Aggregate Representations of Aggregate Games

Martimort, David and Stole, Lars

Paris School of Economics, University of Chicago Booth School of Business

23 June 2011

Working paper

AGGREGATE REPRESENTATIONS OF AGGREGATE GAMES[1]

June 23, 2011

David Martimort and Lars Stole

An aggregate game is a normal-form game with the property that each player's payoff is a function only of his own strategy and an aggregate function of the strategy profile of all players. Aggregate games possess a set of purely algebraic properties that can often provide simple characterizations of equilibrium aggregates without first requiring that one solves for the equilibrium strategy profile. The defining nature of payoffs in an aggregate game allows one to project the n-player strategic analysis of a normal form game onto a lower-dimension aggregate-strategy space, thereby converting an n-player game to a simpler object – a self-generating single-person maximization program. We apply these techniques to a number of economic settings including competition in supply functions and multi-principal common agency games with nonlinear transfer functions.

Keywords: Aggregate games, common agency, asymmetric information, menu auctions.

1. INTRODUCTION

An aggregate game is a normal-form game with the property that each player's payoff is a function only of his own strategy and an aggregate function of the strategy profile of all players. The Cournot quantity game is probably the best known example of such a game (the aggregate for this game is the market supply) and was the motivation behind Selten's (1970) initial treatment of aggregate games. More generally, this class of games encompasses a broad category of strategic settings including public goods games, common resource games, rent-seeking contests, cost sharing games, team games, and patent races, to name a few.[1] Strong comparative statics properties are available in aggregate games if one makes additional assumptions on the relationship between the aggregate and each player's strategy in the payoff functions.[2] In addition, powerful existence theorems easily establish pure-strategy equilibria in these games.[3] That said, the aggregate games that have been typically analyzed have been restricted to finite dimensional spaces and, more often than not, the aggregate space is restricted to a single dimension.[4]

[1]The authors thank Karen Bernhardt-Walther, John Birge, Emir Kamenica, Dan Levin, David Myatt, David Rahman, Michael Schwarz and seminar participants at the University of Chicago and Ohio State University for helpful conversations and suggestions. The present version of this paper consolidates many of the results in the previous working papers "A note on aggregate-strategy games," 2005 and "Tricks with aggregate games," 2007.

[a]Paris School of Economics-EHESS, david.martimort@parisschoolofeconomics.eu

[b]University of Chicago Booth School of Business,lars.stole@chicagobooth.edu

[1]See Bergstrom, et al. (1986), Okuguchi (1993), Corchon (1994), Cornes and Hartley (2000, 2005), Jensen (2009) and Acemoglu and Jensen (2009), and the many references cited therein.

[2]See in particular, Corchon (1994), Dubey, et al. (2006), and Acemoglu and Jensen (2009).

[3]See, for example, Szidarovszky and Yakowitz (1977), Novshek (1984, 1985), Kukushkin (1994, 2004), Dubey, et al. (2006), and Jensen (2009).

[4]Jensen (2009) and Acemoglu and Jensen (2009) briefly consider extensions to larger finite-dimensional aggregates, but the powerful comparative statics and existence properties are compromised by the fact that assumptions of supermodularity and decreasing differences are less likely to be satisfied in many applications that come to mind.

In the games that are of direct interest to the present authors (e.g., competitive agency games with nonlinear transfer functions), strategy spaces are infinite-dimensional and payoffs are not supermodular. Nonetheless, aggregate games possess a set of purely algebraic properties that can often provide simple characterizations of equilibrium aggregates without requiring that one solve for the equilibrium strategy profile. Mathematically, we demonstrate that the defining nature of payoffs in an aggregate game allows one to embed the strategic analysis into the aggregate-strategy space, converting an n-player game to a simpler object − a self-generating single-person maximization program or, equivalently, a two-person, zero-sum game over aggregates. The proofs of these properties are quite straightforward, following directly from algebraic properties of the players' payoffs. Nevertheless, this simple projection of an aggregate game into the space of aggregate strategies allows us to characterize the set of equilibrium aggregates without first solving for equilibrium strategies. We demonstrate this usefulness in a series of examples.

2. AGGREGATE GAMES

As a starting point, consider any n-player normal for game defined by the player set $N = \{1, \ldots, n\}$, each player i's strategy set, X_i, and each player i's payoff function, $u_i : \mathbf{X} \to \mathbb{R}$, where $\mathbf{X} \equiv X_1 \times \cdots \times X_n$ and $\mathbf{x}$ is an arbitrary profile in $\mathbf{X}$. For notational ease, we will use $(x_i, \mathbf{x}_{-i}) \in \mathbf{X}$ to represent the decomposition of a strategy profile into player i's component and the remaining $\mathbf{x}_{-i}$ strategy profile. Accordingly, a strategy profile, $\overline{\mathbf{x}}$, is an equilibrium of a normal-form game if and only if for each $i \in N$,

$$u_i(\overline{x}_i, \overline{\mathbf{x}}_{-i}) \geq u_i(x_i, \overline{\mathbf{x}}_{-i}),\ \forall\, x_i \in \mathbf{X}_i.$$

An **aggregate game** is a normal-form game with the additional requirement that each player i's payoff can be represented as a function which depends only upon his own strategy and an aggregate of the full strategy profile. Formally, we require that there exists a proper[5] aggregate $\phi : \mathbf{X} \to Y$ such that each player's payoff function can be further specialized to the aggregate form

$$\mathbf{x} \mapsto u_i(x_i, \phi(\mathbf{x})).$$

We denote an aggregate normal-form game by the collection

$$G = \left\{ (X_i, u_i)_{i \in N}, \phi \right\}.$$

A defining characteristic of an aggregate game is that player i's preference ordering over X_i is is invariant over the equivalence classes of $\mathbf{X}_{-i}$ for which $\phi(x_i, \mathbf{x}_{-i}) = \phi(x_i, \mathbf{x}'_{-i})$ for all $x_i \in \mathbf{X}_i$. At this stage in the analysis, the definition of an aggregate game is abstract without any mathematical structure placed on the strategy spaces, $\mathbf{X}$, the aggregate space, Y, or the payoff functions $(u_1, \ldots, u_n)$, other than the existence of a common payoff aggregate. Even at this level of abstraction we can state a few important properties of the equilibria set. Later, we will impose a

[5] We say that ϕ is a "proper" aggregate to mean that ϕ is not one-to-one; i.e., the aggregate $y = \phi(\mathbf{x})$ does not identify a unique profile $\mathbf{x}$ but instead is a meaningful "aggregate" of $\mathbf{x}$. If we did not require such properness for ϕ, then every normal-form game would technically be an aggregate game.

linear-symmetric structure on strategies and payoffs, allowing us to characterize additional details of equilibria. As we will see, the class of aggregate games satisfying this additional linear-symmetric structure is quite large. We begin, however, with a very concrete and well-known game for motivation.

Example 1 - Cournot games: Consider the output game of Cournot. Each of the n firms simultaneously produces an output, $q_i \geq 0$, at a cost of $C_i(q_i) = c_i q$. The market price is set such that total demand, $D(p)$, is equal to the quantity supplied, $\sum_{i \in N} q_i$. In this case, using the the aggregate market supply, $\phi(q_1, \ldots, q_n) = \sum_{i \in N} q_i = Q$, we see that the Cournot game is also an aggregate game.[6] Note that the aggregate we chose is not unique. Any strictly monotone function of market output can also generate an aggregate game. In the present case, of course, such a transformed aggregate game is isomorphic to the original.[7] $\diamond$

In defining the class of aggregate games, we require that ϕ is a proper aggregate of the profile $\mathbf{x}$ (i.e., ϕ is not one-to-one) so that the aggregate y cannot be used to perfectly infer the profile, $\mathbf{x}$. In many interesting aggregate games, however, we have the additional property that for a given $\mathbf{x}_{-i}$, the aggregate $y = \phi(x_i, \mathbf{x}_{-i})$ can be inverted to reveal player i's strategy as a function of y. Such invertibility from Y to X_i (if well defined) adds considerable power to the aggregate representation approach we develop. There are a few related notions of invertibility which we define in turn.

For any given $\mathbf{x}_{-i}$, we can define the mapping of ϕ restricted to $\mathbf{x}_{-i}$ by the function $\phi(\cdot, \mathbf{x}_{-i}) : X_i \to Y$. Inverting this mapping, we have a (possibly empty-valued) correspondence

$$\phi_i^{-1}(y, \mathbf{x}_{-i}) \equiv \{x_i \in X_i \mid y = \phi(x_i, \mathbf{x}_{-i})\}.$$

This inverse correspondence is non-empty for all $(y, \mathbf{x}_{-i})$ satisfying $y \in \phi(X_i, \mathbf{x}_{-i})$, where $\phi(X_i, \mathbf{x}_{-i})$ is the image of X_i under ϕ restricted to $\mathbf{x}_{-i}$. If this inverse correspondence is also single-valued for all $y \in \phi(X_i, \mathbf{x}_{-i})$ and $i \in N$, we say that ϕ is **pairwise-injective**. If ϕ is pairwise injective and also surjective on Y (i.e., $\phi(X_i, \mathbf{x}_{-i}) = Y$) for all $\mathbf{x}_{-i} \in \mathbf{X}_{-i}$ and $i \in N$, then we say that ϕ is **pairwise-bijective**. The techniques we develop for aggregate games will be especially fruitful for games with pairwise-bijective aggregates.

Example 1 - Cournot games, continued: Returning to our Cournot game, the aggregate function $\phi(q_1, \ldots, q_n) = \sum_i q_i$ is pairwise-injective and the inverse mapping for each player i is $\phi_i^{-1}(Q, \mathbf{q}_{-i}) = Q - \sum_{j \neq i} q_j$. This aggregate is not pairwise-bijective, however, because firm i's output must be nonnegative, $q_i \geq 0$. Nonnegativity implies that the aggregate which firm i can implement is constrained to be at least $\sum_{j \neq i} q_j$. $\diamond$

Example 2 - Beauty contests:[8] Another example of an aggregate game that is pairwise injective but not pairwise bijective is a variation of Keynes's beauty-contest

[6] Although these aggregates for the Cournot game are symmetric on their domains, symmetry is not required in the definition of an aggregate game.

[7] More generally, one might imagine other normal-form games for which there are two proper aggregate representations that are not isomorphic; we are unaware, however, of any such examples.

[8] We thank David Myatt for suggesting this example.

game. Suppose that the true "beauty" of a person is random variable θ_0 drawn from a publicly known, prior distribution with convex support, $\Theta_0 \subset \mathbb{R}$. Each player in the game is an expert (a " judge" in the beauty contest) who observes a private signal, $s_i \in \Sigma_i$, which is informative about θ_0 and the other players' signals. Player i's strategy is a mapping from private signals to a public action or "prediction" of θ_0, which we denote as the announcement, $x_i(s_i)$. Player i cares both about being correct (i.e., x_i close to the true θ_0) and about his closeness to the average assessment, $y = \frac{1}{n}\sum_j x_j$.

Formally, player i's strategy space is $X_i = \{\tilde{x}_i \mid \tilde{x}_i : \Sigma_i \to \Theta\}$ and $x_i \in X_i$ is an arbitrary strategy mapping signals to predictions. Denote the set of all signal profiles by $\mathbf{\Sigma} \equiv \prod_{i\in N} \Sigma_i$ and let $\mathbf{s} = (s_1, \ldots, s_n) \in \mathbf{\Sigma}$ represent an arbitrary signal profile. Then the aggregate prediction space is the set of additively-separable mappings, $\mathbf{s} \in \mathbf{\Sigma} \mapsto y(\mathbf{s}) \in \Theta_0$:

$$Y \equiv \left\{ y : \mathbf{\Sigma} \to \Theta \;\middle|\; \exists \mathbf{x} \in \mathbf{X} \;\; s.t. \;\; y(\mathbf{s}) = \frac{1}{n}\sum_{i\in N} x_i(s_i) \;,\; \forall \mathbf{s} \in \mathbf{\Sigma} \right\}.$$

Following Morris and Shin (2002), Angeletos and Pavan (2007) and Myatt and Wallace (2008), we suppose that player i's preferences are quadratic but with the additional generality that the weights may vary across players.

$$u_i(x_i, y) = -\int_{\Theta\times\mathbf{\Sigma}} \left((1-\alpha_i)\,(x_i(s_i) - \theta_0)^2 + \alpha_i\,(x_i(s_i) - y(\mathbf{s}))^2 \right) dG(\theta_0, \mathbf{s}),$$

where $G(\theta_0, \mathbf{s})$ is the joint cumulative distribution defined over $\Theta \times \mathbf{\Sigma}$. Because the average includes the player's own assessment, we assume that each α_i is bounded above by $n^2/(2n-1)$ which guarantees that each player's objective function is concave. This upper bound exceeds unity and holds generally for any α_i if n is sufficiently large; there is no corresponding lower bound.[9] Unlike these cited papers, we do not assume that the information structure is Gaussian, nor do we require that $\alpha_i \in (0,1)$. For example, $\alpha_i < 0$ characterizes a game in which player i prefers to be correct about θ but also prefers to be unique relative to the "actions" of rival players.

The aggregate mapping $\phi(\mathbf{x}) = \frac{1}{n}\sum_i x_i$ is pairwise injective because

$$x_i = \phi^{-1}(y, \mathbf{x}_{-i}) = ny - \sum_{j\neq i} x_j$$

is defined for $y \in \phi(X_i, \mathbf{x}_{-i})$. The aggregate prediction is not pairwise bijective, however, as player i is restricted to implementing aggregates in the set $\phi(X_i, \mathbf{x}_{-i}) \subsetneq Y$, which restricts the aggregate to coincide with $\mathbf{x}_{-i}$ over the signals $\mathbf{s}_{-i}$. ◇

In contrast to the games of Cournot and beauty contests, the following two examples illustrate aggregate games with pairwise-bijective aggregates.

[9]If the player's preferences were rewritten as only depending upon the average of *other* players' reports, then this additional technical requirement on α_i would be unnecessary; in such as case, however, the game would not immediately be a member of the aggregate game class as we have defined it and so we prefer to define the beauty contest game as an aggregate game from the beginning.

Example 3 - Supply function games: Consider the supply function game studied in Klemperer and Meyer (1989). Suppose that each of n firms chooses a supply function as its strategy, $S_i : \mathbb{R}_+ \mapsto \mathbb{R}$, where S_i is a bounded function on the domain of nonnegative prices mapping to the number of units firm i commits to supply at the price p. We allow that supply commitments may be negative, which has the interpretation of a net demand by the firm. Each firm chooses its supply function, S_i, prior to knowing the state of demand. We denote demand in state θ as $D(p, \theta)$. After the supply functions are chosen and θ is realized, a market maker chooses the price to equate supply to demand: $\sum_{i \in N} S_i(p) = D(p, \theta)$.[10] The interesting aggregate in this game is the derived market supply function, $S(p) = \sum_{i \in N} S_i(p)$. It is pairwise-bijective because each firm is allowed to choose negative supply at any given price, though in equilibrium we will see that no firm will choose to do so providing marginal costs are nonnegative. $\diamond$

Example 4 - Intrinsic common agency games: Common agency settings with private information are another class of aggregate games with infinite-dimensional strategy spaces.[11] Each of n principals offer a nonlinear contract, T_i, $i = 1, \ldots, n$, which is a commitment by principal i to pay a common agent $T_i(q)$ for any choice $q \in \mathcal{Q}$. The agent's preference ordering over $\mathcal{Q}$ and monetary transfers, however, is private information at the time of contracting, parameterized by $\theta \in \Theta$. In this setting, each principal will typically gain from distorting the choice of q in order to capture some of the agent's information rent. In the simplest contracting game, we assume that common agency is *intrinsic*: i.e., the agent must either accept all contract offers, obtaining the aggregate transfer $T(q) = \sum_{i \in N} T_i(q)$, or reject all contracts and obtain a reservation payoff normalized to zero. Thus, the agent's optimal choice of q will be a function of θ and the aggregate transfer function, $T = \sum_i T_i$. Because principals are allowed to demand payments from the agent (i.e., transfers may be negative), the aggregate transfer function is a pairwise-bijection. $\diamond$

3. CHARACTERIZING EQUILIBRIUM AGGREGATES

Let $\mathcal{E}(G) \subseteq \mathbf{X}$ be the (possibly empty) set of Nash equilibrium strategy profiles for the game G. Let $\mathcal{E}_Y(G) \subseteq Y$ be the corresponding set of aggregate equilibrium outcomes:

$$\mathcal{E}_Y(G) \equiv \{y \in Y \mid y = \phi(\mathbf{x}),\ \mathbf{x} \in \mathcal{E}(G)\}.$$

In this section, we present two closely related approaches to characterize the aggregate equilibrium set $\mathcal{E}_Y(G)$ in aggregate-strategy games. The first characterization result (Theorem 1) is built upon a principle of aggregate-strategy concurrence and is broadly applicable to the entire class of aggregate games. A second collection of stronger results (Corollaries 1 and 2) characterize $\mathcal{E}_Y(G)$ using various self-generating maximization programs. These results, however, rely on an aggregate-invariance assumption that is available only for a strict (but interesting) subset of aggregate games.

[10] To simplify the analysis, we assume that if no such price can be found, then no trade takes place.

[11] The case of intrinsic common agency in a moral hazard setting was first studied in Bernheim and Whinston (1986a) with general contracts but finite outcomes, and later specialized to infinite-dimensional (but linear) contracts by Holmström and Milgrom (1988) and Dixit (1996).

3.1. *Aggregate concurrence*

Our first characterization is based on the notion that at an equilibrium profile $\overline{\mathbf{x}} \in \mathcal{E}(G)$, each player i has the ability to implement any $y \in \phi(X_i, \overline{\mathbf{x}}_{-i})$ and therefore must find $\overline{y} = \phi(\overline{x}) \in \mathcal{E}_Y(G)$ to be optimal among all alternatives $y \in \phi(X_i, \overline{\mathbf{x}}_{-i})$. Because $\overline{y}$ must satisfy similar optimality conditions for each player, in this sense all players must concur over the choice of $\overline{y}$. Indeed, this aggregate concurrence requirement is both necessary and sufficient for $\overline{y} \in \mathcal{E}_Y(G)$.

The key ingredient for the aggregate concurrence result is an observation about best-response correspondences in aggregate games.

Lemma 1 *For each player $i \in N$ and $\mathbf{x}_{-i} \in \mathbf{X}_{-i}$,*

$$x_i \in \arg\max_{\tilde{x}_i \in X_i} u_i(\tilde{x}_i, \phi(\tilde{x}_i, \mathbf{x}_{-i})) \iff \phi(x_i, \mathbf{x}_{-i}) \in \arg\max_{y \in \phi(X_i, \mathbf{x}_{-i})} \left\{ \max_{\tilde{x}_i \in \phi_i^{-1}(y, \mathbf{x}_{-i})} u_i(\tilde{x}_i, y) \right\}.$$

From Lemma 1, we can characterize the set of equilibria using two-stage aggregate best responses. This allows us to restate the definition of equilibrium as the requirement that no player can improve his payoff by changing the aggregate. It follows that the equilibrium aggregate, $\phi(\overline{\mathbf{x}})$, must lie in the best-response correspondence of every player (i.e., in the intersection of the players' aggregate best-response correspondences). Thus, $\overline{\mathbf{x}}$ is a Nash equilibrium if and only if

$$\phi(\overline{\mathbf{x}}) \in \bigcap_{i \in N} \arg\max_{y \in \phi(X_i, \overline{\mathbf{x}}_{-i})} \max_{x_i \in \phi_i^{-1}(y, \overline{\mathbf{x}}_{-i})} u_i(x_i, y).$$

Notice we have not required that ϕ is pairwise injective. The preceding argument suggests that in cases in which ϕ is not a pairwise-injection, we can nevertheless project each player's strategy space onto aggregate-equivalence sets and analyze the game with the strategies chosen from the quotient set, $X_i/_\sim$, embedding the first step of optimization into each player's payoff function.[12] If we are only interested in the equilibrium aggregate, $\phi(\overline{\mathbf{x}})$, and not the component strategies $\overline{\mathbf{x}}$, this reduction is without any loss.

Collecting these results together, we have our first characterization theorem.

Theorem 1 *If $G = \{(X_i, u_i)_{i \in N}, \phi\}$ is an aggregate-strategy game, then*

$$\overline{\mathbf{x}} \in \mathcal{E}(G) \iff \phi(\overline{\mathbf{x}}) \in \bigcap_{i \in N} \arg\max_{y \in \phi(X_i, \overline{\mathbf{x}}_{-i})} \max_{x_i \in \phi_i^{-1}(y, \overline{\mathbf{x}}_{-i})} u_i(x_i, y).$$

If G is a pairwise-injective, aggregate-strategy game, then

$$\overline{\mathbf{x}} \in \mathcal{E}(G) \iff \phi(\overline{\mathbf{x}}) \in \bigcap_{i \in N} \arg\max_{y \in \phi(X_i, \overline{\mathbf{x}}_{-i})} u_i\left(\phi_i^{-1}(y, \overline{\mathbf{x}}_{-i}), y\right). \tag{3.1}$$

[12]Here, the equivalence relation is defined with respect to the aggregate, ϕ. That is, $\mathbf{x} \sim \mathbf{x}' \iff \phi(\mathbf{x}) = \phi(\mathbf{x}')$.

Going forward, we will assume that ϕ is pairwise injective and therefore (3.1) gives the key result.[13] A remarkable consequence of the theorem is that in any equilibrium, all players must agree on the choice of the aggregate. This *principle of aggregate concurrence* provides a powerful perspective for analyzing equilibria in aggregate games.[14]

3.2. *Self-generating maximization programs*

The aggregate concurrence principle suggests that a maximization program over Y might be constructed for which the set of maximizers corresponds with $\mathcal{E}_Y(G)$.[15] If so, such a program would allow conclusions about equilibria to be drawn directly from the maximization program rather than the more complex strategic game. With few exceptions, unfortunately, such a maximization program would also be self-referential, requiring a fixed-point to be satisfied as part of the solution.

To fix ideas, consider first the general case of normal-form games. The definition of equilibrium can be stated equivalently as a *self-generating maximization* (SGM) program: For any vector of positive weights $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$, $\overline{\mathbf{x}}$ is an equilibrium iff

$$\overline{\mathbf{x}} \in \arg\max_{x \in \mathbf{X}} \sum_{i \in N} \lambda_i u_i(x_i, \overline{\mathbf{x}}_{-i}).$$

Denoting the right-hand side more succinctly as $\Lambda(\mathbf{x}, \overline{\mathbf{x}}, \boldsymbol{\lambda})$, we have the statement that for any $\boldsymbol{\lambda} \in \mathbb{R}^n_{++}$,

$$\overline{\mathbf{x}} \in \mathcal{E}(G) \iff \overline{\mathbf{x}} \in \arg\max_{\mathbf{x} \in \mathbf{X}} \Lambda(\mathbf{x}, \overline{\mathbf{x}}, \boldsymbol{\lambda}).$$

In the case of general normal-form games, the solution to this SGM program is of the same order of complexity as finding the fixed points to the n-player game. Nothing is gained by reformulating the problem in this manner.

In the case of aggregative normal-form games, however, *if one is only interested in the equilibrium aggregates*, $\mathcal{E}_Y(G)$, then an aggregate SGM program may reduce the complexity of the problem considerably. In particular, we may be able to project the SGM program defined over $\mathbf{X}$ onto the smaller aggregate space Y. Ideally, we seek conditions for which

$$\overline{y} \in \mathcal{E}_Y(G) \iff \overline{y} \in \arg\max_{y \in Y} \Lambda(y, \overline{y}, \boldsymbol{\lambda}). \tag{3.2}$$

As a starting point, given that ϕ is pairwise-injective, at an equilibrium point $\overline{\mathbf{x}} \in \mathcal{E}(G)$ we may define the collective surplus of the players by the welfare function

$$W(y, \overline{\mathbf{x}}, \boldsymbol{\lambda}) \equiv \sum_{i \in N} \lambda_i u_i(\phi_i^{-1}(y, \overline{\mathbf{x}}_{-i}), y).$$

---

[13] For applications in which pairwise injectivity is not an appropriate assumption, our results can be reinterpreted as holding on the aggregate-quotient space of the game.

[14] While others have used this idea indirectly, to our knowledge Bernheim and Whinston (1986a) are the first to recognize explicitly the force of this principle in their study of moral hazard in a common agency game with finite-dimensional strategies: "We underscore the need to make the principals' objectives congruent in equilibrium: since all principals can effect the same changes in the aggregate incentive scheme, none must find any such change worthwhile." (*Ibid.*, p.929.)

[15] There is a related literature on best-response equivalences between normal-form games and identical-interest (a.k.a. best-response potential) games. Voorneveld (2000) and Morris and Ui (2004) have proven several interesting results in this direction. We address the relationship between this literature and the aggregate game results in the present paper in the final section of this paper.

Theorem 1 states that an equilibrium aggregate, $\overline{y} = \phi(\overline{\mathbf{x}})$, must maximize W over a suitably restricted feasible set of aggregates for any non-negative vector $\boldsymbol{\lambda} \in \Delta$. We have the immediate corollary to Theorem 1.

COROLLARY 1 *If G is a pairwise-injective aggregate game, then*

$$\overline{\mathbf{x}} \in \mathcal{E}(G) \iff \forall \boldsymbol{\lambda} \in \Delta, \quad \phi(\overline{\mathbf{x}}) \in \bigcap_{\substack{i \in N: \\ \lambda_i > 0}} \arg\max_{y \in \phi(X_i, \overline{\mathbf{x}}_{-i})} W(y, \overline{\mathbf{x}}, \boldsymbol{\lambda}). \tag{3.3}$$

Stated in this way, the aggregate concurrence principle is an invariance property: The subset of optimal aggregates that are invariant to the welfare weights is exactly the set of equilibrium aggregates. The value of Corollary 1 is most clear if we place additional structure on ϕ and W.

DEFINITION 1 *We say that the game G is* **aggregate-invariant** *if there exists a vector of welfare weights, $\boldsymbol{\lambda} \in \Delta$, such that*

$$\arg\max_{y \in Y_0} W(y, \mathbf{x}, \boldsymbol{\lambda}) = \arg\max_{y \in Y_0} W(y, \mathbf{x}', \boldsymbol{\lambda})$$

for any $Y_0 \subseteq Y$ and any $\mathbf{x}$, $\mathbf{x}'$ such that $\phi(\mathbf{x}) = \phi(\mathbf{x}')$.

The immediate implication of aggregate invariance is that we may take any such $\boldsymbol{\lambda}$ and define the function

$$\Lambda(y, \phi(\mathbf{x}), \boldsymbol{\lambda}) = W(y, \mathbf{x}, \boldsymbol{\lambda})$$

with the property that for any $\hat{\mathbf{x}}$ in the aggregate equivalence class $[\hat{y}] = \{\mathbf{x} \in \mathbf{X} \mid \phi(\mathbf{x}) = \hat{y}\}$, we have

$$\arg\max_{y \in Y_0} \Lambda(y, \hat{y}, \boldsymbol{\lambda}) \equiv \arg\max_{y \in Y_0} W(y, \hat{\mathbf{x}}, \boldsymbol{\lambda}).$$

Using such a construction, we can deduce a particularly useful corollary from Theorem 1 which gives necessary conditions in form of (3.2).

COROLLARY 2 *If G is an aggregate-invariant game for weight vector $\boldsymbol{\lambda}$, then*

$$\overline{\mathbf{x}} \in \mathcal{E}(G) \implies \phi(\overline{\mathbf{x}}) \in \bigcap_{i \in N} \arg\max_{y \in \phi(X_i, \overline{\mathbf{x}}_{-i})} \Lambda(y, \phi(\overline{\mathbf{x}}), \boldsymbol{\lambda}). \tag{3.4}$$

If in addition ϕ is pairwise-bijective, then

$$\overline{y} \in \mathcal{E}_Y(G) \implies \overline{y} \in \arg\max_{y \in Y} \Lambda(y, \overline{y}, \boldsymbol{\lambda}). \tag{3.5}$$

The import of the corollary is that if we know an equilibrium exists and the game is aggregate invariant, then it suffices to restrict attention to the set of self-generating maximizers of Λ. If G is aggregate invariant for multiple welfare weights, then for each such $\boldsymbol{\lambda}$ we can construct a corresponding Λ objective. An equilibrium aggregate must be a solution to all such SGM programs. If the solution set to any SGM program has a unique element, then it represents the unique Nash equilibrium outcome. If the set contains multiple equilibrium candidates, then we may consider other aggregate-invariant welfare weights, and restrict attention to the intersection of the SGM solution sets.

Example 2 - Beauty contests, continued: Consider again player i's payoff function,

$$u_i(x_i, y) = -\int_{\Theta\times\mathbf{\Sigma}} \left((1-\alpha_i)\,(x_i(s_i)-\theta_0)^2 + \alpha_i\,(x_i(s_i)-y(\mathbf{s}))^2\right) dG(\theta_0,\mathbf{s}).$$

Substituting $x_i \mapsto n(y-\overline{y}) + \overline{x}_i$, expanding the quadratic terms, and eliminating all terms which are independent of both x_i and y, we obtain a simpler, equally faithful, representation of player i's preferences over y given i''s rivals are playing $\overline{\mathbf{x}}_{-i}$:

$$u_i\left(n(y-\overline{y})+\overline{x}_i, y\right) =$$
$$\int_{\Theta\times\mathbf{\Sigma}} \left\{-\frac{\psi_i}{2}y(\mathbf{s})^2 + y(\mathbf{s})(n-\alpha_i)(n\overline{y}(\mathbf{s})-\overline{x}_i(s_i)) - y(\mathbf{s})n(1-\alpha_i)\theta\right\} dG(\theta_0,\mathbf{s}).$$

Above we have introduced the notation $\psi_i = (n-\alpha_i)^2 + \alpha_i(1-\alpha_i)$. Given the upper bound on α_i, it follows that $\psi_i > 0$ and that $n > \alpha_i$. Using the weight of $\lambda_i = \frac{1}{n(n-\alpha_i)}$ for each player i, we can aggregate the players' payoffs to obtain an aggregate-invariant representation:

$$\Lambda(y,\overline{y},\boldsymbol{\lambda}) = -\int_{\Theta\times\mathbf{\Sigma}} \left\{\frac{1}{2}y(\mathbf{s})^2\left(\frac{1}{n}\sum_{i\in N}\frac{\psi_i}{n-\alpha_i}\right) - y(\mathbf{s})\overline{y}(\mathbf{s})(n-1)\right.$$
$$\left. - y(\mathbf{s})n\theta\left(\frac{1}{n}\sum_{i\in N}\frac{1-\alpha_i}{n-\alpha_i}\right)\right\} dG(\theta_0,\mathbf{s}).$$

Define the scalar

$$\kappa \equiv \sum_{i\in N}\frac{1-\alpha_i}{n-\alpha_i},$$

and note that $\kappa > 1-n$. Using the expression for ψ_i, it can be established that

$$\frac{1}{n}\sum_{i\in N}\frac{\psi_i}{n-\alpha_i} - (n-1) = \sum_{i\in N}\frac{1-\alpha_i}{n-\alpha_i} = \kappa.$$

This algebraic fact allows us to simplify the aggregate objective:

$$\Lambda(y,\overline{y},\boldsymbol{\lambda}) = -\int_{\Theta\times\mathbf{\Sigma}} \left\{\frac{1}{2}y(\mathbf{s})^2\left(\kappa+(n-1)\right) - y(\mathbf{s})\left(\overline{y}(\mathbf{s})(n-1)+\kappa\,\theta_0\right)\right\} dG(\theta_0,\mathbf{s}).$$

The integrand in the definition of Λ is continuously differentiable and strictly concave in $y(\mathbf{s})$ pointwise, and so the solution set to the SGM program in Corollary 2 may be computed by optimizing pointwise with respect to $y(\mathbf{s})$ over the set of feasible aggregates and imposing $y(\mathbf{s}) = \overline{y}(\mathbf{s})$ on the solution for each $\mathbf{s}\in\mathbf{\Sigma}$. Because Θ is convex, the local first-order condition therefore a necessary condition for any equilibrium aggregate. Remarkably, aggregate concurrence implies that for any generic preference vector (i.e., $\kappa \neq 0$), for any arbitrary information structure $(G, \Theta\times\mathbf{\Sigma})$, and for every equilibrium, the aggregate estimate $\overline{y}(\mathbf{s})$ is an unbiased estimate of θ_0.

PROPOSITION 1 *For any beauty-contest game with $\kappa \neq 0$, every equilibrium aggregate, $\overline{y}$, is an unbiased estimate of θ_0:*

$$\int_{\Theta\times\mathbf{\Sigma}} (\overline{y}(\mathbf{s}) - \theta_0) dG(\theta_0, \mathbf{s}) = 0. \tag{3.6}$$

For any beauty-contest game with $\kappa = 0$, there is an equilibrium in which the aggregate is unbiased if the conditional expectation $\mathbf{E}[(\theta_0, \mathbf{s}_{-i}) \mid s_i]$ is linear in s_i for each $i \in N$.

The genericity requirement, $\kappa \neq 0$, rules out pathological cases such as when $\alpha_i = 1$ for all $i \in N$. In this particular case, there is an uncountable number of biased equilibrium outcomes and any aggregate report can arise in equilibrium. Less obvious pathologies are also ruled out. Consider a two player game with preferences given by $\alpha_1 = \frac{7}{6}$, $\alpha_2 = \frac{3}{4}$ and assume that the players' signals are independent, conditional on θ_0. For every $\hat{\theta} \in \Theta$, there exists a corresponding equilibrium in which $\mathbf{E}[\overline{y}(\mathbf{s})] = \hat{\theta}$.[16] For $\hat{\theta} = \mathbf{E}[\theta_0]$, the aggregate is unbiased; otherwise, there are a uncountable number of biased equilibria. By using an aggregate-game framework to characterize the equilibria, the key genericity condition, $\kappa \neq 0$, is immediate. If this genericity condition fails, an unbiased equilibrium aggregate still exists if conditional expectations are linear in signals (e.g., a Gaussian information structure), but unbiasedness is no longer assured for all equilibria. Technically, there is a failure of aggregate lower-hemicontinuity over the parameter space at $\kappa = 0$.

While this simple "unbiasedness" result does not tell us about the social value of information, it applies to a much larger class of beauty-contest games than has been studied in such papers as Morris and Shin (2002), Angeletos and Pavan (2007) and Myatt and Wallace (2008). Notice that the number of players is finite, so every player's action has a measurable impact on the aggregate. The players may also have very different preferences about the desirability of being close to the aggregate $\overline{y}(\mathbf{s})$. The information structure in the present analysis is not necessarily Gaussian, nor is it necessarily symmetric across players. Finally, there is no requirement that the players' signals are independent conditional on θ_0. This generality makes our unbiased-aggregate result more surprising.

Suppose, for example, that signals are not independent, conditional on θ_0. In particular, consider a two-player game and suppose that player 1's signal is $s_1 = \theta_0 + \varepsilon_1$, but that player 2's signal is $s_2 = \varepsilon_1$. Hence, player 2 knows nothing about θ_0 that is not contained in the public prior, but player 2 knows the exact bias in player 1's signal. In equilibrium, $\overline{x}_1(s_1)$ cannot be everywhere constant, and so if $\alpha_2 \neq 0$, player 2's action will depend upon $s_2 = \varepsilon_1$ with the result that x_2 will be a biased estimate of θ_0. The unbiased-aggregate result, however, implies that on average such individual biases cancel each other out.[17]

Finally, we emphasize again that aggregate unbiasedness also holds even if $\alpha_i < 0$ for some players. With such preferences, players prefer to make accurate predictions

[16] The equilibrium strategies are given by $\overline{x}_1(s_1) = \frac{7}{3}\mathbf{E}[\theta_0] - \frac{4}{3}\mathbf{E}[\theta_0 \mid s_1] + \frac{7}{5}\hat{\theta}$ and $\overline{x}_2(s_2) = \mathbf{E}[\theta_0] + \frac{3}{5}\hat{\theta}$.

[17] If the prior distribution of θ_0 is normal with mean μ and variance σ^2, and if player 1's signal noise is normally distributed with mean zero and variance τ^2, then the linear equilibrium of this game is $\overline{x}_1(s_1) = a_1 s_1 + b_1$ and

$$\overline{x}_2(s_2) = \frac{2(1-\alpha)}{(2-\alpha)}\mu + \frac{\alpha}{(2-\alpha)}(a_1(\mu + s_2) + b_1),$$

(i.e., $1-\alpha > 0$), but also prefer to distinguish themselves from the mob.[18] Remarkably, equilibrium unbiasedness is also necessary for games in which players prefer to be dissimilar. ◇

Comparative statics. The SGM programs in (3.4) and (3.5) suggest that one may be able to directly compute various marginal properties about equilibrium aggregates when Λ is differentiable. For example, when ϕ is pairwise bijective and Λ is smooth, we can differentiate the expression in (3.5), and apply the Envelope Theorem to conclude that any equilibrium aggregate $\overline{y}$ must satisfy

$$\Lambda_y(\overline{y}, \overline{y}, \boldsymbol{\lambda}) = 0. \tag{3.7}$$

If the game G is parameterized by ρ, we can incorporate ρ into an SGM program directly and redefine the objective as a function of ρ: $\Lambda(y, \overline{y}, \boldsymbol{\lambda};\, \rho)$. The comparative statics on any aggregate equilibrium selection, $\overline{y}(\rho)$, can be obtained using standard comparative static analysis. For example, under assumptions of differentiability, an immediate comparative static is

$$\text{sign}\left\{\overline{y}'(\rho)\right\} = \text{sign}\left\{\Lambda_{\overline{y}\rho}(\overline{y}(\rho), \overline{y}(\rho), \boldsymbol{\lambda};\, \rho)\right\}.$$

Differentiability is not necessary, of course. It is sufficient that $\Lambda(y, \overline{y}, \boldsymbol{\lambda};\, \rho)$ exhibits either increasing or decreasing differences in (y, ρ).

Equivalence with zero-sum games. The self-generating program in (3.5) can be recast as finding the minmax value of a fictitious two-person, zero-sum game. Suppose that G is aggregate-invariant and ϕ is pairwise-bijective. Let player A have the payoff $\Phi(y_a, y_b) \equiv \Lambda(y_a, y_b, \boldsymbol{\lambda}) - \Lambda(y_b, y_b, \boldsymbol{\lambda})$ and player B the negative of Φ. By construction, it follows that

$$\min_{y_b \in Y} \max_{y_a \in Y} \; \Phi(y_a, y_b) \geq 0 \geq \max_{y_a \in Y} \min_{y_b \in Y} \; \Phi(y_a, y_b).$$

The minmax solution (if one exists[19]) corresponds to values of $\overline{y}_a$ and $\overline{y}_b$ such that $\Phi(\overline{y}, \overline{y}) = 0$ and $\Lambda(\overline{y}, \overline{y}, \boldsymbol{\lambda}) = \max_{y \in Y} \Lambda(y, \overline{y}, \boldsymbol{\lambda})$. In this sense, when an equilibrium exists, finding the equilibrium aggregates to an n-player aggregate game over $\mathbf{X}$ is equivalent to solving a two-person zero-sum game over the smaller aggregate space, Y.

The preceding results for SGM programs offer useful tools for characterizing equilibria, but require that G be aggregate-invariant. We now look for primitive assumptions on G which guarantee aggregate-invariance and therefore afford us the simpler SGM program. Linearity will be the crucial ingredient.

---

where

$$a_1 = \frac{2(1-\alpha)(2-\alpha)\sigma^2}{4(1-\alpha)\tau^2 + (2-\alpha)^2\sigma^2}, \quad b_1 = \mu\left(\frac{4(1-\alpha)\tau^2 + (2-\alpha)\alpha\sigma^2}{4(1-\alpha)\tau^2 + (2-\alpha)^2\sigma^2}\right).$$

We have $\mathrm{E}_{(s_1,s_2,\theta_0)}[\frac{1}{2}(\overline{x}_1(s_1) + \overline{x}_2(s_2)) - \theta_0] = 0$ as implied by Proposition 1.

[18] For example, as in Prendergast and Stole (1996), the player prefers to signal his information is of higher quality by taking more extreme positions.

[19] If $\Phi(y_a, y_b)$ is quasi-concave in y_a and quasi-convex in y_b, then we are guaranteed such a solution exists. Sion (1958).

4. LINEAR AGGREGATE GAMES

4.1. *Definitions*

The main representation result in Theorem 1 is based on an abstract notion of aggregation, with little additional algebraic or topological structure on the game. The more powerful characterizations which emerge from a self-generating maximization program, however, required that the game has an aggregate-invariant welfare function. Understanding this last assumption, in particular, requires that we place more structure on the aggregate game, G. We proceed in this direction by first defining a large class of aggregate games with a linear payoff structure that includes our various working examples. We then introduce symmetry restrictions that guarantee aggregate invariance.

In what follows, it is helpful to introduce the notion of a bilinear form. A bilinear form is a mapping defined on two (possibly distinct) linear spaces which is a linear function of each argument. We use $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{W} \to \mathbb{R}$ to denote such a mapping, with the defining property that for every $f \in \mathcal{V}$, $\langle f, \cdot \rangle$ is a linear functional on $\mathcal{W}$ and for every $g \in \mathcal{W}$, $\langle \cdot, g \rangle$ is a linear functional on $\mathcal{V}$. The notation we use is evocative of an inner product, which is an example of a bilinear form that is also symmetric and positive definite. In the simplest such case, $\mathcal{V} = \mathcal{W} = \mathbb{R}$ and the bilinear form reduces to scalar multiplication. A bilinear form, however, is more general than an inner product as it allows asymmetries and semi-definiteness. [20] In Section 5, for example, we define a large class of distributional games by specializing the bilinear form to the Lebesgue-Stieltjes integral of a function, $f \in \mathcal{V}$, with respect to the measure derived from the distribution $g \in \mathcal{W}$, and thus $\langle f, g \rangle \equiv \int f dg$. In this case, $\mathcal{V}$ denotes the space of bounded Borel-measurable functions and $\mathcal{W}$ is a space of Radon measures.

Presently, we use the bilinear form in a general sense to define payoffs in the class of linear aggregate games.

DEFINITION 2 *An aggregate game G is a* **linear aggregate game** *if*

1. $X_1, \ldots, X_n$ *are subsets of a real linear space,* $(\mathcal{V}, +, \cdot)$*;*
2. $\phi : \mathbf{X} \to Y$ *is a linear operator which can be written (up to a normalization) as*[21]

$$\phi(\mathbf{x}) = \sum_{i \in N} x_i$$

where $Y = X_1 + \cdots + X_n$*, and "+" is denotes set addition.*[22]

3. *preferences for player* $i \in N$ *are representable by,* $u_i : X_i \times Y \to \mathbb{R}$*,*

$$u_i(x_i, y) \equiv \alpha_i(y) + \langle x_i, \beta_i(y) \rangle,$$

where $\alpha_i : Y \to \mathbb{R}$ *is a real function,* $\beta_i : Y \to \mathcal{W}$ *is a mapping to the real linear space* $\mathcal{W}$*, and* $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{W} \to \mathbb{R}$ *is a bilinear form.*

[20] As an example of semi-definiteness, $\langle f, g \rangle = \langle f', g \rangle$ for $g \neq \mathbf{0}_W$ need not imply that $f = f'$.

[21] A linear operator has the form $\phi(\mathbf{x}) = \sum_{i \in N} x_i z_i$. We can transform $x_i \mapsto x_i z_i$ and adjust the payoff functions to obtain our simpler form. More generally, with some effort at defining an algebraic group and the operation or "addition", we can extend the class of aggregate games to include more general operators (e.g., aggregates that are products of strategies, etc.) by considering isomorphic strategy spaces and corresponding payoff functions.

[22] If A, B are sets, $b \in B$, then $A + \{b\} = \{a + b \mid a \in A\}$ and $A + B = \{a + b \mid a \in A,\ b \in B\}$.

The first requirement guarantees that scalar multiplication and addition are well defined on the strategy and aggregate spaces and that strategy spaces can be positively scaled.

The second assumption restricts attention to additive operators, which are naturally pairwise-injective. Recall that although ϕ is additive, it is not necessarily pairwise-bijective (e.g., aggregate output in a Cournot game illustrates this relationship). If, however, each X_i is closed under addition *and* subtraction, then ϕ is pairwise bijective. In what follows, the set of equilibrium aggregates for games in which ϕ is pairwise bijective (rather than simply pairwise-injective) are significantly simpler to characterize.

The third assumption is the most restrictive as it requires that $u_i(x_i, y) - u_i(0, y)$ is linear in x_i. While the payoff linearity restriction still allows for nonlinearities in y (and indirectly in x_i through ϕ), the restriction does rule out more general aggregate games. For example, Cournot games with non-constant marginal costs would give rise to payoff functions that are not affine in a player's own strategy (holding the aggregate constant). Fortunately, many aggregate games do satisfy this linearity requirement, including three of our working examples.

To this linear structure, it will be productive to add various notions of payoff symmetry on the bilinear representation

$$u_i(x_i, y) = \alpha_i(y) + \langle x_i, \beta_i(y)\rangle.$$

DEFINITION 3 *G is a* **weakly symmetric** *linear aggregate game if it is a linear aggregate game and for each player $i \in N$*

$$\beta_i(y) = \beta(y) + \delta_i,$$

where $\beta : X \to \mathcal{V}$ and $\delta_i \in \mathcal{V}$.

G is **strongly symmetric** *if, in addition, payoffs satisfy*

$$\beta_i(y) = \beta(y)$$

$$\alpha_i(y) = \alpha(y) + \langle \gamma_i, \beta(y)\rangle,$$

where $\alpha : X \to \mathbb{R}$ and $\gamma_i \in \mathcal{V}$, and the strategy spaces are symmetric,

$$X_1 = \cdots = X_n = Y = X \subseteq \mathcal{V}.$$

The import of weak symmetry implies the difference in payoff to player i between $x_i = x$ and $x_i = x'$ minus the difference in payoff to player j between the same x and x' is linear in $x - x'$ and independent of y:

$$(u_i(x, y) - u_i(x', y)) - (u_j(x, y) - u_j(x', y)) = \langle x - x', \delta_i - \delta_j\rangle.$$

With the additional requirement of strong symmetry this difference is zero. Moreover, the difference in payoff to player i between y and y' minus the difference in payoff to player j between the same y and y' is linear in $\beta(y) - \beta(y')$ and independent of x:

$$(u_i(x, y) - u_i(x, y')) - (u_j(x, y) - u_j(x, y')) = \langle \gamma_i - \gamma_j, \beta(y) - \beta(y')\rangle.$$

Strong symmetry has a further requirement that the players' strategy spaces are identical. Note that in any linear aggregate game, if ϕ is pairwise bijective, then there is a linear bijection between any two players' strategy sets and so, without loss of generality, we may take the players' strategy spaces to be equal.

Supply-function games (with symmetric marginal costs) and intrinsic common agency games are both strongly-symmetric, linear aggregate games. This is not true for our other examples.

Example 1 - Cournot games, continued: The class of weakly-symmetric, linear aggregate games contains many interesting applications, including the Cournot game with payoffs

$$\pi_i(q_, Q) = q_i(P(Q) - c_i) - F_i.$$

The assumption of strong symmetry is more restrictive, further requiring that marginal costs are symmetric across firms, $c_i = c_j$. Strategy spaces are equal across players, but the aggregate is not pairwise bijective. ◇

Example 2 - Beauty contests, continued: The Beauty contest game is, strictly speaking, not a linear aggregate game because it is quadratic in x_i. That said, after making the substitution from x_i to $n(y - \overline{y}) + \overline{x}_i$ and simplifying, the resulting preferences are linear in $\overline{x}_i$, and therefore we are able to establish aggregate invariance. In this sense, the transformed Beauty-contest game is a linear aggregate game. Note, however, that while the preferences of the players satisfy the conditions of strong symmetry, the strategy spaces of the players are not themselves symmetric. For example, player 1's strategy set is the set of mappings from player 1's signal set to reports; this is distinct from principal 2's strategy space. The consequence is that the aggregate space is richer than any individual player's strategy space – the aggregate is an additively-separable mapping from all signal spaces. Hence, beauty contests are only weakly symmetric. ◇

4.2. *Necessary conditions for $\mathcal{E}_Y(G)$*

We restrict attention to weakly-symmetric, linear aggregate games and ask what are the characteristics of any equilibrium aggregate. To this end, the following lemma which establishes aggregate-invariance for the class of linear games is invaluable.

Lemma 2 **(Aggregate-invariance)** *Suppose that G is a weakly-symmetric, linear aggregate game. For any $M \subseteq N$ and any $Y_0 \subseteq X_1 + \cdots + X_n$,*

$$\arg\max_{y \in Y_0} \frac{1}{|M|} \sum_{i \in M} u_i(y - \overline{y} + \overline{x}_i, y) = \arg\max_{y \in Y_0} \Lambda_M(y, \overline{y}_M, \overline{y}),$$

where

$$\Lambda_M(y, \overline{y}_M, \overline{y}) \equiv \langle \beta(y) + \delta_M, y - \overline{y} + \overline{y}_M/|M| \rangle + \alpha_M(y),$$

and $\overline{y}_M = \sum_{i \in M} \overline{x}_i$, $\delta_M = \frac{1}{|M|} \sum_{i \in M} \delta_i$, and $\alpha_M(y) = \frac{1}{|M|} \sum_{i \in M} \alpha_i(y)$.

Given that we are using uniform welfare weights, we have suppressed the argument $\boldsymbol{\lambda}$ from Λ in the Lemma above. We follow this convention for the remainder of this paper.

Setting $M = N$, an immediate implication of the lemma is that G is aggregate-invariant for uniform welfare weights. That is, for any $Y_0 \subseteq X_1 + \cdots + X_n$, we have

$$\arg\max_{y \in Y_0} \frac{1}{n} \sum_{i \in N} u_i(y - \overline{y} + \overline{x}_i, y) = \arg\max_{y \in Y_0} \Lambda(y, \overline{y}),$$

where $\Lambda(y, \overline{y}) \equiv \Lambda_N(y, \overline{y}_N, \overline{y})$.

Consider more generally the subset of aggregates that any player in a given set $M \subseteq N$ can implement, $\bigcap_{i \in M} \phi(X_i, \overline{\mathbf{x}}_{-i})$. If G is a linear aggregate game, then $\phi(X_i, \overline{\mathbf{x}}_{-i}) = X_i + \{\overline{y} - \overline{x}_i\}$. It follows that

$$\bigcap_{i \in M} \phi(X_i, \overline{\mathbf{x}}_{-i}) = \bigcap_{i \in M} X_i + \{\overline{y} - \overline{x}_i\}.$$

If ϕ is pairwise-bijective (and therefore the players' strategy spaces and the aggregate space can be represented as $X = X_1 = \cdots = X_n = Y$), we have for any $M \subseteq N$ the more powerful relationship

$$\bigcap_{i \in M} \phi(X_i, \overline{\mathbf{x}}_{-i}) = X.$$

With these insights, we combine the construction of $\Lambda_N(y, \overline{y}_N, \overline{y})$ from Lemma 2) with Corollary 2 to obtain our main necessity theorem.

Theorem 2 *Suppose that G is a weakly-symmetric, linear aggregate game. Then for every $M \subseteq N$,*

$$\overline{\mathbf{x}} \in \mathcal{E}(G) \implies \phi(\overline{\mathbf{x}}) \in \bigcap_{i \in M} \arg\max_{y \in X_i + \{\overline{y} - \overline{x}_i\}} \Lambda_M(y, \sum_{i \in M} \overline{x}_i, \phi(\overline{\mathbf{x}})) \subseteq \arg\max_{\cap_{i \in M} X_i + \{\overline{y} - \overline{x}_i\}} \Lambda_M(y, \sum_{i \in M} \overline{x}_i, \phi(\overline{\mathbf{x}})), \tag{4.1}$$

and consequently

$$\overline{\mathbf{x}} \in \mathcal{E}(G) \implies \phi(\overline{\mathbf{x}}) \in \arg\max_{\cap_{i \in N} X_i + \{\overline{y} - \overline{x}_i\}} \Lambda(y, \phi(\overline{\mathbf{x}})). \tag{4.2}$$

If ϕ is pairwise-bijective, then

$$\overline{y} \in \mathcal{E}_Y(G) \implies \overline{y} \in \arg\max_{y \in X} \Lambda(y, \overline{y}). \tag{4.3}$$

The sets of solutions to the various SGM programs above are often quite tractable to calculate as we will see when applied to our examples. As an immediate illustration, we apply Theorem 2 to the Cournot game.

Example 1 - Cournot games, continued: As argued above, the Cournot output game with constant unit costs of production is a linear aggregate game with weakly-symmetric payoffs. As a thought experiment, suppose that Cournot aggregation is pairwise-bijective. Then the application of (4.3) implies

$$\overline{Q} \in \arg\max_{Q \geq 0} \frac{1}{n}(P(Q) - c_N)(nQ - (n-1)\overline{Q})$$

where $c_N = \frac{1}{n}\sum_i c_i$ is the average unit cost in the industry. This SGM program can be restated as

$$\overline{Q} \in \arg\max_{Q \geq 0} Q(P(Q) - c_N) + (n-1)(Q - \overline{Q})(P(Q) - c_N). \tag{4.4}$$

where the first term represents industry profit and the second term captures the strategic effects on output. Without any conditions on demand, we could conclude that in any equilibrium with $n > 1$ active firms, there will be over-production relative to the monopoly level. This follows because the second term is increasing in Q around any equilibrium point $Q = \overline{Q}$. Furthermore, providing that $P(Q)$ is concave (or at least not too convex), there exists a unique solution to the SGM program in (4.4).

Of course, using (4.3) is predicated on the false assumption that aggregation is pairwise-bijective. The appropriate SGM program to consider is in fact (4.1):

$$\overline{Q} \in \arg\max_{Q \geq \max_{i \in M} \sum_{j \neq i} \overline{q}_j} Q(P(Q) - c_M) + (m-1)(Q - \overline{Q})(P(Q) - c_M) + (P(Q) - c_M)(\overline{Q}_M - \overline{Q}).$$

If it is known that M contains exactly those firms that produce positive output in equilibrium, then $\overline{Q} = \overline{Q}_M$ and the inequality constraint on Q in the above argmax program is slack. Providing that the objective is quasi-concave, we can replace the slack constraint with the weaker constraint $Q \geq 0$ without changing the solution to the program.

$$\overline{Q} \in \arg\max_{Q \geq 0} Q(P(Q) - c_M) + (m-1)(Q - \overline{Q})(P(Q) - c_M). \tag{4.5}$$

Conditional on the equilibrium set of active firms, M, the SGM program in (4.5) implies that output is greater than that which maximizes the collective profits of the producing firms.

One can also readily see from the statement in (4.5) that a mean-preserving change in the distribution of industry unit costs that leaves the set of active firms unchanged can have no effect on the market price. This result is similar to the well-known neutrality of differential taxation in public finance. Of course, the neutrality result is as fragile as the condition that the set of active firms is invariant, a fact well known in the existing applied-theory literatures. Bergstrom, Blume and Varian (1985) first noted such an invariance in the context of public goods games; Levin (1984) and Bergstrom and Varian (1985a, 1985b) note a similar invariance for Cournot games. ◇

4.3. *Sufficient conditions*

The result in Theorem 2 states that membership in the solution set of a particular SGM program is a necessary condition for an equilibrium aggregate. While useful, this does not tell us which solutions to the SGM program (if any) represent equilibrium aggregates. If payoff functions are strongly symmetric, however, we can conclude that any aggregate in the solution to some SGM program is an equilibrium outcome.

Consider the strategy profile $\overline{\mathbf{x}}$ uniquely defined by the following $n+1$ linear equations:

$$\overline{x}_i = \frac{1}{n}\overline{y} - (\gamma_i - \gamma_N),\ i \in N,$$
$$\overline{y} = \sum_{j\in N} \overline{x}_j.$$

Specializing Theorem 1 to the case of strongly-symmetric, linear aggregate games, this $\overline{\mathbf{x}}$ is an equilibrium if

$$\overline{y} \in \bigcap_{i\in N} \operatorname*{arg\,max}_{y\in X+\{\overline{y}-\overline{x}_i\}} \Lambda(y, \overline{y}),$$

where

$$\Lambda(y, \overline{y}) \equiv \alpha(y) + \langle y - \overline{y}(n-1)/n + \gamma_N, \beta(y)\rangle.$$

The following sufficiency result is an immediate consequence.

THEOREM 3 *If G is a strongly-symmetric linear aggregate game, then*

$$\left\{ \overline{y} \,\middle|\, \overline{y} \in \bigcap_{i\in N} \operatorname*{arg\,max}_{y\in X+\left\{\frac{(n-1)}{n}\overline{y}+\gamma_i-\gamma_N\right\}} \Lambda(y,\overline{y}) \right\} \subseteq \mathcal{E}_Y(G). \tag{4.6}$$

If ϕ is pairwise bijective, then

$$\left\{ \overline{y} \,\middle|\, \overline{y} \in \operatorname*{arg\,max}_{y\in X} \Lambda(y,\overline{y}) \right\} \subseteq \mathcal{E}_Y(G). \tag{4.7}$$

Combining Theorems 2 and 3 together we summarize our main results for the case of strongly-symmetric games in a single Proposition.

PROPOSITION 2 *If G is a strongly-symmetric linear aggregate game, then*

$$\left\{ \overline{y} \,\middle|\, \overline{y} \in \bigcap_{i\in N} \operatorname*{arg\,max}_{y\in X+\{(n-1)\overline{y}/n+\gamma_i-\gamma_N\}} \Lambda(y,\overline{y}) \right\} \subseteq \tag{4.8}$$
$$\mathcal{E}_Y(G) \subseteq \left\{ \phi(\overline{\mathbf{x}}) \,\middle|\, \overline{\mathbf{x}} \in \mathbf{X},\ \phi(\overline{\mathbf{x}}) \in \bigcap_{i\in M} \operatorname*{arg\,max}_{y\in X+\{\overline{y}-\overline{x}_i\}} \Lambda(y, \phi(\overline{\mathbf{x}})) \right\}.$$

Consequently, if the aggregate is pairwise-bijective, then

$$\left\{ \overline{y} \,\middle|\, \overline{y} \in \operatorname*{arg\,max}_{y\in X} \Lambda(y,\overline{y}) \right\} = \mathcal{E}_Y(G). \tag{4.9}$$

For pairwise-bijective, strongly-symmetric, linear aggregate games, (4.9) provides a *complete* characterization of the equilibrium aggregates via an SGM program. Because supply-function games and many intrinsic common agency games are members of this class, we can say quite a bit about equilibrium aggregates by constructing the invariant welfare function, Λ.

We conclude this section once again emphasizing that the characterization in Proposition 2 is fundamentally an algebraic result – no topological assumptions were used in the proofs. The aggregate concurrence principle and the necessary and sufficient characterizations derived from this central result follow from linear (rather than metric) properties.[23]

5. DISTRIBUTIONAL AGGREGATE GAMES

The applications which motivated this paper are symmetric, linear-aggregate games with infinite-dimensional strategy spaces in which the aggregate enters payoff functions through a probability distribution, which we refer to as the class of *distributional aggregate games* for short.

We are interested in characterizing the set of equilibrium aggregates for games in which players choose strategies from subset of real bounded, Borel-measurable functions, $X_i \subseteq \mathcal{B}_b(\Omega)$ defined on some Ω, and for which the aggregate function, $y = \sum_i x_i$, determines a probability distribution on Ω given by $H(\cdot \,|\, y) : \Omega \to [0, 1]$.[24] Hence, we are requiring that H is non-decreasing, has bounded variation, and is right-continuous.

Preferences for player i are represented by

$$u_i(x_i, y) = \alpha_i(y) + \int_\Omega (x_i(\omega) + \gamma_i(\omega)) dH(\omega \,|\, y).$$

Because H is a probability distribution, this integral is well-defined over the set of Borel-measurable, bounded functions.[25] Notice that H may be discontinuous; indeed such discontinuities may arise naturally in equilibrium.[26] By choosing $\langle f, g\rangle \mapsto \int f dg$ as the bilinear form, distributional payoff functions are evidently a special case of linear aggregate game payoffs, $u_i(x_i, y) = \alpha_i(y) + \langle x_i + \gamma_i, H(y)\rangle$. These distributional payoffs are weakly symmetric; they are strongly symmetric if α_i is independent of i and $X_1 = \cdots = X_n$. Thus our previous results for linear aggregate games are immediately available.

DEFINITION 4 *The game G is a* **distributional aggregate game** *if*

- *G is a linear, weakly symmetric aggregate game with strategy spaces $X_i \subseteq \mathcal{B}_b(\Omega)$, where $\mathcal{B}_b(\Omega)$ is the set of real, bounded Borel-measurable functions with domain Ω;*

[23] Of course, establishing that $\mathcal{E}(G)$ is non-empty or that the solution set to the SGM program is non empty, will typically require additional topological assumptions.

[24] We should emphasize at this stage that we allow for discontinuities in such probability distribution functions, but that we restrict attention to *probability* distributions rather than the larger class of generalized distribution functions associated with Sobolev and Schwartz.

[25] See, for example, Aliprantis and Border (2006, Theorem 11.8).

[26] To anticipate results which follow, in the case of common agency H has an upward jump discontinuity at any $q \in \mathcal{Q}$ which is chosen by a positive measure of agents (i.e., at all points of bunching).

- *preferences are representable by*

$$u_i(x_i, y) = \alpha_i(y) + \int_\Omega (x_i(\omega) + \gamma_i(\omega)) dH(\omega \,|\, y), \tag{5.1}$$

where $\alpha_i : X \to \mathbb{R}$*,* $\gamma_i \in \mathcal{B}_b(\Omega)$*.*

- G *is a* **strongly-symmetric distributional aggregate game** *if*

$$\alpha_1 = \cdots = \alpha_n = \alpha,$$

$$X_1 = \cdots = X_n = X.$$

The main import of distributional aggregate games is the influence of the aggregate through its effect on the distribution $H(\omega \,|\, y)$ in (5.1). Both the supply function game (with symmetric unit costs) and the common agency game are examples of strongly-symmetric distributional aggregate games.

It is useful to specialize our characterizations in Theorems 2 and 3 for the case of distributional aggregate games. To this end, redefine the aggregate-invariant welfare function for the class of distributional aggregate games as

$$\Lambda(y, \overline{y}) = \alpha(y) + \int_\Omega \left(y(\omega) - \left(\frac{n-1}{n} \right) \overline{y}(\omega) + \gamma_N(\omega) \right) dH(\omega \,|\, y), \tag{5.2}$$

where we recall that $\gamma_N = \frac{1}{n} \sum_i \gamma_i$. Because our distributional game is linear and symmetric (by definition), we may use the above construction of Λ and directly apply the various results for linear aggregate games.

5.1. *Supply function games*

Consider supply-function games. Let $\Omega = [0, b] \subset \mathbb{R}_+$ represent a bounded set of nonnegative prices and $\mathbb{R}$ the space of supplied output (recall negative supply is interpreted as demand). We assume that b is sufficiently large that at price $p = \hat{p}$, demand is zero regardless of θ. The strategy space is $X = \mathcal{B}_b([0, \hat{p}], \mathbb{R})$. For the moment, take the market maker's price-setting function as an exogenously given measurable function mapping from the demand state, $\theta \in \Theta$, and the sum of the firms' supply functions, $S = \sum_{i \in N} S_i$, to a market price, $\overline{p}_0 : \Theta \times X \to [0, \hat{p}]$. For any aggregate-supply function, S, the distribution of demand states induces a well-defined probability distribution over prices:

$$H(p \,|\, S) = \text{Prob}\left(\theta \in \Theta \,|\, \overline{p}_0(\theta, S) \leq p \right).$$

We assume that each firm's cost of production is constant and equal to c, which implies that a firm's expected profit function takes the form in (5.1):

$$\int_{[0,b]} S_i(p)(p - c) dH(p|S).$$

Because $S = \sum_i S_i$ is pairwise-bijective, we can immediately apply Proposition 2 to obtain a necessary and sufficient condition for any equilibrium aggregate supply function, $\overline{S}$:

$$\overline{S} \in \underset{S \in \mathcal{B}_b([0,\hat{p}])}{\arg\max} \int_{[0,b]} \big(S(p)(p - c) + (n-1)(S(p) - \overline{S}(p))(p - c) \big) \, dH(p \,|\, S). \tag{5.3}$$

Stated as an SGM program, we have a simple decomposition into two economic influences. The first term in the objective is industry profit. The second term characterizes the non-cooperative effect: for $n = 1$ it is absent, and as n grows its weight relative to industry profit increases. Except for the assumption of symmetric marginal costs, we have placed very little structure on the primitives of the game but may nonetheless conclude that *every* equilibrium supply function, $\overline{S}$, must be a solution to the the (self-generated) program in (5.3). In particular, we have said nothing about the character of $\overline{p}_0$ (which is embedded in H), the relationship between θ and market demand, or the distribution of demand states.

With additional structure, we can say more about the non-cooperative effect. Suppose that the market maker's objective is to select a market-clearing price. Thus, $\overline{p}_0(\theta, S)$ is a pointwise solution to the equation, $D(p, \theta) = S(p)$. If no solution exists, we assume that the market maker dictates that no trades take place; we denote the set of no-trade demand states as $\Theta_\emptyset$ and its complement by $\overline{\Theta} = \Theta/\Theta_\emptyset$. If multiple solutions exist, we assume the market-maker chooses the lowest market-clearing price; this selection maximizes the volume traded. In addition, we assume that D is strictly increasing in θ, which implies that $\overline{p}_0$ is a strictly increasing function on $\overline{\Theta}$, with the inverse function denoted $\overline{\theta}_0(p, S)$. If $F : \Theta \to [0, 1]$ is the distribution function over demand states, then the corresponding distribution of prices is

$$H(p|S) = F(\overline{\theta}_0(p, S)).$$

Because $D(\overline{p}_0(\theta), \theta) = \overline{S}(\overline{p}_0(\theta))$ on $\overline{\Theta}$ by construction, we can translate the previous self-generating program into an equivalent problem over the dual space of price functions:

$$\text{(5.4)} \qquad \overline{p} \in \underset{p \in \mathcal{B}_b(\overline{\Theta}, [0, \hat{p}])}{\arg\max} \int_{\overline{\Theta}} D(p(\theta), \theta)(p(\theta) - c) dF(\theta) + \\ (n-1) \int_{\overline{\Theta}} (D(p(\theta), \theta) - \overline{S}(p(\theta)))(p(\theta) - c) dF(\theta).$$

Again, the first term represents industry profit and the second term contains the additional strategic effect.

To find the equilibrium points of the supply function game, we need to find each self-generating fixed point, $\overline{p}$ (or $\overline{S}$, alternatively). We restrict attention to equilibria for which $\Theta_\emptyset$ has measure zero and require that demand functions are (once) differentiable with D_p strictly increasing in θ. The latter monotonicity requirement insures that for any $\overline{S}$, the pointwise optimizing p of the integrand is strictly increasing in θ. If $\overline{p}$ is the optimal price function associated with $\overline{S}$, because it is strictly increasing we can replace $\overline{S}(p)$ with $D(p, \overline{p}^{-1}(p))$ in the integrand. Pointwise maximization yields the following differential equation for the market price:

$$\text{(5.5)} \qquad \frac{d\overline{p}(\theta)}{d\theta} = \frac{(n-1)(\overline{p}(\theta) - c) D_\theta(\overline{p}(\theta), \theta)}{(\overline{p}(\theta) - c) D_p(\overline{p}(\theta), \theta) + D(\overline{p}(\theta), \theta)}.$$

The solution set of this differential equation is exactly the set of market-clearing equilibria (i.e., equilibria for which the market clears with probability one).

Rather than finding explicit solutions to the differential equation, we investigate the properties of equilibrium price functions directly from the optimization program.

Define $p^b(\theta) \equiv c$ as the Bertrand (perfectly competitive) price function and $p^c(\theta)$ as the Cournot price function, which satisfies (pointwise) the equation

$$D_p(p^c(\theta), \theta)(p^c(\theta) - c) + \frac{1}{n} D(p^c(\theta), \theta) \equiv 0.$$

Restricting attention to equilibria with strictly increasing supply functions, we can be assured that $\overline{S}$ is differentiable almost everywhere. The pointwise derivative of the integrand in (5.4) with respect to price in equilibrium is

$$\left(nD_p(\overline{p}(\theta), \theta) - (n-1)\overline{S}_p(\overline{p}(\theta))\right)(\overline{p}(\theta) - c) + D(\overline{p}(\theta), \theta).$$

If, in addition, $D_{\theta\theta} \geq 0$, then the integrand is strictly quasi-concave in p, and so the first-order solution characterizes the optimum. Evaluated at $p^b(\theta) = c$, this derivative is equal to $D(c, \theta) \geq 0$, indicating that $p(\theta) \geq p^b(\theta)$. Evaluating the first-order equation at $p^c(\theta)$, we instead obtain a derivative of $-(n-1)\overline{S}_p(p^c(\theta)) < 0$. Thus, in any equilibrium with increasing aggregate supply, the equilibrium price function must lie between the Bertrand and Cournot solutions:

$$p^b(\theta) \leq \overline{p}(\theta) < p^c(\theta).$$

If we restrict the strategy space of each firm to contain only nondecreasing supply schedules, then any resulting equilibrium price function would satisfy these bounds. Absent such a restriction, a priori, we cannot rule out the possibility that aggregate supply is decreasing in some region, only that it cannot decrease too fast: $\overline{S}_p(p) - D_p \geq 0$. Wherever a non-increasing equilibrium exists, the upper bound on the price function must exceed $p^c(\theta)$ in the decreasing supply regions. Note that the derivative of (5.4) evaluated at the monopoly price, which we denote $p^m(\theta)$, is nonnegative, $(n-1)(D_p - \overline{S}_p)(p^m - c) \leq 0$. We conclude that $p^m(\theta)$ is an upper bound on $\overline{p}(\theta)$ for the general case.

For the case in which firms have different (but constant) marginal costs, c_i, solving explicitly for the set of Nash equilibria involves finding all the solutions to a system of differential equations. Unfortunately, the SGM program approach is not entirely satisfactory either. Evaluated at Cournot prices, the marginal effect of price simplifies to

$$-(n-1)\overline{S}_p(p^c(\theta))(p^c(\theta) - c_N) - \sum_{i \in N} \overline{S}_{i,p}(p^c(\theta))(c_i - c_N),$$

where c_N is the average unit cost. Notice that the new term is the covariance of c_i and $S_{i,p}(p^c)$ across the firms. The aggregate concurrence principle tells us that any two players must have the same marginal returns with respect to the aggregate components. It follows that whenever high-cost firms have lower market share than low-cost firms, then high-cost firms must also exhibit greater supply responses. We may conclude, therefore, that if high cost firms have weakly smaller markets shares in equilibrium, then the covariance must be is positive and $\overline{p}(\theta) < p^c(\theta)$.

5.2. *Intrinsic common-agency games*

The class of games which motivated this paper are intrinsic common agency games with public contracts.[27] Each principal i has a bounded, continuous benefit function

[27]A full treatment of the issues involved is outside the scope of this note and is developed in Martimort and Stole (2009,2010).

of $v_i : \mathcal{Q} \to \mathbb{R}$ which gives the return associated to the agent's choice of $q \in \mathcal{Q}$. We assume that $\mathcal{Q}$ is compact. To motivate the agent, a principal may offer a transfer schedule, $\mathcal{T}_i \in \mathcal{T}$, which is an agreement to pay the agent $T_i(q)$ whenever $q \in \mathcal{Q}$ is chosen. We take the set of available contracts, $\mathcal{T}$, to be the the set of bounded, upper-semicontinuous functions on $\mathcal{Q}$. Because upper-semicontinuous functions are Borel measurable, $\mathcal{T} \subseteq \mathcal{B}_b(\mathcal{Q})$.

Principal i's payoff when the agent accepts the principal's offer and chooses q is $v_i(q) - T_i(q)$. If the agent chooses not to participate, then action $q_\emptyset$ is implemented by default and principal i earns $v_i(q_\emptyset)$, which we normalize to 0.

We assume that the agent has private information, θ, distributed on Θ with distribution $F(\theta)$. An agent of type θ has a bounded, continuous benefit function, $u(\cdot, \theta) : \mathcal{Q} \to \mathbb{R}$, and a net payoff

$$u(q, \theta) + \sum_{i \in N} T_i(q),$$

when accepting all of the contracts and choosing $q \in \mathcal{Q}$. We take common agency to be *intrinsic* in the game; i.e., the agent must either accept all contract offers (and become a common agent to all n principals) or choose not to participate. In the latter case, the agent's reservation utility is normalized to $u(q_\emptyset, \theta) = 0$.[28]

The agent's strategy is a function $q_0(\theta, T)$ which depends upon type and the aggregate contract offer, T. We are assuming in this application that the agent chooses q without regard to the various contributions of the principals to T.[29] For any aggregate contract offer, $T = \sum_{i \in N} T_i$, the agent's best-response correspondence is non-empty, compact-valued and upper semi-continuous. Denote the correspondence by

$$Q_0(\theta, T) \equiv \arg\max_{q \in Q} u(q, \theta) + T(q),$$

and let $q_0(\theta, T) \in Q_0(\theta, T)$ be any measurable selection. Fixing q_0, we can define the distribution function[30]

$$H(q \,|\, T) \equiv \text{Prob}\left(\theta \in \Theta \,|\, q_0(\theta, T) \leq q\right).$$

Thus, taking q_0 as given, we have an n-player distributional aggregate game in which player i's payoff is represented by

$$u_i(T_i, T) = \int_{\mathcal{Q}} (v_i(q) - T_i(q)) dH(q \,|\, T).$$

If we characterize the set of equilibrium aggregates of this n-player game for every selection $q_0 \in Q_0$, we will have characterized the entire set of equilibrium aggregates in the original $n+1$ player game.

---

[28]Note that we have not imposed any restrictions on the agent's benefit function u other than continuity and boundedness; in particular, we do not impose single-crossing or monotonicity in type. Thus, the normalization of $u(q_\emptyset, \theta) = 0$ to a type-independent reservation utility is without loss of generality as it can be embedded in the original benefit function: $\tilde{u}(q, \theta) = u(q, \theta) - u(q_\emptyset, \theta)$.

[29]In other words, if two profiles $(T_1, \ldots, T_n)$ and $(T_1', \ldots, T_n')$ both sum to the same aggregate $T = \sum T_i = \sum T_i'$, then the agent's equilibrium responses are assumed to be identical.

[30]The methodological approach of using a distribution function over output which is a function of nonlinear tariffs is developed in Wilson (1993). Wilson refers to such distributions as "demand profiles".

There are many equilibria to this game. One trivial equilibrium is for two of the n principals to set transfers sufficiently negative such that every agent prefers to choose $q_\emptyset$. Such "non-participation" is an equilibrium, possibly Pareto-dominated by all other equilibria. There are also more interesting equilibria that do not appear to suffer from such coordination failures. We wish to characterize the entire set of equilibria, including possibly equilibria with discontinuous transfer functions.

Because the n-player game is a distributional game, it is also aggregate-invariant and we can construct a weighted payoff function

$$\Lambda(T, \overline{T}) = \frac{1}{n} \int_{\mathcal{Q}} \left(\sum_{i \in N} v_i(q) - T(q) + (n-1)(T(q) - \overline{T}(q)) \right) dH(q \,|\, T) \tag{5.6}$$

and apply our previous results in Theorems 2 and 3.[31]

There are two cases to consider, depending upon whether or not the aggregate is pairwise bijective on $\mathcal{T}$. Consider first the case in which $\mathcal{Q}$ is a finite collection of actions, and thus a contract in $\mathcal{T}$ is a finite sequence of real numbers and the agent's best-response correspondence $Q_0(\theta, T)$ is non-empty. The aggregate function, $T = \sum_{i \in N} T_i$ is pairwise-bijective because $\mathcal{T}$ is closed under addition and subtraction. As such, Proposition 1 applies squarely to the case of common agency games and we can conclude:

Corollary 3 *In the common agency game, G, if the action space $\mathcal{Q}$ is finite, then*

$$\left\{ \overline{T} \,\middle|\, \overline{T} \in \arg\max_{T \in \mathcal{T}} \Lambda(T, \overline{T}) \right\} = \mathcal{E}_Y(G). \tag{5.7}$$

The second case arises when $\mathcal{Q}$ is infinite. Specifically, suppose that $\mathcal{Q}$ is a compact subset of $\mathbb{R}^k$, endowed with the usual topology. Because the difference between two upper semicontinuous functions is not necessarily upper semi-continuous within this larger space, the aggregate function is not pairwise bijective (i.e., $\mathcal{T}$ is not closed under subtraction).[32] That said, we may still apply Theorem 3 to this non-bijective setting. Specializing the equilibrium construction from linear games to the present setting, we have

$$\overline{T}_i = \frac{1}{n}\overline{T} + v_i - \frac{1}{n}\sum_{j \in N} v_j,$$

$$\mathcal{T} + \{\overline{T}_{-i}\} = \mathcal{T} + \left\{ \frac{(n-1)}{n}\overline{T} + \frac{1}{n}\sum_{j \in N} v_j - v_i \right\}.$$

Because $\overline{T} \in \mathcal{T}$ and the principals' benefit functions are continuous, it follows that $\overline{T}_i \in \mathcal{T}$ and, consequently, $\overline{T} \in \mathcal{T} + \{\overline{T}_{-i}\}$. Revealed preference requires

$$\overline{T} \in \arg\max_{T \in \mathcal{T}} \Lambda(T, \overline{T}) \implies \overline{T} \in \arg\max_{T \in \mathcal{T} + \{\overline{T}_{-i}\}} \Lambda(T, \overline{T}).$$

[31] Notice that (5.6) is more generally true for any intrinsic common agency contracting game. When moral hazard is the source of misaligned incentives, one can reinterpret $H(q \,|\, T)$ to be the equilibrium distribution of outcomes given the agent's optimal choice of effort when presented with the incentive contract T. The content of (5.6) is analogous to the result in Lemma 1 of Bernheim and Whinston (1986a).

[32] We are grateful to David Rahman for bringing this important detail to our attention and leading us to a more thorough understanding of our aggregate characterization for infinite-dimensional games.

From Theorem 3, we conclude that the constructed contracts are in fact equilibrium strategies.

COROLLARY 4 *For the common agency game, G,*

$$\left\{ \overline{T} \;\middle|\; \overline{T} \in \arg\max_{T \in \mathcal{T}} \Lambda(T, \overline{T}) \right\} \subseteq \mathcal{E}_Y(G).$$

Unlike the case in which $\mathcal{Q}$ is finite, proving the reverse inclusion requires that we exploit the metric properties in the game. This proof is accomplished by showing that the difference between any pair of upper-semicontinuous functions, $T_i = T - \overline{T}_{-i}$, can be approximated from below by another upper-semicontinuous function, $T_i^\varepsilon \nearrow T_i$, such that $u_i(T_i^\varepsilon, T_i^\varepsilon + \overline{T}_{-i}) \to u_i(T_i, T_i + \overline{T}_{-i})$. In words, the aggregate T is pairwise-bijective on a dense subset of $\mathcal{T}$ for which payoffs are sequentially continuous. From this technical result, we have our main conclusion for intrinsic, common agency games.

PROPOSITION 3 *In the common agency game, G, if the action space $\mathcal{Q}$ is a compact subset of $\mathbb{R}^k$, then*

$$\left\{ \overline{T} \,|\, \overline{T} \in \arg\max_{T \in \mathcal{T}} \Lambda(T, \overline{T}) \right\} = \mathcal{E}_Y(G).$$

Returning to our SGM program, we may define $q(\theta) \equiv \overline{q}_0(\theta, T)$ and $U(\theta) \equiv u(q(\theta), \theta) + T(q(\theta))$ and restate the program in its dual form:

$$(5.8) \qquad (\overline{q}, \overline{U}) \in \arg\max_{(q,U) \in \mathcal{IC}} \int_\Theta \left\{ \left(\sum_{i \in N} v_i(q) + u(q, \theta) \right) + (n-1)(u(q,\theta) + \overline{T}(q)) - nU \right\} dF(\theta),$$

where $\mathcal{IC}$ is the set of all (q, U) pairs that are implementable by some aggregate T, and $\overline{T}$ is the aggregate transfer that implements $(\overline{q}, \overline{U})$. The dual program in (5.8) characterizes the *entire* set of Nash equilibrium pairs $(\overline{q}, \overline{U})$ for the common agency games. Remarkably, the characterization holds regardless of the dimension of Θ or the presence of single-crossing.

Consider as an example the equilibrium in which the agent rejects all contracts because two or more principals make undesirable offers. In the above SGM program, this corresponds to an equilibrium $\overline{T}$ that is extremely large and negative for all $q \neq q_\emptyset$. The solution to the self-generating maximization program therefore is to implement $q_\emptyset$.

Next consider the more economically interesting set of smooth equilibria with full participation. If they exist, the Envelope Theorem together with (5.8) implies that the set of all smooth equilibrium outcomes is characterized by

$$(\overline{q}, \overline{U}) \in \arg\max_{(q,U) \in \mathcal{IC}} \int_\Theta \left\{ \left(\sum_{i \in N} v_i(q) + u(q, \theta) \right) - nU \right\} dF(\theta).$$

This simpler program is no longer self-generating. We may conclude that a smooth equilibrium exists if and only if this maximization program has a solution. Moreover, the smooth program suggests that it is *as if* there is a single principal maximizing collective surplus less the agent's information rent *multiplied by* n. It is tantamount to an n-fold marginalization of rents. A single firm collectively representing the interests of all principals would solve a similar program with n replaced with 1. Unsurprisingly, competition among the principals (weakly) reduces their payoffs relative to what they would obtain if they could maximize payoffs collectively. More surprising, however, the program implies that expected consumer surplus is also lower under competition relative to the collective outcome.

PROPOSITION 4 *In any intrinsic common-agency game, every smooth equilibrium is less socially efficient and exhibits lower agent payoffs in expectation relative to the outcome in which the principals collude in contract design.*

This conclusion is derived entirely from the assumption that the program has a solution and the nature of the objective function without any need to calculate the solution explicitly. When Θ is multi-dimensional or single-crossing fails, this is particularly valuable as closed-form, analytic solutions are typically elusive.

There are still other equilibria that are not smooth and exhibit agent participation. For these equilibria, nonsmooth analysis must be used to characterize the solutions to the SGM programs, which we pursue in companion papers, Martimort and Stole (2010a, 2010b). The approach of aggregate representation, in tandem with these tools, continues to describe the characteristics of equilibrium aggregates.

As a final extension, one can easily augment this model of intrinsic agency to include incomplete information for the principals, similar to the analysis in the beauty-contest game. Sufficiency results are no longer available, but the necessary condition in Theorem 2 continues to apply. We take up this brief extension in the Appendix.

5.3. *Delegated common-agency games: A special case*

In both the "Supply Function" and "Intrinsic Common Agency" games, we made use of the pairwise bijective aggregate to give necessary and sufficient conditions for set of equilibria aggregates. In many applications of multi-principal contracting, however, the agent has the ability to select a subset of contract offers. Consider, for example, the case of lobbyists influencing politicians through contributions. In these *delegated* common-agency games, transfers are effectively constrained to be positive as the agent can reject any disadvantageous offer; the aggregate transfer is not pairwise bijective. That said, using an SGM program we can establish a set of necessary conditions for equilibrium aggregates which in many cases will give a satisfactory answer to the nature of any contractual inefficiencies.

We consider the special case in which $q \in \mathcal{Q}$ is a one-dimensional choice from a convex, compact interval $\mathcal{Q} \subset \mathbb{R}$. In addition, we suppose that each principal has strictly monotonic preferences for q. For $i \in \mathcal{A}$, we assume that $v_i'(q) > 0$; for $i \in \mathcal{B}$, we assume $v_i'(q) < 0$. The agent's benefit functions is assumed to be continuous and quasi-bilinear in (q, θ):

$$u(q, \theta) = \theta q + w(q) + z(\theta),$$

where w is a strictly concave function of q. Lastly, we assume that θ is distributed on $\Theta \equiv [0,1]$ with continuous density function f and differentiable distribution function F. Define the following principal-specific inverse hazard rate:

$$H_i(\theta) \equiv \begin{cases} \frac{F(\theta)-1}{f(\theta)}, & \text{for } i \in \mathcal{A}, \\ \\ \frac{F(\theta)}{f(\theta)}, & \text{for } i \in \mathcal{B}. \end{cases}$$

We make the standard inverse hazard rate assumption that H_i is nondecreasing on Θ.

Without loss of generality, we can restrict attention to equilibria in which each principal's transfer function is nonnegative. For any equilibrium of the delegated contracting game, $\{\overline{q}_0, \overline{T}_1, \ldots, \overline{T}_n\}$, we may define the agent's outside option from rejecting principal i's contract as

$$\overline{U}_{-i}(\theta) \equiv \max_{q \in \mathcal{Q}} \ u(q,\theta) + \overline{T}_{-i}(q),$$

and the agent's corresponding choice $\overline{q}_{-i}(\theta) = \overline{q}_0(\theta, \overline{T}_{-i})$ as

$$\overline{q}_{-i}(\theta) \in \arg\max_{q \in \mathcal{Q}} \ u(q,\theta) + \overline{T}_{-i}(q).$$

In equilibrium, therefore, each principal i chooses a contract offer, $T_i \geq 0$, to induce (q, U) and maximize

$$\int_\Theta \Big\{ \big(v_i(q(\theta)) - v_i(\overline{q}_{-i}(\theta)\big) + \big(u(q(\theta),\theta) + \overline{T}_{-i}(q(\theta)) - \overline{U}_{-i}(\theta)\big) - \big(U(\theta) - \overline{U}_{-i}(\theta)\big)\Big\} f(\theta) d\theta, \tag{5.9}$$

subject to the constraint that $U(\theta) \geq \overline{U}_{-i}(\theta)$ for all $\theta \in \Theta$.

It is helpful at this stage to introduce a mild refinement on the set of equilibria. We say that an equilibrium is **monotone** if $\overline{T}_i$ is everywhere nondecreasing for $i \in \mathcal{A}$ and everywhere nonincreasing for $i \in \mathcal{B}$. Given that a principal $i \in \mathcal{A}$ prefers greater levels of q, paying the agent weakly increasing amounts for higher q is a natural feature to expect in equilibrium.[33] Given our restriction to monotone equilibria, principal i's program in (5.9) can be greatly simplified. In particular, in any monotone equilibrium, for any $i \in \mathcal{A}$, the set of types for whom $\overline{U}(\theta) = \overline{U}_{-i}(\theta)$ must be a lower (possibly degenerate) interval, $[0, \overline{\theta}_i]$. The reverse is true for principal $i \in \mathcal{B}$. Using the standard result that (q, U) is incentive compatible iff q is nondecreasing and $U'(\theta) = q(\theta)$ at all points for which q is continuous, we can reformulate the program for principal $i \in \mathcal{A}$ as

$$\max_{q,\theta_i} \int_{\theta_i}^1 \Big\{ v_i(q(\theta)) - v_i(\overline{q}_{-i}(\theta)) + u(q(\theta),\theta) + \overline{T}_{-i}(q(\theta)) + H_i(\theta)(q(\theta) - \overline{q}_{-i}(\theta))\Big\} f(\theta) d\theta,$$

[33]Indeed, we are not even certain that this monotonicity requirement is a refinement per se, as we have been unable to find any non-monotone equilibria.

subject to q nondecreasing. Given that H_i is nondecreasing, this program has a convenient pointwise solution. Define

$$\overline{J}_i(q,\theta) = v_i(q) + u(q,\theta) + \overline{T}_{-i}(q) + H_i(\theta).$$

One may then write the pointwise optimal choice for principal $i \in \mathcal{A}$ as

$$q(\theta) \in \arg\max_{q \in \mathcal{Q}} \; \max\left\{\overline{J}_i(q,\theta) - \overline{J}_i(\overline{q}_{-i}(\theta),\theta), 0\right\}.$$

Notice that a similar statement holds for principal $i \in \mathcal{B}$.

$$q(\theta) \in \arg\max_{q \in \mathcal{Q}} \; \min\left\{\overline{J}_i(q,\theta) - \overline{J}_i(\overline{q}_{-i}(\theta),\theta), 0\right\}.$$

Now we use the aggregate concurrence principle. Summing the pointwise objectives, we conclude that in any monotone equilibrium, the allocation $\overline{q}$ must solve the following SGM program:

$$\overline{q}(\theta) \in \arg\max_{q \in \mathcal{Q}} \; \sum_{i \in \mathcal{A}} \max\left\{\overline{J}_i(q,\theta) - \overline{J}_i(\overline{q}_{-i}(\theta),\theta), 0\right\} + \sum_{i \in \mathcal{B}} \min\left\{\overline{J}_i(q,\theta) - \overline{J}_i(\overline{q}_{-i}(\theta),\theta), 0\right\}.$$

As a final simplification, assume that each principal's preference for q is a linear function, $v_i(q) = s_i q$. Given linearity, the SGM program may be further reduced to

$$\overline{q}(\theta) \in \arg\max_{q \in \mathcal{Q}} \; \sum_{i \in \mathcal{A}} \max\left\{s_i - \frac{1-F(\theta)}{f(\theta)}, 0\right\} q + \sum_{i \in \mathcal{B}} \min\left\{s_i + \frac{F(\theta)}{f(\theta)}, 0\right\} q + u(q,\theta) + (n-1)\left(u(q,\theta) + \overline{T}(q)\right).$$

This formulation is evocative of Myerson's (1980) optimal auction design. Indeed, in the case of smooth transfer equilibria, we can use the Envelope Theorem to eliminate the final terms. We have the following proposition.

PROPOSITION 5 *In the delegated common-agency game with linear principal preferences, all monotone, smooth equilibria have allocations satisfying*

$$\overline{q}(\theta) \in \arg\max_{q \in \mathcal{Q}} \left(\sum_{i \in \mathcal{A}} \max\left\{s_i - \frac{1-F(\theta)}{f(\theta)}, 0\right\} + \sum_{i \in \mathcal{B}} \min\left\{s_i + \frac{F(\theta)}{f(\theta)}, 0\right\} \right) q + u(q,\theta).$$

A detailed proof of the above proposition, including a constructive proof of existence, is provided in our companion paper Martimort and Stole (2010a). The mathematical intuition, however, can be understood in terms of SGM programs for aggregate games presented above. Economically, the proposition indicates that there is a multiple marginalization, just as in the case of intrinsic agency.

To interpret the result, suppose that the game is one of private provision of public goods. In this case, $N = \mathcal{A}$, and a principal is active precisely when $f(\theta)s_i >$

$1 - F(\theta)$. For any type agent for whom m principals are active, there will be an m-fold marginalization of information rents. One can further show that if $\overline{q}_S(\theta)$ is the allocation chosen in an intrinsic game with principals $S \subseteq N$ (and $\overline{q}_\emptyset(\theta) = 0$), then the equilibrium allocation in the delegated game is

$$\overline{q}(\theta) = \max_{S \subseteq N} \overline{q}_S(\theta).$$

In the case of conflict among the principals, the result is more subtle. Suppose that $n = 2$ and that each principal is a lobbyist trying to influence a politician's choice of q. Let $s_1 > 0 > s_2$ represent the preference conflict and θ represent the agent's preference for q where we now assume θ is uniformly distributed over the set $[-\Delta, \Delta]$. Martimort and Stole (2010a) establish that the monotone, smooth equilibrium is characterized by two critical types, $\overline{\theta}_1$ and $\overline{\theta}_2$, with the property that an agent with type $\theta \in [-\Delta, \min\{\overline{\theta}_1, \overline{\theta}_2\})$ is influenced only by principal 2, an agent $\theta \in (\max\{\overline{\theta}_1, \overline{\theta}_2\}, \Delta]$ is influenced only by principal 1, and if $\overline{\theta}_1 < \overline{\theta}_2$ (which will be the case for Δ sufficiently small), then types in the interval $(\overline{\theta}_1, \overline{\theta}_2)$ are influenced by both principals. In the dual-influence region, the agent's action is doubly-distorted away from efficient allocation.

6. AGGREGATE GAMES AND POTENTIAL GAMES

We conclude this paper with a brief review of the connections between various notions of potential games and the aggregate game concepts developed in this paper. The connection between aggregate and potential games is loose, but their contrast serves to illuminate the key features of aggregate games and their characteristic self-generating programs.

Recalling Monderer and Shapley (1996), a game $G = \{N, (X_i)_{i \in N}, (u_i)_{i \in N}\}$ is said to be an **(exact) potential game** if there exists a potential function $\mathcal{P} : \mathbf{X} \to \mathbb{R}$ such that for every player $i \in N$ and for every profile $\mathbf{x}_{-i} \in \mathbf{X}_{-i}$,

$$u_i(x_i, \mathbf{x}_{-i}) - u_i(x_i', \mathbf{x}_{-i}) = \mathcal{P}(x_i, \mathbf{x}_{-i}) - \mathcal{P}(x_i', \mathbf{x}_{-i}).$$

G is said to be an **ordinal potential game** if there exists a function $\mathcal{P} : \mathbf{X} \to \mathbb{R}$ such that for every player $i \in N$ and for every profile $\mathbf{x}_{-i} \in \mathbf{X}_{-i}$,

$$u_i(x_i, \mathbf{x}_{-i}) - u_i(x_i', \mathbf{x}_{-i}) > 0 \iff \mathcal{P}(x_i, \mathbf{x}_{-i}) - \mathcal{P}(x_i', \mathbf{x}_{-i}) > 0.$$

Clearly every exact potential game is an ordinal potential game.

An even weaker notion than an ordinal potential is that of a best-response potential developed in Voorneveld (2000). A game G is a **best-response potential game** if there exists a function $\mathcal{P} : \mathbf{X} \to \mathbb{R}$ such that for every player $i \in N$ and for every profile $\mathbf{x}_{-i} \in \mathbf{X}_{-i}$,

$$\arg\max_{x_i \in X_i} u_i(x_i, \mathbf{x}_{-i}) = \arg\max_{x_i \in X_i} \mathcal{P}(x_i, \mathbf{x}_{-i}).$$

Morris and Ui (2004) use a best-response equivalence relation to construct and characterize this quotient space for more general games. In the sense of such equivalence classes, G is a best-response potential game precisely when it is best-response equivalent to an "identical-interest" or "coordination" game,

$$\hat{G} = \{N, (X_i)_{i \in N}, (\mathcal{P})_{i \in N}\}.$$

Of course, the sets of Nash equilibria for G and $\hat{G}$ coincide whenever such an equivalence holds.

In the most general definition of an aggregate game, the defining requirement is the existence of an aggregate function $\phi : \mathbf{X} \to Y$ such that preferences can be represented as $u_i : X_i \times Y \to \mathbb{R}$. The equilibrium set of every best-response potential game (and hence every ordinal and exact potential game) coincides with the equilibrium set of the identical-interest game $\hat{G}$. Note that a coordination game is itself an aggregate game upon substituting $\phi(\mathbf{x}) = \mathcal{P}(\mathbf{x})$ and $u_i(x_i, \phi(\mathbf{x})) = \mathcal{P}(\mathbf{x})$. We may thus conclude that every best-response potential game G has an aggregate-game representation $\hat{G}$ that is faithful to the set of Nash equilibria. The class of aggregate games is therefore decidedly larger than the class of best-response potential games and the principle of aggregate concurrence in Theorem 1 has broader applications.[34]

It is illuminating to draw out the implications of the aggregate concurrence principle when applied to a best-response potential game. Clearly, any $\overline{\mathbf{x}} \in \arg\max_{\mathbf{x}\in\mathbf{X}} \mathcal{P}(\mathbf{x})$ is a Nash equilibrium. The set of $\mathcal{P}$ maximizers, however, is possibly a proper subset of the equilibria to G. Indeed, Theorem 1 informs us that the entire set of Nash equilibrium aggregate values is characterized by the self-generating program

$$\mathcal{P}(\overline{\mathbf{x}}) \in \bigcap_{i\in N} \arg\max_{y\in\mathcal{P}(X_i,\overline{\mathbf{x}}_{-i})} y,$$

or more simply

$$\mathcal{P}(\overline{\mathbf{x}}) = \max_{\mathbf{x}\in\mathbf{X}} \frac{1}{n} \sum_{i\in N} \mathcal{P}(X_i, \overline{\mathbf{x}}_{-i}).$$

Stated this way, the possibility of coordination failures arising in equilibrium is evident.

The previous argument suggests that even in linear aggregate games for which there exists a common payoff $\Lambda(y, \overline{y})$ that represents the players' preferences of $y \in X$ relative to $\overline{y}$, we cannot conclude that the there is a common payoff function (a best-response potential), $\mathcal{P}(y)$, that (i) depends only upon y and (ii) faithfully represents *all* of the equilibria in the original game. Typically, for $\Lambda(y, \overline{y})$ to be written independently of $\overline{y}$, either an equilibrium refinement or a restriction on strategy spaces must be introduced. For example, recalling the case of intrinsic common agency, one can restrict attention to equilibria in smooth tariffs (alternatively, we could restrict the strategy spaces in the original game). In this case, (5.8) reduces to the simpler program

$$(\overline{q}, \overline{U}) \in \arg\max_{(q,U)\in\mathcal{IC}} \mathcal{P}(q, U),$$

where

$$\mathcal{P}(q, U) = \int_{\Theta} \left\{ \left(\sum_{i\in N} v_i(q) + u(q, \theta) \right) - nU \right\} dF(\theta).$$

[34]The smaller classes of linear aggregate games and exact potential games are non-nested. First, it is well known that the game of Cournot is an exact potential game only if market demand curve is linear; the only requirement on cost functions is that they are differentiable. The Cournot game is a linear aggregate game, however, only if each firm's cost function is affine, but there are no similar restrictions on demand; indeed, demand can be discontinuous.

Thus, the intrinsic agency game restricted to differentiable tariffs is a best-response potential game, but we are unable to make such a statement about the entire class of intrinsic common agency games. One of the contributions of the present paper is to demonstrate that regardless, aggregate games have a simple SGM characterization derived from preferences over aggregate strategies.

APPENDIX

Proof of Lemma 1: The right-hand side represents player i's strategic problem as an equivalent two-stage maximization program over the space of aggregates. First, for any given aggregate y, the player chooses the best x_i from the set of choices that implement y given $\overline{\mathbf{x}}_{-i}$. Thus, $x_i \in \phi_i^{-1}(y, \overline{\mathbf{x}}_{-i})$. Denote the value function for this step as

$$U_i(y, \overline{\mathbf{x}}_{-i}) \equiv \max_{x_i \in \phi_i^{-1}(y, \overline{\mathbf{x}}_{-i})} u_i(x_i, y).$$

If ϕ is pairwise-injective, then this step is trivial as $\phi_i^{-1}(y, \mathbf{x}_{-i})$ is therefore a single-valued correspondence and $U_i(y, \overline{\mathbf{x}}_{-i}) = u_i(\phi_i^{-1}(y, \overline{\mathbf{x}}_{-i}), \overline{\mathbf{x}}_{-i})$. In the second stage, the player chooses the best aggregate in the set of all implementable aggregates, given $\overline{\mathbf{x}}_{-i}$; that is, y is chosen from the set $\phi(X_i, \overline{\mathbf{x}}_{-i})$ to maximize $U_i(y, \overline{\mathbf{x}}_{-i})$. *Q.E.D.*

Proof of Proposition 1: Applying Corollary 2, if $\overline{y}$ is an equilibrium aggregate, it must maximize $\Lambda(y, \overline{y})$ over the set of feasible aggregates given the equilibrium aggregate $\overline{y}$ strategy profile,

$$y \in \mathcal{Y}(\overline{y}) \equiv \bigcap_{i \in N} \{y \mid \exists z_i \in X_i \text{ s.t. } y = \overline{y} + z_i\}$$

where we have used the property that $X_i = X_i - \{\overline{x}_i\}$. In words, the set of feasible aggregates is the set of variations from $\overline{y}$ that are restricted to a single signal.

Because $\Lambda(\cdot, \overline{y})$ is strictly concave and Θ is a convex set, any solution to the SGM program

$$\overline{y} \in \arg\max_{y \in \mathcal{Y}(\overline{y})} \Lambda(y, \overline{y}),$$

must satisfy the first-order conditions for each permissible variation. It is thus necessary that that for each $i \in N$, the first-order condition for the variation $y(\mathbf{s}) = \overline{y}(\mathbf{s}) + z_i(s_i)$ must be satisfied at $z_i = \mathbf{0}$. Thus, for each i,

$$\mathbf{0} \in \arg\max_{z_i \in X_i} \Lambda(\overline{y} + z_i, \overline{y}).$$

In the context of beauty-contest games, the expression $\Lambda(\overline{y} + z_i, \overline{y})$ specializes to

$$-\int_{\Theta \times \mathbf{\Sigma}} \left\{ \frac{1}{2} (\overline{y}(\mathbf{s}) + z_i(s_i))^2 (\kappa + (n-1)) \right.$$
$$\left. - (\overline{y}(\mathbf{s}) + z_i(s_i)) (\overline{y}(\mathbf{s})(n-1) + \kappa\, \theta_0) \right\} dG(\theta_0, \mathbf{s}).$$

Eliminating all terms that do not involve z_i, we have the simpler requirement that

$$\mathbf{0} \in \arg\min_{z_i \in X_i} \int_{\Theta \times \mathbf{\Sigma}} \left(\kappa\, z_i(s_i)(\overline{y}(\mathbf{s}) - \theta_0) + \frac{1}{2} z_i(s_i)^2 (\kappa + (n-1)) \right) dG(\theta_0, \mathbf{s}).$$

Suppose that $\kappa \neq 0$. Then the first-order necessary condition evaluated at $z_i(s_i) = 0$ yields

$$\int_{\Theta \times \mathbf{\Sigma}} (\overline{y}(\mathbf{s}) - \theta_0)\, dG(\theta_0, \mathbf{s}) = 0,$$

as required.

Suppose instead that $\kappa = 0$. Then the first-order conditions of the aggregate program no longer imply aggregate unbiasedness. Nonetheless, there exists at least one equilibrium with an unbiased aggregate. Consider the linear equilibrium in which player i chooses the strategy $\overline{x}_i(s_i) = a_i\mathbf{E}[\theta_0 \,|\, s_i] + (1-a_i)\mathbf{E}[\theta_0]$. By construction, the associated aggregate is unbiased. To see that such a linear equilibrium exists, note that the first-order condition for each player i requires that $x_i(s_i)$ is a linear function of $\mathbf{E}[\theta_0 \,|\, s_i]$ and $\mathbf{E}[\sum_{j\neq i}\overline{x}_j(s_j) \,|\, s_i]$. If the other players choose linear strategies, then the unique optimal choice for player i is also a linear function of $\mathbf{E}[\theta_0 \,|\, s_i]$ and $(n-1)$-terms, $\mathbf{E}[s_j \,|\, s_i]$. Given that these conditional expectations are linear in s_i, we obtain that $\overline{x}_i$ is linear in s_i. Given that α_i is bounded above so that $\psi_i > 0$, the system of first-order linear equations has full rank and can be inverted to obtain a unique profile of $(\overline{a}_1, \ldots, \overline{a}_n)$ which supports this unbiased linear equilibrium. *Q.E.D.*

Proof of Proposition 3: Given Corollary 4, we need only to show that

$$\mathcal{E}_Y(G) \subseteq \left\{ \overline{T} \,|\, \overline{T} \in \arg\max_{T\in\mathcal{T}} \Lambda(T, \overline{T}) \right\}.$$

Suppose not. Then there is a $\overline{T} \in \mathcal{E}_Y(G)$ and a $\hat{T} \in \mathcal{T}$ such that

$$\Lambda(\hat{T}, \overline{T}) > \Lambda(\overline{T}, \overline{T}).$$

Let $(\overline{T}_1, \ldots, \overline{T}_n)$ be the equilibrium offers which generate $\overline{T}$ and define $\hat{T}_i \equiv \hat{T} - \overline{T}_{-i}$. Because Λ is the sum of the players' payoff functions, there must exist some player i such that

$$\int_{\mathcal{Q}} (v_i(q) - \hat{T}_i(q)) dH(q \,|\, \hat{T}_i + \overline{T}_{-i}) >$$
$$\int_{\mathcal{Q}} (v_i(q) - \overline{T}_i(q)) dH(q \,|\, \overline{T}_i + \overline{T}_{-i})$$
$$= \max_{T_i\in\mathcal{T}} \int_{\mathcal{Q}} (v_i(q) - T_i(q)) dH(q \,|\, T_i + \overline{T}_{-i}).$$

The latter equality is implied by the hypothesis that $\overline{T}_i$ is an equilibrium strategy against $\overline{T}_{-i}$. Given that $\hat{T}_i$ yields greater payoff than $\overline{T}_i$, it follows that $\hat{T}_i \notin \mathcal{T}$.

The following lemma provides the key technical step in establishing a contradiction with the optimality of $\overline{T}_i$ over $\mathcal{T}$.

LEMMA 3 *Suppose Ω is a compact subset of $\mathbb{R}^k$, (Ω, d) is a metric space and (Ω, Σ, μ) is a probability space. For any bounded, lower-semicontinuous function, $f : \Omega \to \mathbb{R}$, there exist a sequence of upper-semicontinuous functions $(f_m)_{m\in\mathbb{N}}$ such that $f_m(\omega) \nearrow f(\omega)$ for each $\omega \in \Omega$, and there exists a corresponding sequence of open sets, $(\mathcal{O}_m)_{m\in\mathbb{N}}$, such that for each $m \in \mathbb{N}$, $f(\omega) = f_m(\omega)$ for $\omega \in \Omega \setminus \mathcal{O}_m$ and $\lim_{m\to\infty} \mu(\mathcal{O}_m) = 0$.*

Proof of Lemma: Denote the discontinuity points of f by the set Ω_0 and a typical element as ω_0. For each such point ω_0, we construct a ε-neighborhood entirely contained in Ω:

$$B_\varepsilon(\omega_0) = \{\omega \in \Omega \,|\, d(\omega_0, \omega) < \varepsilon\}.$$

$B_\varepsilon(\omega_0)$ is an open set in Ω. Define the open covering of Ω_0 as

$$B_\varepsilon = \bigcup_{\omega_0 \in \Omega_0} B_\varepsilon(\omega_0).$$

From B_ε we remove the discontinuity points of Ω_0 and denote the resulting set

$$\mathcal{O}_\varepsilon = B_\varepsilon \setminus \Omega_0.$$

Because Ω_0 is a closed set in Ω, the set $\mathcal{O}_\varepsilon$ is open in Ω and its complement $\Omega \setminus \mathcal{O}_\varepsilon$ is closed. By construction, f is continuous on $\Omega \setminus \mathcal{O}_\varepsilon$.

We next construct an ε-approximation to f as

$$f_\varepsilon(\omega) = \begin{cases} f(\omega), & \omega \in \Omega \setminus \mathcal{O}_\varepsilon, \\ \underline{f}, & \omega \in \mathcal{O}_\varepsilon, \end{cases}$$

where $\underline{f}$ is sufficiently small so that $\underline{f} \leq \inf_{\omega \in \mathcal{O}_\varepsilon} f(\omega)$. Because f is bounded, such an $\underline{f}$ exists. Thus, we have constructed a function that is continuous when restricted to the closed subset $\Omega \setminus \mathcal{O}_\varepsilon$ and also continuous when restricted to the open subset $\mathcal{O}_\varepsilon$.

Choose a sequence (ω_k) which traverses a boundary from $\Omega \setminus \mathcal{O}_\varepsilon$ to $\mathcal{O}_\varepsilon$. At the boundary, the function $f_\varepsilon(\omega_k)$ jumps downward. Because f_ε is a continuous function on the closed set $\Omega \setminus \mathcal{O}_\varepsilon$ and a continuous function on the open set $\mathcal{O}_\varepsilon$, it follows that f_ε is upper semi-continuous at such discontinuity points. Similarly, when (ω_k) traverses the boundary going from $\mathcal{O}_\varepsilon$ to $\Omega \setminus \mathcal{O}_\varepsilon$, the approximation $f_\varepsilon(\omega_k)$ jumps upward from a continuous function defined on an open set to a continuous function defined on a closed set. We conclude that the approximation f_ε is upper semi-continuous on Ω. Abusing slightly our notation, we set $\varepsilon = \frac{1}{m}$ and redefine the sequence of upper semicontinuous approximations in terms of $m \in \mathbb{N}$. This sequence, $(f_m)_{m \in \mathbb{N}}$, converges to f from below as $m \to \infty$.

To establish that $\lim_{m \to \infty} \mu(\mathcal{O}_m) = 0$ (here we again substitute ε with $\frac{1}{m}$ and retask the notation), it is sufficient to note that the sequence of open sets $(\mathcal{O}_m)_{m \in \mathbb{N}}$ in Σ decreases to $\emptyset$, and thus the probability measure $\mu(\mathcal{O}_m) \searrow 0$ as required. *Q.E.D.*

Returning to the main task, we wish to show that an upper-semicontinuous approximation to $\hat{T}_i = \hat{T} - \overline{T}_{-i}$ exists, $\hat{T}_i^m$, for which

$$\int_{\mathcal{Q}} (v_i(q) - \hat{T}_i^m(q)) dH(q \,|\, \hat{T}_i^m + \overline{T}_{-i}) > \int_{\mathcal{Q}} (v_i(q) - \overline{T}_i(q)) dH(q \,|\, \overline{T}_i + \overline{T}_{-i}).$$

The previous lemma establishes that a sequence of u.s.c. approximations exists for the lower-semicontinuous function, $-\overline{T}_{-i}$, which we denote $(-\overline{T}_{-i}^m)_{m \in \mathbb{N}}$. Thus, we can set $\hat{T}_i^m = \hat{T} + (-\overline{T}_{-i}^m)$ and construct $(\hat{T}_i^m)_{m \in \mathbb{N}}$ as a sequence of upper semi-continuous approximations to $\hat{T}_i$ such that $\hat{T}_i^m \nearrow \hat{T}_i$.

Using the approximation in the proof of the Lemma, the aggregate sequence $\hat{T}^m = \hat{T}_i^m + \overline{T}_{-i}$ is given by

$$\hat{T}^m(q) = \begin{cases} \hat{T}(q), & q \in \mathcal{Q} \setminus \mathcal{O}_m \\ \underline{t}, & q \in \mathcal{O}_m, \end{cases}$$

where $\underline{t}$ is a constant satisfying $\underline{t} \leq \inf_{\tilde{q} \in \mathcal{Q}} \hat{T}(\tilde{q})$, and $\mathcal{O}_m$ is an open set which can be made arbitrarily small.

Because $\hat{T}^m \leq \hat{T}$ with equality on $q \notin \mathcal{O}_m$, revealed preference implies

$$\overline{q}_0(\theta, \hat{T}) \notin \mathcal{O}_m \ \Rightarrow \ \overline{q}_0(\theta, \hat{T}^m) = \overline{q}_0(\theta, \hat{T}).$$

Given that $\mu(\mathcal{O}_m) \to 0$, we have pointwise convergence

$$\overline{q}_0(\theta, \hat{T}^m(q)) \to \overline{q}_0(\theta, \hat{T}(q)),$$

$$\text{Prob}\left(\theta \in \Theta \,|\, \overline{q}_0(\theta, \hat{T}) \in \mathcal{O}_m\right) \to 0,$$

and $H(\cdot \,|\, \hat{T}^m)$ converges in distribution to $H(\cdot \,|\, \hat{T})$.

Consider the limiting payoff to player i from playing $\hat{T}_i^m \in \mathcal{T}$:

$$\lim \inf_{m \to \infty} \int_{\mathcal{Q}} (v_i(q) - \hat{T}_i^m(q)) dH(q \,|\, \hat{T}_i^m + \overline{T}_{-i}).$$

Because $\hat{T}_i^m \leq \hat{T}_i$, a lower bound to the limiting payoff is

$$\lim \inf_{m \to \infty} \int_{\mathcal{Q}} (v_i(q) - \hat{T}_i(q)) dH(q \,|\, \hat{T}_i^m + \overline{T}_{-i}).$$

Because $H(\cdot \,|\, \hat{T}^m)$ converges in distribution to $H(\cdot \,|\, \hat{T})$, the limit of this lower bound exists and is

$$\int_{\mathcal{Q}} (v_i(q) - \hat{T}_i(q)) dH(q \,|\, \hat{T}_i + \overline{T}_{-i}),$$

which exceeds

$$\int_{\mathcal{Q}} (v_i(q) - \overline{T}_i(q)) dH(q \,|\, \overline{T}).$$

Thus, $\overline{T}_i$ cannot be a best response to $\overline{T}_{-i}$, contradicting the hypothesis that $\overline{T} \in \mathcal{E}_Y(G)$. *Q.E.D.*

Extension: Intrinsic common-agency games with general incomplete information:

We demonstrate how one can introduce more complex incomplete information structures and still obtain some of the the key results in Theorems 1 and 2. In particular, we generalize the application of intrinsic common agency to allow for private information by the agent *and* the principals.[35]

Suppose that each principal has a private type, $s_i \in S_i$, $i \in N$. This is in addition to the agent's type, denoted $\theta \in \Theta$. Each principal's benefit function is $v_i : \mathcal{Q} \times S_i \to \mathbb{R}$. For simplicity, we require that $\mathcal{Q}$, Θ and each S_i are finite and ordered. Let $(\theta, \mathbf{s}) \in \Theta \times \mathbf{S}$ be jointly distributed according to G.

Each principal i chooses a behavioral strategy that is a bounded transfer function, $T_i : \mathcal{Q} \times S_i \to \mathbb{R}$, which is optimal given the eventual realization of his private type, s_i. This strategy set is denoted by $\mathcal{T}_i$ for each player. Denote $T(q, \mathbf{s}) = \sum_{i \in N} T_i(q, s_i)$ and

[35] The full treatment of Bayesian games between informed principals is outside of the scope of this paper. On this, see Martimort and Moreira (2010).

$T \in \mathcal{T}$, where $\mathcal{T}$ is the set of aggregate transfers that are feasible for some strategy profile:

$$\mathcal{T} = \left\{ T : \mathcal{Q} \times \mathbf{S} \to \mathbb{R} \;\middle|\; \exists (T_i, \ldots, T_n) \in \mathcal{T}_1 \times \cdots \times \mathcal{T}_n,\; T = \sum_{i \in N} T_i \right\}.$$

Notice that this game is a weakly-symmetric distributional game with distribution function given by

$$H(q, \mathbf{s} | T) = \text{Prob}\left(\tilde{\theta} \in \Theta, \tilde{\mathbf{s}} \in \mathbf{S} \;\middle|\; q_0(\tilde{\theta}, T(\cdot, \tilde{\mathbf{s}})) \leq q, \tilde{\mathbf{s}} \leq \mathbf{s} \right).$$

If the players' strategy spaces were identical, this would be a strongly-symmetric distributional game.

Principal i chooses T_i to maximize (for any realization of type, $\mathbf{s}_i$)

$$\max_{T_i \in \mathcal{T}_i} \int_{\mathcal{Q} \times S} (v_i(q, s_i) - T_i(q, s_i)) dH(q, \mathbf{s} \,|\, T_i + \overline{T}_{-i}).$$

Theorem 1 immediately provides that

$$(\overline{q}_0, \overline{T}_1, \ldots, \overline{T}_n) \in \mathcal{E}(G) \iff \overline{T} \in \bigcap_{i \in N} \underset{T \in \mathcal{T}_i + \{\overline{T}_{-i}\}}{\arg\max}\; u_i(T - \overline{T}_{-i}, T).$$

The aggregate in this game is not pairwise bijective because of the measurability restrictions on the aggregate that each player can individually induce. That said, we can nonetheless establish the necessary condition in Theorem 2. To this end, define the aggregate welfare function (analogous to (5.6)) as

$$\Lambda(T, \overline{T}) = \int_{\mathcal{Q} \times \mathbf{S}} \left(\sum_{i \in N} v_i(q, \theta_i) - T(q, \mathbf{s}) + (n-1) \left(T(q, \mathbf{s}) - \overline{T}(q, \mathbf{s}) \right) \right) dH(q, \mathbf{s} \,|\, T).$$

Immediately, Theorem 2 implies that any equilibrium aggregate, $\overline{T} \in \mathcal{T}_1 + \cdots + \mathcal{T}_n$, must solve the following SGM program:

$$\overline{T} \in \underset{T \in \cap_{i \in N} \mathcal{T}_i + \{\overline{T}_{-i}\}}{\arg\max} \Lambda(T, \overline{T}).$$

This result is similar in spirit to the necessary condition for beauty contest games which delivered the "unbiased aggregate" result. Unfortunately, it is difficult to say more without placing additional structure on the aggregate game. In terms of sufficiency, we cannot appeal to Theorem 3 because the players' strategy spaces are not identical and hence the game is not strongly symmetric. Nevertheless, the fact that the more complex incomplete information games have a structure similar to the SGM program derived from (5.6) leaves us hopeful that these aggregate techniques will be of use in an even larger class of applications.

REFERENCES

Acemoglu, D. and M. Jensen, (2009): "Aggregate comparative statics," working paper.

Aliprantis, C. and K. Border, (2006): *Infinite Dimensional Analysis: A Hitchhiker's guide*, 3rd edition, Springer, Berlin.

Angeletos, G-M, and A. Pavan, (2007): "Efficient use of information and social value of information," *Econometrica*, 75, 1103-42.

Bergstrom, T. and H. Varian, (1985a): "Two remarks on Cournot equilibria," *Economic Letters*, 19, 5-8.

Bergstrom, T. and H. Varian, (1985b): "When are Nash equilibria independent of the distribution of agents' characteristics," *Review of Economic Studies*, 52, 715-718.

Bergstrom, T., Blume, L., and H. Varian, (1985): "On the private provision of public goods," *Journal of Public Economics*, 29, 25-49.

Bernheim, D. and M. Whinston, (1986a): "Common agency," *Econometrica*, 54, 923-942.

Bernheim, D. and M. Whinston, (1986b): "Menu auctions, resource allocations and economic influence," *Quarterly Journal of Economics*, 101, 1-31.

Corchon, L., (1994): "Comparative statics for aggregative games: The strong concavity case," *Mathematical Social Sciences*, 28, 151-165.

Cornes, R. and R. Hartley, (2000): "Joint production games and share functions," working paper.

Cornes, R. and R. Hartley, (2005): "Asymmetric contests with general technologies," *Economic Theory*, 26, 923-936.

Dixit, A., (1996): *The Making of Economic Policy: A Transaction-Cost Politics Perspective*, MIT Press.

Dubey, P., Haimanko, O. and A. Zapehelnyuk, (2006): "Strategic complements and substitutes, and potential games," *Games and Economic Behavior*, 54, 77-94.

Holmström, B. and P. Milgrom, (1988): "Common agency and exclusive dealing," working paper, April 1988.

Jensen, M., (2009): "Aggregative games and best-reply potentials," *Economic Theory*, forthcoming.

Klemperer, P and M. Meyer, (1989): "Supply function equilibria in oligopoly under uncertainty," *Econometrica*, 57, 1243-1277.

Kukushkin, N., (2004): "Best response dynamics in finite games with additive aggregation," *Games and Economic Behavior*, 48, 94-110.

Levin, D., (1984): "Taxation within Cournot oligopoly," *Journal of Public Economics*, 27: 281-290.

Martimort, D. and H. Moreira, (2010): "Common agency and public good provision under asymmetric information," forthcoming *Theoretical Economics.*

Martimort, D. and L. Stole, (2010a): "Public contracting in delegated agency games," working paper.

Martimort, D. and L. Stole, (2010b): "Public contracting in intrinsic common agency games," working paper.

Monderer, D. and L. Shapley, (1996): "Potential games," *Games and Economic*

Behavior, 14, 124-143.

Morris, S., and T. Ui, (2004): "Best response equivalence," *Games and Economic Behavior*, 49, 260-287.

Nikaido, H, and K. Isoda, (1955): "Note on noncooperative convex games," *Pacific Journal of Mathematics*, 5, 807-815.

Novshek, W., (1984): "Finding all n-firm Cournot equilibria," *International Economic Review*, 25, 61-70.

Myatt, D. and C. Wallace, (2008): "On the sources of value of information: Public announcements and macroeconomic performance," Nuffield College, Oxford University, Working Paper, 27 October 2008.

Novshek, W., (1985): "On the existence of Cournot equilibrium," *Review of Economic Studies*, 52, 85-98.

Okuguchi, K., (1993): "Unified approach to Cournot models: oligopoly, taxation and aggregate provision of a pure public good," *European Journal of Political Economy*, 9, 233-245.

Prendergast, C. and L. Stole, (1996): "Impetuous youngsters and jaded old timers: acquiring a reputation for learning," *Journal of Political Economy*, 104, 1105-1134.

Selten, R., (1970), *Preispolitik der Mehrproduktenunternehmung in der statischen Theorie*, Springer-Verlag.

Sion, M. (1958): "On general minimax theorems," *Pacific Journal of Mathematics*, 8, 171-176.

Szidarovszky, F. and S. Yakowitz, (1977): "A new proof of the existences and uniqueness of the Cournot equilibrium," *International Economic Review*, 18, 787-789.

Voorneveld, M., (2000): "Best-response potential games," *Economic Letters*, 66, 289-295.

Wilson, Robert, (1993): *Nonlinear Pricing*, Oxford University Press.