Information Acquisition, Efficiency, and Non-Fundamental Volatility

Benjamin Hébert Jennifer La'O

July 11, 2022

This Paper

- we consider a class of beauty-contest games with rationally inattentive agents
- we introduce: partial monotonicity & partial invariance
  - we define these as properties of divergences
  - move around conditional distributions of some variables, keeping the marginal of others intact
  - generalizes standard invariance concept to particular dimensions
- we ask: when is there always an equilibrium with zero non-fundamental volatility?
  - monotonicity in noisy public signals: sufficient and necessary
- we ask: when is there always a constrained-efficient equilibrium?
  - invariance in endogenous actions: sufficient and necessary

Agents and Actions

- continuum of agents $i \in [0,1]$, each agent takes an action

$$a^i \in A \subset \mathbb{R}$$

  - A is non-empty, convex, and compact.
  - aggregate action $\bar{a} = \int_0^1 a^i di \in \bar{A} \subseteq \mathbb{R}$

- "mean-critical" class of payoffs, finite set of payoff-relevant states $s \in S \subset \mathbb{R}$:

$$u(a,\bar{a},s) = g(a,s) + G(\bar{a},s) + (a-\bar{a})\frac{\partial}{\partial \bar{a}}G(\bar{a},s),$$

  - payoffs defined over own action, aggregate action, and payoff-relevant state
  - the functions g and G are $\mathcal{C}^1$ and $\mathcal{C}^2$ in a and $\bar{a}$, respectively, for each $s \in S$
  - G convex (concave): strategic complementarity (substitutability)
  - this class is special: it has efficiency in the use of information (we generalize Angeletos Pavan 2007)

- leading example: linear-quadratic (e.g. as in Morris Shin 2002),

$$u(a^i,\bar{a},s) = -(1-\beta)(a^i - s)^2 - \beta(a^i - \bar{a})^2$$

Exogenous States and Priors

- finite set of "noisy public signals" $r \in R \subset \mathbb{R}$
  - assume there are shocks $e \in \mathcal{E} \subset \mathbb{R}$ that are orthogonal to the fundamentals $s \in S$
  - public signals generated by some mapping $\mathcal{E} \times S \to R$
- common prior over exogenous state space

$$\mu_0(s,r), \qquad \mu_0 \in \mathcal{U}_0 \equiv \Delta(S \times R)$$

- endogenous aggregate action function $\bar{\alpha} \in \bar{\mathcal{A}}$

$$\bar{a} = \bar{\alpha}(s,r), \qquad \bar{\alpha} : S \times R \to \bar{A},$$

- $\mu_0 \in \mathcal{U}_0$ and $\bar{\alpha} \in \bar{\mathcal{A}}$ induce a prior over the larger space

$$\mu \in \mathcal{U} \equiv \Delta(S \times R \times \bar{A}),$$

$$d\mu(s,r,\bar{a}) = d\mu_0(s,r)d\delta_{\bar{\alpha}(s,r)}(\bar{a})$$

Strategies

- strategies: joint measures over actions and posteriors

$$\pi \in \Delta(A \times \mathcal{U})$$

  - let $\mu' \in \mathrm{supp}(\pi) \subseteq \mathcal{U}$ denote a posterior measure in the support of π

- a strategy $\pi \in \Delta(A \times \mathcal{U})$ is "Bayes-consistent" with the prior $\mu \in \mathcal{U}$ if

$$E^{\pi(a,\mu')}[\mu'] = \mu$$

  and if all $\mu' \in \mathrm{supp}(\pi) \subseteq \mathcal{U}$ are absolutely continuous with respect to μ

- let $\Pi(\mu) \subset \Delta(A \times \mathcal{U})$ be the set of all strategies Bayes-consistent with the prior

  - set of all feasible strategies

Information Costs

- The cost function is posterior-separable (Caplin, Dean, Leahy 2022)

$$C(\pi,\mu) = E^{\pi(a,\mu')}[D(\mu'||\mu)]$$

  - $D : \mathcal{U} \times \mathcal{U} \to \mathbb{R}_+$ is a divergence, convex in its first argument

- posterior-separable class: cost is expected value of the divergence from prior μ to posterior $\mu' \ll \mu$
- why? allows us to define properties of cost functions as properties of divergences

Posterior-Separable class: what does it nest?

- mutual information (Sims 2003): associated divergence is Kullback-Leibler (KL)
- Tsallis entropy costs (Caplin, Dean, Leahy 2020)
- neighborhood-based cost function (Hébert Woodford, 2021)
- LLR cost function (Pomatto, Strack, Tamuz 2018)

Signals vs. Posteriors

- many signal structures can generate the same $\pi \in \Pi(\mu)$
  - could condition on (pay attention to) exogenous (s, r) or endogenous $\bar{a}$
- D describes cost of optimal (least costly) way to reach the given posterior
- Could write $D : \mathcal{U}_0 \times \mathcal{U}_0 \times \bar{\mathcal{A}} \to \mathbb{R}$ instead of $D : \mathcal{U} \times \mathcal{U} \to \mathbb{R}$
  - posteriors always degenerate in $\bar{a}$ conditional on (s, r)
- Why D might depend on $\bar{\mathcal{A}}$: if least costly way to obtain info. involves conditioning on $\bar{a}$, changes in $\bar{\alpha} \in \bar{\mathcal{A}}$ affect information costs
  - post. sep. costs could be defined on $S \times R$ (excluding this possibility) or $S \times R \times \bar{A}$ (allowing it)

The Agent's Problem

- agent's expected utility from taking action a under posterior $\mu' \in \mathcal{U}$

$$V(a,\mu') \equiv E^{\mu'(s,r,\bar{a})}[u(a,\bar{a},s)]$$

- agent's problem

$$\max_{\pi\in\Pi(\mu)} E^{\pi(a,\mu')}[V(a,\mu') - D(\mu'||\mu)]$$

Assumption

$D(\phi\{\mu_0',\bar{\alpha}\}||\phi\{\mu_0,\bar{\alpha}\})$ *is continuously differentiable on* $(\mu_0',\bar{\alpha}) \in int(\mathcal{U}_0 \times \bar{\mathcal{A}})$.

Strategy Profiles and Mean-Consistency

- let $\xi \equiv (\pi, \bar{\alpha})$ denote a symmetric strategy profile

Definition

A symmetric strategy profile $\xi = (\pi, \bar{\alpha})$ is mean-consistent if, for all $(s,r) \in S \times R$,

$$\bar{\alpha}(s,r) = E^{\pi(a,\mu')}[a|s,r]$$

- Independence of realizations $\rightarrow$ apply the law of large numbers (Uhlig 1996)
- mean action in the population $=$ expected action

Equilibrium Definition

Definition

Given $\mu_0 \in \mathcal{U}_0$, a symmetric Bayesian Nash equilibrium (BNE) of the game is a mean-consistent strategy profile $\xi = (\pi, \bar{\alpha})$ such that $\pi \in \Pi(\mu)$ is a best response to the prior $\mu = \phi\{\mu_0, \bar{\alpha}\}$,

$$\max_{\pi \in \Pi(\mu)} E^{\pi(a,\mu')}[V(a,\mu') - D(\mu'||\mu)],$$

Constrained Efficiency Definition

Definition

Given a common prior $\mu_0 \in \mathcal{U}_0$, a strategy profile ξ^* is **constrained efficient** if it solves

$$\max_{\bar{\alpha}\in\bar{\mathcal{A}},\pi\in\Pi(\mu)} E^{\pi(a,\mu')}[V(a,\mu') - D(\mu'||\mu)],$$

subject to $\mu = \phi\{\mu_0', \bar{\alpha}\}$ and mean-consistency,

$$\bar{\alpha}(s,r) = E^{\pi(a,\mu')}[a|s,r], \qquad \forall(s,r) \in S \times R$$

- planner internalizes impact of actions on $\bar{\alpha}$

Three departures from the standard RI framework

- posterior-separable class of cost functions
  - allows us to define properties of cost functions as properties of divergences
- agents can pay attention to noisy public signals r
  - builds a bridge to the exogenous information beauty contest literature (Morris Shin 2002, Angeletos Pavan 2007, Bergemann Morris 2013)
  - potential source of non-fundamental volatility in equilibrium
- agents can pay attention to the endogenous aggregate action
  - reasonable in many applied economic settings: agents learn from prices, macro statistics, etc.

Properties of Divergences

- want to show: properties of equilibria related to properties of information cost functions
- partial monotonicity & partial invariance: properties of divergences
  - divergence: function of the prior and posterior
  - describe how a divergence responds to certain transformations of its inputs (prior and posterior)
- we need first to describe the transformations we have in mind

Information about R

- consider a probability measure

$$\mu \in \mathcal{U} = \Delta(S \times R \times \bar{A})$$

- $\gamma_R : \mathcal{U} \to \Delta(S \times \bar{A})$ be the function that generates the marginal distribution on $(S \times \bar{A})$ from $\mu \in \mathcal{U}$

$$\gamma_R\{\mu_1\}(s,\bar{a}) = \sum_{r \in R} \mu(s,r,\bar{a}), \qquad \forall s,\bar{a}$$

  - removes information about $r \in R$

Compositional operator

- consider a pair of measures

$$\mu_1, \mu_2 \in \mathcal{U}, \qquad \text{with} \qquad \gamma_R\{\mu_1\} \ll \gamma_R\{\mu_2\}$$

- we define a compositional operator $\eta_R : \mathcal{U} \times \mathcal{U} \to \mathcal{U}$ by

$$\frac{d\eta_R\{\mu_1, \mu_2\}}{d\mu_2} = \frac{d\gamma_R\{\mu_1\}}{d\gamma_R\{\mu_2\}}$$

  - replaces the conditional distribution of r given $(s, \bar{a})$ from μ_1 with that of μ_2
  - but preserves marginal over $s, \bar{a}$ of μ_1
  - note that by construction: $\eta_R\{\mu_1, \mu_1\} = \mu_1$

- essentially replaces information about r, while keeping marginals over $s, \bar{a}$ intact

Example: Compositional operator

$\bar{a} = \bar{a}_0$	r_1	r_2
s_1	0.45	0.15
s_2	0.30	0.10

Table: $\mu_1 \in \mathcal{U}$

$\bar{a} = \bar{a}_0$	r_1	r_2
s_1	0.10	0.10
s_2	0.20	0.60

Table: $\mu_2 \in \mathcal{U}$

	$\bar{a} = \bar{a}_0$
s_1	3
s_2	1/2

Table: $\frac{d\gamma_R\{\mu_1\}}{d\gamma_R\{\mu_2\}}$

$\bar{a} = \bar{a}_0$	r_1	r_2
s_1	0.30	0.30
s_2	0.10	0.30

Table: $\eta_R\{\mu_1, \mu_2\} \in \mathcal{U}$

Example: Compositional operator

$\bar{a} = \bar{a}_0$	r_1	r_2
s_1	0.45	0.15
s_2	0.30	0.10

Table: $\mu_1 \in \mathcal{U}$

$\bar{a} = \bar{a}_0$	r_1	r_2
s_1	0.10	0.10
s_2	0.20	0.60

Table: $\mu_2 \in \mathcal{U}$

	$\bar{a} = \bar{a}_0$
s_1	3
s_2	1/2

Table: $\frac{d\gamma_R\{\mu_1\}}{d\gamma_R\{\mu_2\}}$

$\bar{a} = \bar{a}_0$	r_1	r_2
s_1	0.30	0.30
s_2	0.10	0.30

Table: $\eta_R\{\mu_1, \mu_2\} \in \mathcal{U}$

Monotonicity in R

Definition

A divergence D is monotone in R given $\mu \in \mathcal{U}$ if, for all $\mu' \in \mathcal{U}$ with $\mu' \ll \mu$,

$$D(\mu'||\mu) \geq D(\eta_R\{\mu', \mu\}||\mu)$$

A divergence is monotone in R if it is monotone in R given any $\mu \in \mathcal{U}$.

- replaces conditionals $(r|s, \bar{a})$ of the posterior with those of the prior
- if this operation reduces their divergence, then monotone in R

Invariance in R

Definition

A divergence D is invariant in R given $\mu \in \mathcal{U}$ if, for all $\mu', \mu'' \in \mathcal{U}$ with $\mu' \ll \mu$ and $\gamma_R\{\mu\} \ll \gamma_R\{\mu''\}$,

$$D(\eta_R\{\mu', \mu\}||\mu) = D(\eta_R\{\mu', \mu''\}||\eta_R\{\mu, \mu''\}).$$

A divergence is invariant in R if it is invariant in R given any $\mu \in \mathcal{U}$.

- replaces conditionals $(r|s, \bar{a})$ of μ and μ' with those of another dist
- invariance in R $\Leftrightarrow$ if prior and posterior share a common conditional dist $(r|s, \bar{a})$, then the exact values of this conditional distribution are immaterial for their divergence.

Example: Kullback-Leibler (Mutual Information)

- KL divergence from prior μ to posterior μ'

$$D_{KL}(\mu'||\mu) = \sum_{s\in S}\sum_{r\in R}\mu'_0(s,r)\ln\left(\frac{\mu'_0(s,r)}{\mu_0(s,r)}\right)$$

- KL divergence from prior μ to $\eta_R\{\mu',\mu\}$

$$D_{KL}(\eta_R\{\mu',\mu\}||\mu) = \sum_{s\in S}\sum_{r\in R}\mu'_0(s,r)\ln\left[\frac{\sum_{r\in R}\mu'_0(s,r)}{\sum_{r\in R}\mu_0(s,r)}\right]$$

- $D_{KL}(\mu'||\mu) \geq D_{KL}(\eta_R\{\mu',\mu\}||\mu)$ equivalent to

$$\sum_{r\in R}\mu'_0(s,r)\ln\left[\frac{\mu'_0(s,r)}{\sum_{r\in R}\mu'_0(s,r)}\right] \geq \sum_{r\in R}\mu'_0(s,r)\ln\left[\frac{\mu_0(s,r)}{\sum_{r\in R}\mu_0(s,r)}\right], \qquad \forall s$$

  - holds by Jensen's inequality

- therefore, the KL divergence is monotone in R

Example: Kullback-Leibler (Mutual Information)

- KL divergence from $\eta_R\{\mu,\mu''\}$ to $\eta_R\{\mu',\mu''\}$

$$D_{KL}(\eta_R\{\mu',\mu''\}||\eta_R\{\mu,\mu''\}) = \sum_{s\in S}\sum_{r\in R}\mu_0'(s,r)\ln\left[\frac{\sum_{r\in R}\mu_0'(s,r)}{\sum_{r\in R}\mu_0(s,r)}\right]$$

- μ'' drops out
- therefore, KL divergence is invariant in R

Example: Neighborhood-Based Cost function (Hebert Woodford, 2021)

- let $S = \{s_1, \ldots, s_{|S|}\}$ and $R = \{r_1, \ldots, r_{|R|}\}$
  - with s_i strictly increasing in i and r_j strictly increasing in j.
- treat each adjacent pair (s_{i-1}, s_i) and (r_{j-1}, r_j) as a "neighborhood"
- cost penalizes differences between μ, μ' within each neighborhood using the KL divergence

$$D_{SR}(\mu'||\mu) = \sum_{i=2}^{|S|} \sum_{j=1}^{|R|} \frac{\mu_0'(s_i, r_j) + \mu_0'(s_{i-1}, r_j)}{d_{ij}^s} D_{KL,i,r_j}(\mu_0'||\mu_0) + \sum_{i=1}^{|S|} \sum_{j=2}^{|R|} \frac{\mu_0'(s_i, r_j) + \mu_0'(s_i, r_{j-1})}{d_{ij}^r} D_{KL,j,s_i}(\mu_0'||\mu_0)$$

  - with "perceptual distances"

$$d_{ij}^s = s_i - s_{i-1}, \qquad \text{and} \qquad d_{ij}^r = r_j - r_{j-1}$$

  - if difference $r_j - r_{j-1}$ is large, then it is less costly to distinguish between states
- the divergence D_{SR} is not monotone in R

Information about $\bar{A}$

- $\gamma_{\bar{A}} : \mathcal{U} \to \Delta(S \times R)$ be the function that generates the marginal distribution on $(S \times R)$ from $\mu \in \mathcal{U}$
- we define a compositional operator $\eta_{\bar{A}} : \mathcal{U} \times \mathcal{U} \to \mathcal{U}$ by

$$\frac{d\eta_{\bar{A}}\{\mu_1, \mu_2\}}{d\mu_2} = \frac{d\gamma_{\bar{A}}\{\mu_1\}}{d\gamma_{\bar{A}}\{\mu_2\}}, \quad \text{with} \quad \gamma_{\bar{A}}\{\mu_1\} \ll \gamma_{\bar{A}}\{\mu_2\}$$

  - replaces the conditional distribution of $\bar{a}$ given (s, r) from μ_1 with that of μ_2
  - preserves marginal over s, r of its first argument μ_1

Invariance in $\bar{A}$

Definition

A divergence D is invariant in $\bar{A}$ given $\mu \in \mathcal{U}$, if for all $\mu', \mu'' \in \mathcal{U}$ with $\mu' \ll \mu$ and $\gamma_{\bar{A}}\{\mu\} \ll \gamma_{\bar{A}}\{\mu''\}$,

$$D(\eta_{\bar{A}}\{\mu', \mu\} || \mu) = D(\eta_{\bar{A}}\{\mu', \mu''\} || \eta_{\bar{A}}\{\mu, \mu''\}).$$

A divergence is invariant in $\bar{A}$ if it is invariant in $\bar{A}$ given any $\mu \in \mathcal{U}$

- replaces conditionals $(\bar{a}|s, r)$ of μ and μ' with those of another dist
- invariance in $\bar{A}$ $\Leftrightarrow$ if prior and posterior share a common conditional dist $(\bar{a}|s, r)$, then the exact values of this conditional distribution are immaterial for their divergence.

Example: Neighborhood-Based Cost function

- abstract from the r dimension
- consider a different neighborhood based cost function

$$D_{S\bar{A}}(\mu'||\mu) = \sum_{i=2}^{|S|} \frac{\mu_0'(s_i) + \mu_0'(s_{i-1})}{d_i^s(\bar{\alpha})} D_{KL,i}(\mu_0'||\mu_0)$$

  - with "perceptual distances"

$$d_i^s(\bar{\alpha}) = (s_i - s_{i-1}) \left[1 + \left(\frac{\bar{\alpha}(s_i) - \bar{\alpha}(s_{i-1})}{s_i - s_{i-1}} \right)^2 \right].$$

- if difference $\bar{\alpha}(s_i) - \bar{\alpha}(s_{i-1})$ is large, then it is less costly to distinguish between states
  - agent pays attention to the aggregate action and extreme actions are easier to observe
- the divergence $D_{S\bar{A}}$ is not invariant in $\bar{A}$.

Where we're headed

- show that these properties are directly related to properties of equilibria
  - monotone in R: related to zero non-fundamental volatility in equilibrium
  - invariant in $\bar{A}$: related to constrained efficiency

Implication of Monotonicity in R

Lemma

The posterior-separable cost function $C(\pi,\mu)$ is associated with a divergence D that is monotone in R given $\mu \in \mathcal{U}$ if and only if, for all strategies $\pi \in \Pi(\mu)$, the strategies $\pi' \in \Pi(\mu)$ induced from π by $(a,\mu') \mapsto (a,\eta_R\{\mu',\mu\})$ satisfy

$$C(\pi',\mu) \leq C(\pi,\mu).$$

- strategies that ignore r are the least costly iff the cost function is monotone in R.

Implication of Invariance in $\bar{A}$

Lemma

The posterior-separable cost function $C(\pi,\mu)$ *is associated with a divergence* D *that is* $\bar{A}$*-invariant given* $\mu \in \mathcal{U}$ *if and only if, for all* $\mu'' \in \mathcal{U}$ *with* $\gamma_{\bar{A}}\{\mu\} \ll \gamma_{\bar{A}}\{\mu''\}$ *and all strategies* $\pi \in \Pi(\mu)$*, the strategies* $\pi'' \in \Pi(\eta_{\bar{A}}\{\mu,\mu''\})$ *induced from* π *by* $(a,\mu') \mapsto (a,\eta_{\bar{A}}\{\mu',\mu''\})$ *satisfy*

$$C(\pi'',\eta_{\bar{A}}\{\mu,\mu''\}) = C(\pi,\mu).$$

- there is no channel by which $\bar{\alpha}$ affects the cost of information acquisition iff the cost function is invariant in $\bar{A}$.

Equilibrium Existence

Proposition

A symmetric BNE exists.

Proof: Application of Kakutani's fixed point theorem, relying on finiteness of $S \times R$, continuity of the utility function, and convexity and continuity of the information cost function.

Uniqueness of equilibrium $\bar{\alpha}$

Proposition

If $\beta < 0$ and the divergence D is invariant in $\bar{A}$, then there is a unique aggregate action function $\bar{\alpha}$ common to all symmetric BNE.

- both conditions essentially rule out strategic complementarity in payoffs and information costs
- when D is not invariant in $\bar{A}$, strategic complementarity in information acquisition can arise

Fundamentals-driven equilibria

Definition

An aggregate action function is s-measurable if

$$\bar{\alpha}(s,r) = \bar{\alpha}(s,r') \quad \forall \text{ and } r,r' \in R$$

A BNE $(\pi, \bar{\alpha})$ is s-measurable if $\bar{\alpha}$ is s-measurable and π does not condition on r (conditional on s).

- s-measurable equilibrium $=$ zero non-fundamental volatility

Existence of s-measurable equilibria

Proposition

An s-measurable symmetric BNE exists for all mean-critical utility functions if and only if, for all s-measurable $\bar{\alpha} \in \bar{\mathcal{A}}$, the divergence D is monotone in R given $\mu = \phi\{\mu_0, \bar{\alpha}\}$.

- if costs are monotone in R, then an equilibrium exists with zero non-fundamental volatility
- if not, there exists a utility function such that all equilibria have non-fundamental volatility
- If in addition D is invariant in $\bar{A}$, and either G is strictly concave or if D is "convex enough" relative to G, all equilibria are s-measurable

Sketch of proof/intuition

- consider an s-measurable equilibrium. $\bar{a} = \bar{\alpha}(s)$
- are best responses measurable in r? depends on monotonicity in R
  - payoffs: agents do not care about r per se, only insofar as it affects $\bar{a}$
  - monotone in R: conditional on $s, \bar{a}$, the least-costly strategy ignores r
    - same fixed-point argument for existence proves that an s-measurable eq. exists
  - non-monotone: strategies that pay some attention to r are less costly
    - cannot have equilibrium given this $\bar{\alpha}$
    - construct utility function such that this $\bar{\alpha}$ is only possible s-measurable equilibrium

Payoff Externalities

- recall the agent's problem

$$\max_{\pi \in \Pi(\mu)} E^{\pi(a,\mu')}[V(a,\mu') - D(\mu'||\mu)],$$

  - there are two places in which externalities can show up: $V(a,\mu')$ and $D(\mu'||\mu)$

- our "mean-critical" utility functions rule out payoff externalities
  - the mean value $E[a|s]$ is a critical point of $h(\bar{a}) = E[u(a,\bar{a},s)|s]$
  - imposing efficiency in the *use* of information (Angeletos Pavan 2007)
  - appendix: all such functions have our functional form (subject to some regularity conditions)
  - appendix: interpretation as production economy with quasi-linear preferences (Angeletos Sastry 2020)

Constrained Efficiency

Proposition

If the divergence D is invariant in $\bar{A}$, then a constrained efficient symmetric BNE exists.

If in addition $\beta \leq 0$, all symmetric BNE are constrained efficient.

- no externalities in payoffs and no externalities in information costs
- necessity: C must be locally-invariant in $\bar{A}$ on the relevant domain (planner's optimum)

Intuition

- if costs are non-invariant in $\bar{A}$:
  - $\bar{\alpha}, \bar{\alpha}' \in \bar{\mathcal{A}}$ induce different information costs
  - then an informational externality arises
  - agents do not take into account how their actions affect other agents' info. costs
  - similar in spirit to learning externality that arises with signals about endog. objects
    Laffont (1985), Amador Weill (2010), Vives (2017)
- if costs are invariant in $\bar{A}$:
  - $\bar{\alpha}, \bar{\alpha}' \in \bar{\mathcal{A}}$ induce the same information costs
  - no such informational externality arises

Conclusion

- equil. with zero non-fundamental volatility iff monotonicity in R ("public signals")
- constrained efficient equil. iff invariance in $\bar{A}$ (endogenous actions)
- linear-quadratic-Gaussian examples in the paper illustrate these points
- equil. properties driven by separate properties of information costs
  - mutual info $\rightarrow$ zero non-fundamental vol. and efficiency
  - Fisher info (Hébert Woodford, 2021) $\rightarrow$ non-fundamental vol and inefficiency
  - Tsallis entropy (Caplin, Dean, Leahy, 2019) $\rightarrow$ non-fundamental vol, but efficient