Modeling Ignorance without Bayesian beliefs

Olivier Compte and Andrew Postlewaite

July 11, 2018

General objective: an attempt to challenge the way we deal with ignorance in models

This talk:

- Discuss some issues with standard Bayesian models
and in particular, the risk of "overspecification"
- Review a number of examples in which we propose an alternative path

Some notation

a action $a \in A$

s underlying state, $s \in S$

z signal/perception $z \in Z$

ω distribution over (z, s)

$\sigma : Z \longrightarrow A$ strategy

Standard Bayesian model:

$\max_\sigma E_\omega u(\sigma(z), s)$

Agent behaves as if at each z he held "bayesian belief"

$\beta_z = \omega(. \mid z) \in \Delta(S)$...

and maximized subjective expected utility given that belief

$\max_a E_{\beta_z} u(a, s)$

Two kinds of misspecification in economic models

1. Underspecification

Lucas's critique. Effect of policy changes should not be based on naive model of economic behavior in which agents would not anticipate the consequences of policy changes.
Meaningful revenue comparisons in auctions. Agent should be assumed to adjust bidding to auction format.

Response to the risk of underspecification

- agents should not miss important aspects of the environment (unless we can document why)
- agents know the model...and have common priors.

Myerson (2004): "If we can assume any arbitrary characteristics for the individuals in our model, then why could we not explain behavior even more simply by assuming that each individual has a payoff function that is maximized by this behavior? Thus, to avoid trivialization, applied theorists have generally limited themselves to models that satisfy Harsanyi's consistency assumption.

Comment:

Consistency between agent's beliefs is not the only concern.

Consistency with the "actual" environment is another one.

- **In applied work on auctions**, $\omega(v_1, ..., v_n)$ is interpreted as empirical distribution, i.e. a fictitious but somewhat/hopefully accurate description of the "typical" environment that bidders face.
- **Betting at the horse track**: we may explain that one bets all his fortune on a horse by assuming crazy beliefs about the odds of winning... but we are generally interested in explaining behavior in which beliefs are consistent with "actual" odds (whatever that means...).

2. Full knowledge of model and the risk of overspecification

Pb with full knowledge: it may be too easy for agents to exploit the structure of the model.
(in particular because models are simplified representation of the environment)

(a) Some modeling "details" may acquire undue importance.

- "simple" auction model with only two values, 10 or 100?

  $->$ **Pb:** behavior driven by particular choice of 10 and 100

- **Fix?** values drawn in a continuum $[10, 100]$

But...not always a good fix because we increase the number of instruments: $b(v)$ for each v.

Getting $v = 99$ becomes evidence of top valuation bidder $->$ may matter a great deal in all pay auction....
All pay generates efficient allocation... but this relies on implausible ability to exploit the value distribution (and rank in that value distribution)

In the same vein:

- Information aggregation (knowledge of what it implies to be pivotal) : $z_i = \eta_i v$ estimate of common value v. Large population

If told that $x\%$ of the agents are more optimistic than him, agent can perfectly infer v.

- mechanism design, contract theory (allows fine tuning of transfers as if there were symmetric information about signal distribution)

- repeated games (allows fine tuning of continuation values to monitoring technology).

2. Full knowledge and the risk of overspecification

(b) Prior bias

It may become too easy for agents to "undo" errors that we build into the signal technology.
Behavior ends up being (overly?) driven by priors (a somewhat fictitious theoretical construct).

Illustration I (with no underlying variability)

- Analyst runs 74 draws of a die

$x_k \in \{1, ..., 6\}$, $X = \sum_{k=1}^{74} x_k$, $s = \Pr(X > 140)$

- Agent wins 50α if $X > 140$, he loses 10α otherwise.
- Assume that agent can choose the stake α.

That is, assume he solves

$$\max_{\alpha} v(\alpha, s) \equiv s\ u(50\alpha) + (1 - s)\ u(-10\alpha)$$

Define $\boldsymbol{\alpha}^*(s)$ that solves

$$\frac{5u'(50\alpha)}{u'(-10\alpha)} = \frac{1-s}{s}$$

Underlying characteristic of the decision problem is s.
Agent chooses $\alpha^* = \boldsymbol{\alpha}^*(s)$.

What if you wish to model noisy perceptions of s?

For example, z such that $\frac{z}{1-z} = \eta \frac{s}{1-s}$, where η is lognormal.
This defines distribution $\omega(z, s)$, but the marginal on s is degenerate:

$$\max_{\sigma(.)} Ev(\sigma(z), s) = E(su(50\sigma(z)) + (1-s)u(-10\sigma(z)))$$

yields

$$\sigma^*(z) = \alpha^* \text{ for all } z.$$

Agent behaves as if he knew s, because $\beta_z = \omega(. \mid z)$ is degenerate.

Illustration II (with some underlying randomness)

s constant return on investiment, y size of investment

$\pi(y) = sy - y^2/2$

optimal investment $\mathbf{y}^*(s) = s$

state and observations: $s \sim \mathcal{U}[0,1]$, $z = \eta s$, $\omega(z,s)$

optimal investment: $y = \phi(z) = E_\omega[s \mid z]$

Signals affect investment decisions

yet in expectation: $Ey = E_z[E[s \mid z]] = Es$

- **Consequence I.** Quality of signals does not affect expected investment

Illustration II (with some underlying randomness)

s constant return on investiment, y size of investment

$\pi(y) = sy - y^2/2$

optimal investment $\mathbf{y}^*(s) = s$

state and observations: $s \sim \mathcal{U}[0,1]$, $z = \eta s$, $\omega(z,s)$

optimal investment: $y = \phi(z) = E_\omega[s \mid z]$

Signals affect investment decisions

yet in expectation: $Ey = E_z[E[s \mid z]] = Es$

- **Consequence I.** Quality of signals does not affect expected investment
  When signals' quality is poor, agent ignores them.
  Quality of "Prior Information" ω about (z,s) is perfect, and $E_\omega s$ is solely driving investment.
- **Consequence II**. For another player who would try to predict y, behavior of poorly informed firm is **more predictable**: behavior of lesser informed is biased toward the prior.

Comment: implication for comparative statics w.r.t. information

Agent has two kinds of information: z and ω.
Noisier $z \Rightarrow \omega$ acquires importance, decision biased towards prior.

Implication: lack of information$\longrightarrow$more predictible behavior
This is a modeling artifact... that may give rise to questionable comparative statics

Examples

- **Information rents**: easier to extract rents from poorly informed players, better to sign a contract before agent acquires information (Cremer Khalil).
  In principle, poor information could instead make it more difficult to target offers.

- **Auctions.** Symmetric second price auction, two bidders, competition fearcer between uninformed (Ganuza)

$v_i = \theta_i$, independent draws.

Uninformed bidders bid $E\theta_i$, $R = E\theta$ (symmetry) Informed bidders bid θ_i, $R = E \min \theta_i$
What creates rents is dispersion of bids... and *lack of information* could create dispersion.
Alternative model: $z_i = \eta_i v_i$, $b_i = \lambda_i z_i$. Equilibrium λ^*?
Noisier estimates potentially create more rents (through increased dispersion, and increased caution – smaller λ^*)

- **Coordination** is easier between uninformed divisions (Dessein)

Alternative path I:

Mitigate ability to adjust to priors (or model specifications) by adding strategy restrictions

- **Investment problem:**

$\sigma^\lambda : z \to \mathbf{y}^*(\lambda z)$
$v(\lambda) = E[s\sigma^\lambda(z) - (\sigma^\lambda(z))^2] = E(s\lambda z - z^2\lambda^2/2)$

$$\lambda^* = \frac{Esz}{Ez^2} = \frac{E\eta}{E\eta^2}$$

hence $E\lambda^* z = \frac{(E\eta)^2}{E\eta^2} Es < Es$
Caution in treating estimate, caution increases when estimate is noisier
No need to define Bayesian beliefs: estimate z acts as a "belief" about s, and the agent learns to take belief z with caution.

- **Stakes example** α?

$\alpha^*(z)$ =naive decision rule
$\sigma^\lambda : z \longrightarrow \boldsymbol{\alpha}^*(\lambda z)$, where $\lambda \neq 1$ means agent uses z with caution
Optimal caution?

$$\max_\lambda E_\omega u(\alpha^*(\lambda z), s)$$

- Decision is partially driven by correct s, partially driven by perception z, with caution λ adjusted to preferences and quality of estimate.

- In both examples, we assume that agent can deal with noisy observations...*without simultaneously assuming perfect ability to exploit the process that generates them.*

- No need to introduce arbitrary variability in underlying state to model ignorance about underlying state.

- **All pay auction** with v_i i.i.d draws

standard model: $b(v)$ such that
$u(v) = q(v)v - b(v) = q(v)E[v^{(2)} \mid v^{(2)} < v]$
with restriction $b(v) = \lambda v$:
$u_i(\lambda_i, \lambda) = \Pr(\lambda_i v_i > \lambda \max v_j) E[v_i \mid \lambda_i v_i > \lambda \max v_j] - \lambda_i E v$

- **All pay auction** with v_i i.i.d draws

standard model: $b(v)$ such that
$u(v) = q(v)v - b(v) = q(v)E[v^{(2)} \mid v^{(2)} < v]$
with restriction $b(v) = \lambda v$:

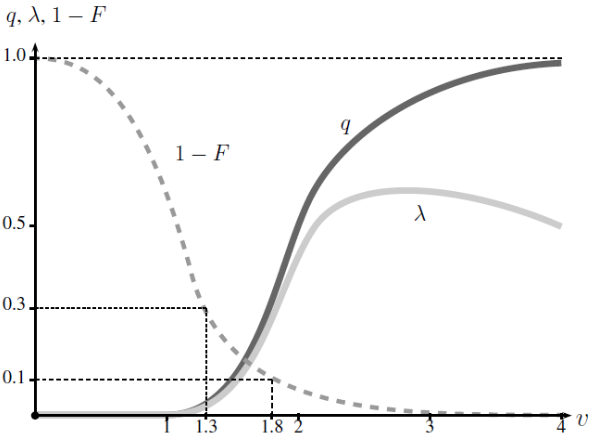
$u_i(\lambda_i, \lambda) = \Pr(\lambda_i v_i > \lambda \max v_j)E[v_i \mid \lambda_i v_i > \lambda \max v_j] - \lambda_i Ev$
no symmetric equilibrium in pure strategies when $n > 2$ and
$\log v_i \sim \mathcal{N}(0, 0.5)$

In standard model: only "high" valuation bid seriously (because $q(v)$ remains small for many values of v)

Restricted model incorporates some information about rank, but not enough...

Intuition: no negative profit: $\frac{1}{n}E \max v_i - \lambda Ev \geq 0$
This pushes λ down quickly with n.
Bidding higher becomes profitable.

Note: Similar insight could obtain in model with aggregate uncertainty (i.e. $v_i = w_i\theta$, with θ random), but at the cost of greater mathematical complexity.


$q, \lambda, 1-F$
1.0
0.5
0.3
0.1
$1-F$
q
λ
1
1.3
1.8
2
3
4
v

Alternative path II.

Mitigate ability to adjust to priors (or model specifications) by defining family of noisy strategies

- **Investment problem:**

Agent targets y, but implements $\eta y \Rightarrow$
$\max_y E_\eta u(\eta y, s) = \max_y syE\eta - y^2 E\eta^2/2$

- **Choosing stake α.**

Agent targets α, but implements $\eta\alpha \Rightarrow \max_\alpha E_\eta u(\eta\alpha, s)$

- **Auctions.** Agent target λv, but implements ϕv where $\phi = \frac{\lambda\eta}{\lambda+(1-\lambda)\eta}$

Rather than introducing randomness in the underlying state, introduce randomness in strategic choice.
This permits to examine the first order consequence of mistakes, without endogenizing the structure of mistakes.

Restrictions as a way to mitigate the risk of overspecification.

- **Econometrics**: The statistician looks for a model that predicts the (stochastic) relationship between exogenous variables z and the quantity s that he cares about. A statistical model λ provides a prediction

$$\widehat{s} = \sigma^{\lambda}(z)$$

that can be evaluated according to a loss function $L(s, \widehat{s})$.

Based on data $S = \{(z_t, s_t)_t\}$, each model λ can be evaluated

$$L(\lambda) = \sum_t L(s_t, \sigma^{\lambda}(z_t))$$

and $L(\lambda)$ minimized.
But there is a risk of finding a model λ^* that works well on sample S, and not on another sample S'.

To mitigate that risk, econometricians typically limit the set of models considered (or put costs on complex ones – Lasso), restrict number of instruments z, or construct auxiliary ones $\widehat{z} = \gamma(z)$

Restrictions as a way to mitigate the risk of overspecification.

In theory papers, agents face a similar problem: they choose action $\widehat{s}$ as close as possible to the ideal one s, they experience a loss $L(s, \widehat{s})$ otherwise:

$$L(s, \widehat{s}) = u(s, s) - u(\widehat{s}, s)$$

They get signals/perceptions z that may guide their choices, and evaluate each strategy $\sigma^\lambda(z)$:

$$L(\lambda) = EL(s, \sigma^\lambda(z)) = \int L(s, \sigma^\lambda(z))\omega(z, s)dsdz$$

Same risk of overfitting σ^λ to ω, unless one puts constraints on Σ (or add a cost structure $c(\lambda)$ as in rational inattention models)

Comparison with standard Bayesian model

Comparison with standard Bayesian model

Standard bayesian model: agents behave as if they knew ω (or agent knows ω)... and could draw all consequences of that knowledge.

With direct strategy restrictions: agents behave as if they knew ω... *but were only able to use that knowledge to compare "few" strategies.*

Primitives: ω and restriction $\Sigma = \{\sigma^\lambda\}_\lambda$
(ω, Σ) is a **joint restriction** on agent's ability to target optimal action for each s.

Comparison with standard Bayesian model

Standard bayesian model: agents behave as if they knew ω (or agent knows ω)... and could draw all consequences of that knowledge.

With direct strategy restrictions: agents behave as if they knew ω... *but were only able to use that knowledge to compare "few" strategies.*

Primitives: ω and restriction $\Sigma = \{\sigma^\lambda\}_\lambda$
(ω, Σ) is a **joint restriction** on agent's ability to target optimal action for each s.
Restriction on information? on sophistication? Whichever you like.
Information is not solely about what you observe, but also about how to take advantage of what you observe.

Interpretation of ω

Standard Bayesian model: Mostly, ω is viewed as reflecting *the agent's knowledge*.
Agent holds belief β_z at z, or behaves as if ...
+ *consistency condition* that ties beliefs $\beta_z = \omega(. \mid z)$ to one another (harsanyi)
+ *consistency condition* that ties beliefs to real data (in empirical work or when discussing revenues in auctions, ω implicitly reflects typical problems faced).

With restrictions: our view is that $\omega(z, s)$ reflects *an outsider's perspective* (not the agent's knowledge), capturing a (somewhat ficticious) distribution over pbs faced by agent.

Interpretation of ω (continued)

Classic Bayesian interpretation: agent "knows" ω and draws inferences from **signal** z

Our interpretation: only the analyst knows ω, agent **cannot** make inferences from **perception** z; he is just assumed to know/learn which strategy $\sigma(z)$ is best, within Σ.

Interpretation of ω (continued)

Classic Bayesian interpretation: agent "knows" ω and draws inferences from **signal** z

Our interpretation: only the analyst knows ω, agent **cannot** make inferences from **perception** z; he is just assumed to know/learn which strategy $\sigma(z)$ is best, within Σ.

comment: Does this mean agents never draw inferences? NO
One could model inferences from signals and the process by which agents make inferences, say β... but then we would define $\omega(\beta, s)$ i.e. the joint process over states and inferences...
... and Σ the set of strategies available, for example, a set of $\lambda-$robust/prudent decision rules
$\sigma^{\lambda}(\beta) = \max_a \min_{\beta'} E_{\beta'} u(a, s) + d(\beta, \beta')/\lambda$
... and optimize over prudence λ.
$\max_{\lambda} E_{\omega} u(\sigma^{\lambda}(\beta), s)$

Other popular means of mitigating knowledge of underlying distributions

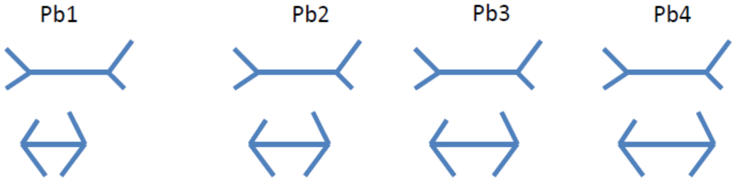
- **Rational inattention**

This corresponds to endogenizing $\omega(z, s)$ under informational flow constraint.

Pb: (i) Not tractable even in simple investment problem $u(y, s) = sy - y^2/2$

(ii) Suffers from prior bias: $Ey = Es$ independently of constraint.

(iii) Sometimes/often? the signals that we get may not be the result of an optimal coding problem



In each pb, you have the option to bet on which line is larger.
You gain 10 if you succeed, you lose 100 otherwise.
If aware that an optical illusion might be present, the asymetry in gains and loss make you cautious, betting on fewer problems when the penalty is increased.
But the asymetry is unlikely to change the way your eyes process information (unless possibly this becomes a recurrent problem).
The coding of information by the eyes likely applies to a context broader than that defined by the experiment, and it seems fine to consider it exogenous and next assume that agents process their perceptions and adjust decision making to take into account the possibility that perceptions might be misleading.

- **Robust decision making**
- **Robust DM**: at each z, specify belief set B_z, then

$$\max_y \min_{\beta \in B_z} E_\beta u(y, s)$$

What structure on distribution over (B_z, z, s) ?

Pbs:

If this is left unstructured, we get an underspecified model, with too many degrees of freedom

$+$ Maxmin is a mechanical rule that could bear little relationship with actual welfare.

Pause

Strategy restriction as robustness check.

- Cf: all pay auction example
- **It also matters in first price...**

Existence of pure strategy equilibrium is no longer guaranted**...**
$u_i(\lambda_i, \lambda) = (1 - \lambda_i) D_i(\lambda_i, \lambda)$
with $D(\lambda_i, \lambda) = \Pr(\lambda_i v_i > \lambda \max v_j) E[v_i \mid \lambda_i v_i > \lambda \max v_j]$
With v_i are very concentrated with proba p, then λ^* close to 1 (Bertrand competition, $\frac{\partial D_i(\lambda, \lambda)}{\partial \lambda_i}$ high)
But if v_i's are rather dispersed with proba $1 - p$, equilibrium profits cannot be very small. **Contradiction.**

Strategy restriction as robustness check.

- Cf: all pay auction example
- **It also matters in first price...**

Existence of pure strategy equilibrium is no longer guaranted**...**
$u_i(\lambda_i, \lambda) = (1 - \lambda_i) D_i(\lambda_i, \lambda)$
with $D(\lambda_i, \lambda) = \Pr(\lambda_i v_i > \lambda \max v_j) E[v_i \mid \lambda_i v_i > \lambda \max v_j]$
With v_i are very concentrated with proba p, then λ^* close to 1 (Bertrand competition, $\frac{\partial D_i(\lambda,\lambda)}{\partial \lambda_i}$ high)
But if v_i's are rather dispersed with proba $1 - p$, equilibrium profits cannot be very small. **Contradiction.**

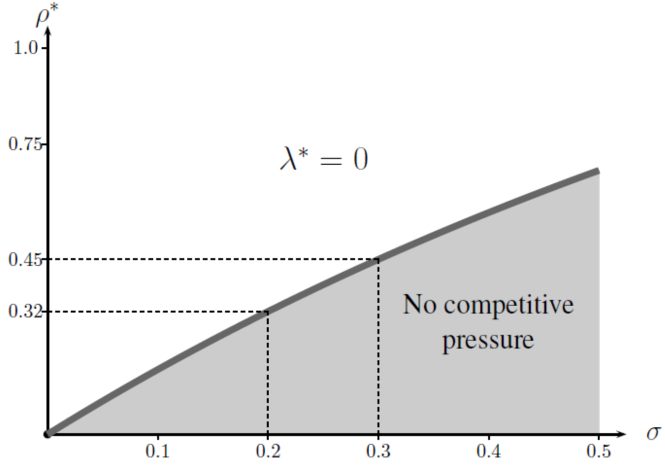
$\rightarrow$ **Dispersion uncertainty** may lead to inexistence (for the same reason that for some demand functions, price competition for differentiated products may not have pure strategy equilibria...
$\rightarrow$ Similar results can be obtained in standard models with correlated values, but interpretation of distributional assumptions more difficult in these models

Centipede games

No deadline and δ close to 1.
Exit date t_i for player i.
If $t_1 < t_2$, then $v_1 = t_1$ and $v_2 = t_1 - \rho t_1$ with $\rho \in (0, 1)$
–> unraveling to the start of the game (even if ρ is small)
Strategy restriction: Each player i is restricted to stochastic strategies $t_i = \eta_i \lambda_i$ with $\log \eta_i \sim \mathcal{N}(0, \sigma^2)$

$$
\begin{aligned}
v_1(\lambda_1, \lambda_2) &= E v_1(\eta_1 \lambda_1, \eta_2 \lambda_2) \\
&= E \min(\eta_1 \lambda_1, \eta_2 \lambda_2) - \rho \Pr(\eta_1 \lambda_1 > \eta_2 \lambda_2) E[\eta_2 \lambda_2 \mid \eta_1 \lambda_1 > \eta \\
v(x\lambda, \lambda) &= \lambda [H(x) - \rho G(x)]
\end{aligned}
$$

so for ρ small $H'(1) > \rho G'(1)$ and there are no competitive pressures, for ρ large, it unravels....

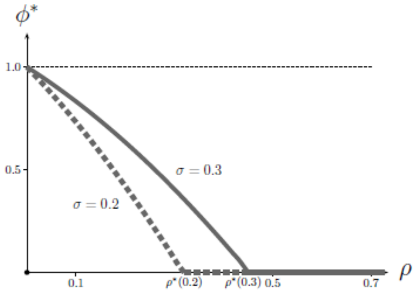


Comment: structure of noise matters: as t gets large, stakes increase, but noise term increases too. Story would be different with a deadline: as we get close to it, noise must vanish.

With a deadline: target date $\phi_i T$, but ϕ_i implemented with noise:

$$t_i = \widetilde{\phi}_i T \text{ with } \widetilde{\phi}_i = \frac{\phi_i \eta_i}{1 - \phi_i + \phi_i \eta_i}$$

with $\log \eta_i \sim N(0, \sigma^2)$. Randomness smaller near 0 or near terminal date.
$v_1(\phi_1, \phi_2) = E v_1(\widetilde{\phi}_1 T, \widetilde{\phi}_2 T)$
Equilibrium ϕ^*? It depends on σ and ρ: little unravelling (i.e. ϕ^* large) when noise is large or ρ small.

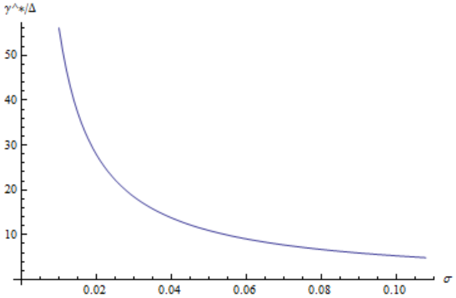


When ρ is not too large, or σ not too small, ϕ^* is neither 0 nor 1.

Similar insights with repeated games with a deadline T.
Continuous time $t \in [0, T]$, playoff flows

$$\begin{array}{ccc} & C & D \\ C & 1,1 & -L, 1+g \\ D & 1+g, -L & 0,0 \end{array}$$

Assume it takes d to detect a defection.
Similar structure as centipede: when $t_2 > t_1 + d$, $v_1 = t_1 + gd$ and $v_2 = v_1 - (g+L)d$
Main difference: the stake $\Delta = (g+L)d$ is constant (rather than increasing with t_1) and this pushes termination to the end. How far from the end? *What's the length of the "end game"?*
With large T, $t_i = T - \eta_i \gamma_i$, $v_1(\gamma_1, \gamma_2) = Ev_1(T - \eta_1\gamma_1, T - \eta_2\gamma_2)$ with $\log \eta_i \sim \mathcal{N}(0, \sigma^2)$
Equilibrium γ^*? When γ gets large, $\sigma\gamma$ gets large compared to Δ, and delaying defection pays



That noise plays a role in undermining unravelling is well-known. The point is that we often use complex information structures to generate noise, spending much time investigating how agents optimally react to each and every signal given the structure of noise, without necessarily understanding well the resulting noise structure.
Here, analysis is focused on "length of end game", and on how it depends on exogenous structure of noise and payoff parameters.

Coordination games

Classic global game set up:

	H	L
H	$r - c_1, r - c_2$	$-c_1, 0$
L	$0, -c_2$	$0, 0$

c_1 and c_2 fixed, $r > 0$ drawn from a (rather) flat distribution.
- **threshold strategy:** play H if $r > c_i(1 + x_i)$. Any $x > 0$ is a symmetric equilibrium
- **noisy thresold strategy**: $x_i = \lambda_i \eta_i$ where η_i is a positive random variable (or $x_i = \lambda_i + \varepsilon_i$)
- $V_i(\lambda_i, \lambda_i) = Ev_i(\lambda_i \eta_i, \lambda_i \eta_i)$

Unique interior equilibrium $(\lambda_1^*, \lambda_2^*)$, essentially because noise creates risk that each one wants to mitigate, and because
- risk is small when λ_2 is high $\Longrightarrow$ 1 wants to set λ_1 lower than λ_2
- risk is large when λ_2 is small $\Longrightarrow$ 1 wants to set λ_1 above λ_2

Differences with global game approach:
- Threshold strategy assumed, no attempt to endogenize it
- uniqueness (of interior eq.) obtains without use of dominance relations.
- ($\lambda^* = \infty$) is an eq., and a weak one by standard arguments
- **Key force:** chance of miscoordination is positive, and players have incentives to reduce it.
- miscoordination could stem from other sources... such as limits in handling complex games

If c_1 and c_2 are random what rule $\sigma(c_1, c_2)$ should you use? risk dominance? Seems implausible for asymmetric games+not well defined for larger games

Assume players play L is $r < \max(c_1, c_2)$, and evaluate strategies:

$u_i^\lambda(a_i) = (1-\lambda)\min_{a_j} u_i(a_i, a_j) + \lambda \max_{a_j} w(a_i, a_j)$

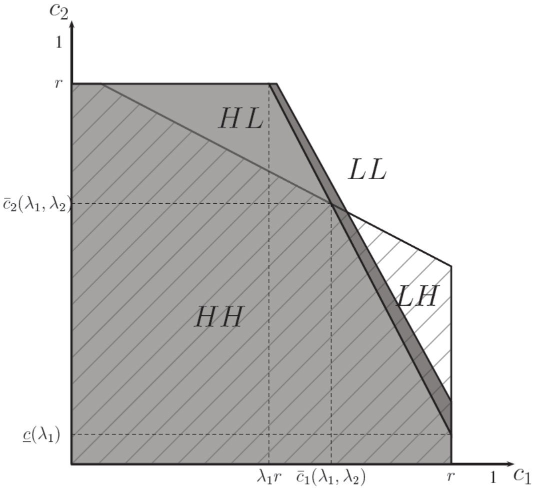
$u_i^\lambda(H) = (1-\lambda)(-c_i) + \lambda(2r - c_1 - c_2)$ and $u_i^\lambda(L) = 0$

$\sigma_i^\lambda = \arg\max_{a_i} u_i^\lambda(a_i) =$ play H if $c_i + \lambda c_j < 2r\lambda$ and $\max(c_1, c_2) < r$.

$V_i(\lambda_1, \lambda_2) = E_{c_1,c_2} u_i(\sigma_i^{\lambda_i}, \sigma_j^{\lambda_j})$

Eq? unique λ_1^*, λ_2^*.

Uniqueness stems from miscoordination being always present, and each player optimally mitigating the risk and costs of miscoordination.


c_2
1
r
HL
LL
$\bar{c}_2(\lambda_1, \lambda_2)$
LH
HH
$\underline{c}(\lambda_1)$
$\lambda_1 r$
$\bar{c}_1(\lambda_1, \lambda_2)$
r
1
c_1

Handling more complex environment (I)
Information transmission, Crawford Sobel
player 1 (Decision maker) $u(d, s) = -(d - s)^2$
player 2 (adviser) $u(\theta, s) = -(d - \theta)^2$
bias: $b \equiv \theta - s$.
Observations: $x = s + \varepsilon \qquad b_i = \eta_i b$.
player 1 observes x, b_1, **player** 2 observes θ, b_2.
s, b, ε and each η_i are independent random variables. $E\eta_i = 1$ and $\varphi_i \equiv E\eta_i^2$.
$1/\varphi_i \equiv$ *ability* parameter.

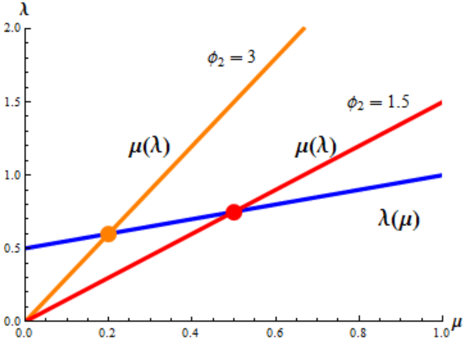
Model 1. manipulation/debiasing

$$y = \theta + \mu b_2.$$

$$d = y - \lambda b_1.$$

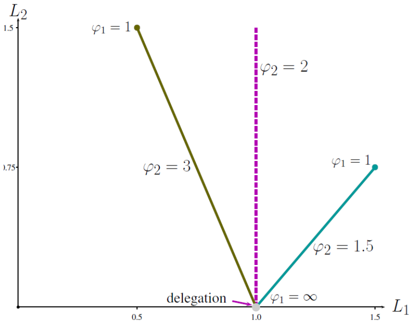
μ : degree to which 2 manipulates
λ : degree to which 1 attempts to de-bias the suggestion y.
$d = \theta + \mu b_2 - \lambda b_1 = s + b(1 + \mu\eta_2 - \lambda\eta_1)$.
$L_1(\lambda,\mu) = Eb^2(1 + \mu\eta_2 - \lambda\eta_1)^2$ and $L_2(\lambda,\mu) = Eb^2(\mu\eta_2 - \lambda\eta_1)^2$.
Equilibrium (λ^*,μ^*)**?** (recall $\varphi_i = E\eta_i^2 > 1$):
Best responses: player 1: $\lambda(\mu) = (1+\mu)/\varphi_1$; player 2: $\mu(\lambda) = \lambda/\varphi_2$
Higher noise (i.e. higher φ) diminishes incentives to manipulate or debias.
Equilibrium losses: $L_1 = 1 + \frac{2-\varphi_2}{\varphi_1\varphi_2 - 1}$ and $L_2 = \frac{\varphi_2}{\varphi_1\varphi_2 - 1}$
$L_1 + L_2 = 1 + \frac{2}{\varphi_1\varphi_2 - 1}$
Welfare decreases with ability of either player.
2 benefits both from own noise ($\varphi_2 \nearrow$) **and** 1's noise ($\varphi_1 \nearrow$)
1 benefits from own noise iff $\varphi_2 < 2$, otherwise he prefers smaller φ_1

Setting $\varphi_1 = 2$, and varying φ_2 we get:



Large φ_2 implies a smaller μ : less manipulation by 2, lesser debiasing by player 1..

Equilibrium losses L_1 and L_2 when φ_1 changes (for $\varphi_2 = 1.5, 2, 3$)



Two regimes:

$\varphi_2 < 2$. Player 2 has a high ability to influence the decision, and better ability by player 1 hurts both players. The decision maker (player 1) would actually *benefit from delegating* the decision to player 2.

$\varphi_2 > 2$. Player 2 has little ability to influence the decision, he makes less distorted suggestions, and the decision maker's ability to assess the bias accurately benefits him (and hurts the adviser). *Communication is better than delegation.*

Model 2. Skepticism/Trust under competing opinions.

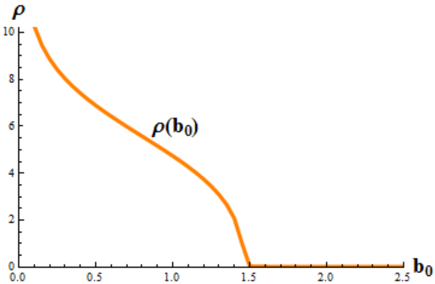
For any $\rho \geq 0$, define σ_1^ρ as follows:

$$\begin{aligned} d &= y \text{ if } \mid y - x \mid < \rho \\ &= x \text{ otherwise} \end{aligned}$$

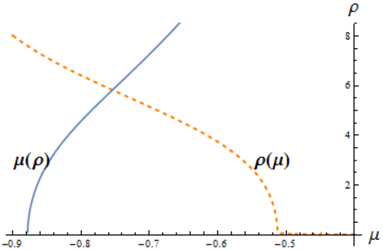
$\rho =$ degree to which the decision maker trusts the suggestion y.

$$\Sigma_1 = \{\sigma_1^\rho\}_{\rho \geq 0}; \Sigma_2 \text{ as before.}$$

Optimal response ρ as b_0 varies (setting $\sigma_z = 1$) with $b = b_0 z$, $\log z \sim \mathcal{N}(0, \sigma_z^2)$



Equilibrium? Player 2 has to take into account the risk that his advice will be ignored. This risk is limited when ρ is large (figure drawn for $\eta_2 = 1$)



Handling more complex environment (II)

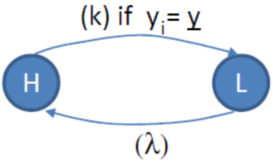
Repeated game with autocorrelated shocks on payoff structure.

$$\begin{array}{ccc} & C & D \\ C & x-\gamma, x-\gamma & -\gamma, x \\ D & x, -\gamma & 0,0 \end{array} \quad \text{with } x \in \{0, g\},\ g > \gamma$$

Autocorrelation: In each period, with small probability α, new draw of x. $Q = \Pr(x = 0)$.

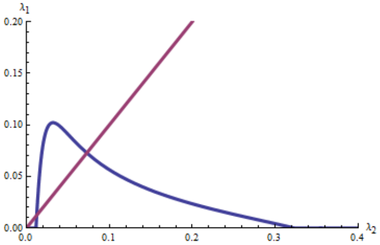
Observations: two private signals $\overline{y}_i, \underline{y}_i$ $p = \Pr(\overline{y}_i \mid (a_i, C), g) > q = \Pr(\overline{y}_i \mid (a_i, a_j), x)$ for $(a_j, x) \neq (C, g)$

Strategies $\sigma^{k,\lambda}$: two state automaton, play C in H and D in L



values $v_1(\lambda_1, \lambda_2) = q_A^{\lambda_1,\lambda_2} g - \gamma q_B^{\lambda_1,\lambda_2}$ where $A = \{a_2 = C, x = g\}$ and $B = \{a_1 = C\}$

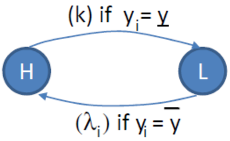
Equilibrium? $p = 0.95, q = 0.1, k = 0.2, g = 1, \gamma = 0.3, \; Q = 0.5$



Insights:

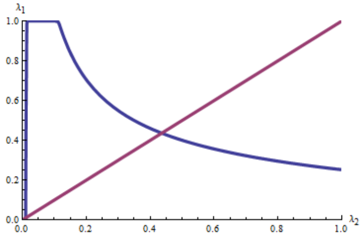
- private monitoring (*and incentive to trigger D*) is not an issue.
- Outcome away from contrained first best $\max_\lambda v_1(\lambda, \lambda)$. Game becomes similar to standard contribution game
- Autocorrelation provides stronger incentives to adopt a dynamic strategy (when the other does so as well) – if Q large enough, *playing C always is a bad strategy.*
- If $Q = 0$, cooperation cannot be sustained (with this strategy set)

Alternative strategy set?



recoordination is made easier, not relying on two independent attempts

Equilibrium:



In expectation, **recoordination** attempts are less frequent, but more successful when attempted −> higher expected welfare obtains

Shaping intuitions differently

Baron Myerson. Procurement: buyer with value vq, seller with cost cq^2. Optimal contract extract all surplus: Efficient quantity $q^e = v/(2c)$ at price $P^e = \alpha v q^e$ with $\alpha = 1/2$.
What if buyer has noisy estimates of q^e and α: $\widehat{q} = \eta q^e$ and $\widehat{\alpha} = \frac{1}{1+\xi}$

- **Fixed contract** (λ, μ) : produce $Q = \lambda \widehat{q}$ at price $P = \mu \widehat{\alpha} v Q$.

Optimal fixed contract (λ^*, μ^*) distorts quantity $\lambda^* < 1$.

- **Flexible contract** (λ, μ) : same as above, with option to modify quantities upward: $Q + \Delta$ at price $P + \Delta v$, with $\Delta > 0$

Flexible contrat at same terms is pareto superior, because it leads to efficient production iff demanded surplus $\widehat{S} = vQ - P$ is below $S = vq^e - P^e$. Buyer is always better off under a flexible contract.

- **Flexible constrat** (λ, μ) **with equal sharing:** $Q + \Delta$ at price $P + \Delta v/2$, $\Delta > 0$.

Quantity distorsions arise out of **caution with respect to own estimates**, rather than **incentive constraints** of different types of sellers.

With $\log \eta \sim N(0, \sigma^2)$ and $\log \xi \sim N(0, \sigma^2)$, $\sigma = 0.3$

We get:

Fixed contract: $(\lambda^*, \mu^*) \simeq (0.68, 1.08)$

Flexible contract: $\simeq (0.56, 0.66) \Rightarrow$ **3.7%** gain for the buyer

Flexible with equal sharing: $\simeq (0.58, 1) \Rightarrow$ **12.9%** gain for the buyer

Conclusion:

- comparative statics w.r.t. information less subject to a "prior" bias.
- robustness of standard models/insights
- Avoid irrealistic unraveling (centipede and alike), or coordination
- Focus on first order effect (auctions, centipede)
- Deal with more complex environment (information transmission/repeated games)
- Deal with bounded rationality through **structure put** on strategy space (what's the agent's frame of mind, how does this shape his naive/anchor strategy, which strategies does he compare?)
- Shape intuitions differently (quantity distorsions in contract theory)
- model incomplete learning (through family of noisy strategies)

Conclusion

Morris and Shin (2006, Chapter 3, page 56).
"*In principle, optimal strategic behavior should be analyzed in the space of all infinite belief hierarchies.*"

Attributed to John von Neumann
"There is no sense in being precise when you don't even know what you are talking about"

- **Probabilistic sophistication**.

Standard modelling has agents holding probabilistic beliefs:

$$z \equiv \widehat{\beta}_z \in \Delta(S)$$

$\rightarrow$ Then, if $\widehat{\beta}_z \neq \omega(\cdot \mid z)$, taking beliefs with caution seems a good idea (robust decision theory or maxmin)

But

1. Perceptions need not be probabilistic
2. Any perception, probabilistic or not, should probably be taken with caution,
   and the special form that caution takes depends on the nature of the perception assumed

Information aggregation:

$z_i = \eta_i v$
you don't know the common value v, but you know that if there are x% people more optimistic than you then $v \simeq z_i / \eta^x$ where η^x solves $\Pr(\eta_i > \eta^x) \equiv x$ (Pesendorfer Swinkels, Feddersen Pesendorfer)
When x is small (you are among the most optimistics), you should adjust downward your estimate...
... but the degree to which you should adjust is probably hard to determine.
By a modeling artifact, you know almost exactly your error η_i, just from being told x.
$->$ the model lacks aggregate uncertainty, about dispersion for example.

Folk theorems.

Cooperate=no price cut; Defect=price cut; $y \in \{1, 2\}$=who gets the consumer
$p = \Pr(y = 1 \mid CC)$, $\Pr(y = 1 \mid DC) = \Pr(y = 2 \mid CD) = q > p$.
p defines a particular public monitoring structure, with the following implication: when agents cooperate, market shares in the long run should be very close to p: if you defect a small fraction α of the time, this can be perfectly detected. Equilibrium strategies that approximate efficiency exploit that particular assumption.
- σ^* that yields approximate efficiency for all $p \in [\underline{p}, \overline{p}]$?
- Impossible. Essentially because playing C with proba $1 - \alpha$ rather than 1 when (CC) is called for cannot be detected (it generates same path as $p' = \alpha q + (1 - \alpha)p$ – hence C played most of the time if $p' < \overline{p}$)

Example of equilibrium strategies with $p = 1/2$

- Each starts with K credits.
- One credit transfer from j to i if j gets the consumer.
- One period punishment when a player has 0 credit –> $(0, 3.5)$ or $(3.5, 0)$ depending on who is punished.

$+$ 1 credit transfer to punished at the end of punishment period.

$\delta = 0.99$

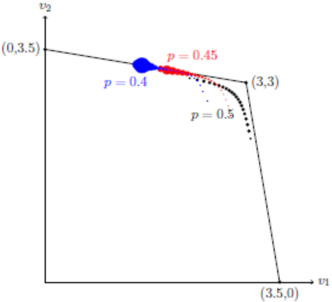


Figure 1: Location of W for $p = 0.5$, $p = 0.45$ and $p = 0.4$.