x}}(θ) \) must jump downwards. Complementary slackness (condition (B1), however, imposes that \( \bar{x}(τ) = 0 \),
with the implication that a downward discontinuity at $\tau$ would violate the state constraint $x(\theta) \geq 0$ in the neighborhood to the immediate right of $\tau$. Hence, continuity must hold for all points $\theta \in [\theta, \theta)$ under Condition (B2'). Furthermore, because $\mathcal{M}$ is left continuous at $t = 1$, no jump in $\hat{x}(\theta)$ is possible at this endpoint. We conclude that Condition (B2') implies that $\hat{x}(\theta)$ is continuous for all $\theta \in \Theta$. The weaker Condition (B2) allows $\hat{x}(\theta)$ to violate the maximization condition on sets of measure zero, including at $\theta = \theta$. But such violations have no effect on the solution $\bar{x}$ which is absolutely continuous. Thus, $\bar{x}$ is smooth as posited. Q.E.D.