The Collective Wisdom of Beauty Contests

Martimort, David and Stole, Lars
Paris School of Economics, University of Chicago Booth School of Business

June 2011

Working paper

THE COLLECTIVE WISDOM OF BEAUTY CONTESTS[1]

June 2011

DAVID MARTIMORT AND LARS STOLE

This note uses techniques developed for aggregate games to characterize the set of equilibria for a beauty contest or prediction game in which the experts' preferences are quadratic, but with an otherwise unrestricted information structure for private signals and the state variable. We show that, on aggregate, the experts' collective estimate of the unknown parameter to be estimated is unbiased for *every* equilibrium.

KEYWORDS: Aggregate games, beauty contests, prediction games.

1. A BEAUTY-CONTEST 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 context of a beauty contest) who observes a private signal, $s_i \in \Sigma_i$, which is possibly informative about θ_0 and the other players' signals. Player i's strategy is a mapping from i's private signals to a public action or "prediction" of θ_0, which we denote as the announcement or prediction, $x_i(s_i)$. Player i cares both about being correct (i.e., x_i close to the true θ_0) and about his distance to the average assessment, $y = \frac{1}{n}\sum_j x_j$. Importantly, we do not assume that player i prefers to be close to the average; indeed, in many strategic settings it seems likely that a player would benefit the most by making an accurate prediction that is distinctive from the "herd." For example, one can think of this game as being played by debt rating agencies, each of which desires to correctly estimate a firm's underlying probability of default. Holding the accuracy $||x_i - \theta_0||$ fixed, it seems plausible that agency i prefers that its report is further away from the average assessment. In short, a player in a prediction game may want to be both accurate and distinctive. Of course, one can easily imagine situations where making an inaccurate prediction has greater costs to a player when the average assessment of the other players is closer to the truth. This may be the case when the experts are conservative enough and fear a reputation loss from moving away form that average. Here, a player's payoffs may be for accuracy and herding.

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 which maps 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. The aggregate prediction space is the set of additively-separable mappings,

[1]Parts of this note appeared in an earlier version of "Aggregate Representations of Aggregate Games" (2010) which has been revised without the beauty contests application and is no longer circulating in its previous form. The authors thank John Birge, Emir Kamenica, David Myatt, Michael Schwarz and seminar participants at the University of Chicago for helpful conversations and suggestions.

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

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

$y : \Sigma \mapsto \Theta$, that can be generated from some strategy profile of the players:

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

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 they put in their objective function on being correct may vary across players and the joint distribution of signals is arbitrary. Formally, player i's payoff is

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

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.[1] Unlike these cited papers, we emphasize that 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.

2. AGGREGATE CONCURRENCE IN AGGREGATE GAMES

The beauty contest game thus described is an example of an *aggregate game* as defined in Martimort and Stole (2011a).[2] An aggregate game is a normal form game with the additional restriction that each player i's payoff can be represented by a function of his own strategy, $x_i \in X_i$, and a proper (not one-to-one) aggregation of every player's strategy (including i), $y = \phi(\mathbf{x})$. In the case of beauty contests, the individual strategies and the aggregate are functions mapping from signal spaces to reports and the aggregate is simply $y(\mathbf{s}) = \phi(\mathbf{s}) = \frac{1}{n} \sum_i x_i(s_i)$. The immediate import of Martimort and Stole (2011a) is that every equilibrium aggregate must satisfy an aggregate concurrence principle: each player must find the equilibrium aggregate optimal over the space of aggregate variations available to that player. In particular, if $\overline{\mathbf{x}} = (\overline{\mathbf{x}}_1, ...\overline{\mathbf{x}}_n)$ is an equilibrium of the game with equilibrium aggregate $\overline{y} = \frac{1}{n} \sum_i \overline{\mathbf{x}}_i$, then a consequence of Theorem 1 in Martimort and Stole (2011a), is that for any nonnegative weighting vector $(\lambda_1, \ldots, \lambda_n)$,

$$\overline{y} \in \underset{y \in \bigcap i \in N \phi(X_i, \overline{\mathbf{x}}_{-i})}{\arg\max} \sum_{i \in N} \lambda_i u_i(\phi_i^{-1}(y, \overline{\mathbf{x}}_{-i}), y),$$

where

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

[1] 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.

[2] See Martimort and Stole (2011b) for applications of aggregate techniques to the class of common agency games with adverse selection and public contracts.

3. A NECESSARY CONDITION

Replacing x_i in player i's payoff with $n(y - \overline{y}) + \overline{x}_i$, player i's payoff can be stated as

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

where $\psi_i \equiv (n-\alpha_i)^2 + \alpha_i(1-\alpha_i) > 0$ and R_i is a remainder function that is independent of y and x_i. Using the weight of $\lambda_i = \frac{1}{n(n-\alpha_i)}$ for each player i (this weight is positive given the assumed upper bound on α_i), and ignoring the residual terms R_i, it must be that if $\overline{\mathbf{x}}$ is a Nash equilibrium with aggregate $\overline{y}$, then

$$\overline{y} \in \underset{y\in Y(\overline{y})}{\arg\max} - \int_{\Theta\times\mathbf{\Sigma}} \left\{\frac{1}{2}y(\mathbf{s})^2\left(\frac{1}{n}\sum_{i\in N}\frac{\psi_i}{n-\alpha_i}\right) - y(\mathbf{s})\overline{y}(\mathbf{s})(n-1) - y(\mathbf{s})n\theta\left(\frac{1}{n}\sum_{i\in N}\frac{1-\alpha_i}{n-\alpha_i}\right)\right\} dG(\theta_0,\mathbf{s}),$$

where $Y(\overline{y}) \equiv \bigcap_{i\in N}\{y \mid \exists z_i \in X_i \text{ s.t. } y = \overline{y} + z_i\}$ is the set of feasible variations in the aggregates given $\overline{y}$. 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 weighted objective function above to

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

The integrand in the definition of Λ is continuously differentiable and strictly concave in $y(\mathbf{s})$ pointwise, and so the necessary condition requires that $\overline{y}(\mathbf{s})$ maximizes this objective pointwise over the set of feasible aggregates for each $\mathbf{s} \in \mathbf{\Sigma}$. Because Θ is convex, the local first-order condition is 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:*

$$\int_{\Theta\times\mathbf{\Sigma}}(\overline{y}(\mathbf{s}) - \theta_0)dG(\theta_0,\mathbf{s}) = 0. \tag{3.1}$$

For any beauty-contest game with $\kappa = 0$, there is an equilibrium in which the aggregate is unbiased if $\mathbf{E}[\mathbf{E}[\theta_0|s_j]|s_i] = \mathbf{E}[\theta_0|s_i]$ *for all* $i \in N$.

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 measure-zero 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 $\mathbf{E}[\overline{y}(\mathbf{s})] = \hat{\theta}$.[3] 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 the players' signals are independently distributed conditional on θ_0, 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 with probability one. The unbiased-aggregate result, however, implies that on average such individual biases cancel each other out.[4]

Finally, we emphasize again that aggregate unbiasedness also holds even if $\alpha_i < 0$

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

[4]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

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

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

for some players. With such preferences, players prefer to make accurate predictions (i.e., $1-\alpha > 0$), but also prefer to distinguish themselves from the mob.[5] Remarkably, equilibrium unbiasedness is also necessary for games in which players prefer to be dissimilar.

APPENDIX

Proof of Proposition 1: If $\overline{y}$ is an equilibrium aggregate, it must maximize $\Lambda(y,\overline{y})$ over the set of feasible aggregates $y \in Y(\overline{y})$. 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

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

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 $\kappa = 0$. We will prove that $\overline{x}_i(s_i) = \mathbf{E}[\theta_0|s_i]$ for $i \in N$ is an unbiased equilibrium under the assumption that $\mathbf{E}[\mathbf{E}[\theta_0|s_j]|s_i] = \mathbf{E}[\theta_0|s_i]$ for all $i \in N$. It is immediate that such a profile of strategies will generate an unbiased aggregate. To see that it is an equilibrium, we need only consider an agent's individual first-order condition with respect to x_i; there is no additional value to using an aggregate approach in this case. Individual i's first-order condition requires that

$$\alpha_i\mathbf{E}[x_i(s_i)-\theta_0|s_i] + (1-\alpha_i)\left(\frac{n-1}{n}\right)\mathbf{E}[x_i(s_i)-y(\mathbf{s})|s_i] = 0.$$

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

If $\overline{x}_j(s_j) = \mathbf{E}[\theta_0|s_j]$, then $\mathbf{E}[\overline{x}_j(s_j)|s_i] = \mathbf{E}[\theta_0|s_i]$, from which we have

$$\mathbf{E}[y(\mathbf{s})|s_i] = \frac{1}{n}x_i(s_i) + \frac{1}{n}\sum_{j\neq i}\mathbf{E}[\overline{x}_j(s_j)|s_i] = \frac{1}{n}x_i(s_i) + \frac{n-1}{n}\mathbf{E}[\theta_0|s_i].$$

Substituting into player i's first-order condition yields

$$\alpha_i\mathbf{E}[x_i(s_i) - \theta_0|s_i] + (1-\alpha_i)\left(\frac{n-1}{n}\right)\mathbf{E}[x_i(s_i) - y(\mathbf{s})|s_i]$$
$$= \alpha_i\left(x_i(s_i) - \mathbf{E}[\theta_0|s_i]\right) + (1-\alpha_i)\left(\frac{n-1}{n}\right)^2\left(x_i(s_i) - \mathbf{E}[\theta_0|s_i]\right)$$

which is zero as required for $x_i(s_i) = \mathbf{E}[\theta_0|s_i]$. Given each player's minimization program is strictly convex, this is sufficient for optimality and proves that the hypothesized strategy profile is an equilibrium. *Q.E.D.*

REFERENCES

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

Martimort, D. and L. Stole, (2011a): "Aggregate representations of aggregate games with applications to common agency," Applied Theory Initiative Working Paper #2011-05, June 2011.

Martimort, D. and L. Stole, (2011b): "Public contracting in delegated common-agency games ," Applied Theory Initiative Working Paper #2011-06, June 2011.

Morris, S., and H. Shin, (2002): "The social value of public information," *American Economic Review*, 92, 1521-1534.

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.

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