SELF-CONFIDENCE AND PERSONAL MOTIVATION

Roland Bénabou and Jean Tirole

Princeton University IDEI-Toulouse

Quarterly Journal of Economics (August 2002)

Introduction

"Egonomics" : role of *self-perception / self image*
in social & economic interactions

- Why is self-confidence / self-esteem so valuable to people?

- How do they try to maintain or enhance it?

- What implications for their information-gathering information /processing & decision-making?

•Develop simple economic model building on & unifying several themes in psychology $\Rightarrow$

•*Explain* variety of apparently "irrational" behaviors documented by psychologists:

unwillingness to know, self-handicapping,

distorted/selective memory or attention, self-deception.

•*Economic implications:* uninformed or biased decision-making by consumers, students, workers, investors. Principal-agent relationships, bargaining situations.

1. SELF-CONFIDENCE MODEL

- **Hypothesis 1 (imperfect information):** *People have imperfect knowledge about their enduring personal characteristics, or more generally about the costs and payoffs of their efforts*

Model: ability θ, cost c, payoff V.

Sources of information: experience (recall of past successes & failures), observation (attribution theory), feedback from others (possibly strategic), social comparisons (benchmarking)

Hypothesis 2 (cognitive ability): *The individual is an information processor who does not fool him/herself on average.*

•Bayesian inference. More generally, not too unaware of how he gathers & processes information. Richer implications than purely naïve view.

•Long-standing debate over possibility of *self-deception* or *motivated beliefs.*

Hypothesis 3 (relevance of self-knowledge): *Ability and effort are complementary factors in the production of performance.*

•Executive function view of the self.

• Self-confidence = key to *motivation*. Endogenous instrumental value.

• Alternative: “reflective consciousness”: self-esteem has (exogenous) consumption value.

• **Corollary (manipulative behavior):** *The individual's imperfect self-knowledge creates, for himself and/or others, incentives to manipulate information concerning the self.*

1) Manipulator = another person: parent, teacher, spouse, friend, colleague, boss,.... Principal-agent relationship, bargaining, etc.

$\Longrightarrow$ *Interpersonal strategies.*

2) Manipulator = oneself: conflict between different temporal "selves" due to time- inconsistent preferences

$\Longrightarrow$ *Intrapersonal strategies.*

We shall assume: *hyperbolic discounting* (Strotz 1956, Phelps-Pollack 1968, Laibson 1997). Otherwise, completely standard framework.

• Intrapersonal conflict: *procrastination:* enjoy today, work tomorrow.

• Information manipulation: substitute for commitment (build on Carrillo-Mariotti 1997)

2. BASICS AND A FIRST APPLICATION

Risk neutral individual with three-period horizon :

0 | 1 | 2

Date 0:

Date-0 "action":
- simplest = choice of information structure
- more generally, date-0 activity (date-0 flow payoff as well).

$\xrightarrow{\text{SELF-}}$ KNOWLEDGE

Date 1:

Undertakes date-1 activity (effort), or does not (no effort).

Cost of effort c

$$\Pr(\text{success}) = \begin{cases} \theta & \text{if effort} \\ 0 & \text{if no effort} \end{cases}$$

Date 2:

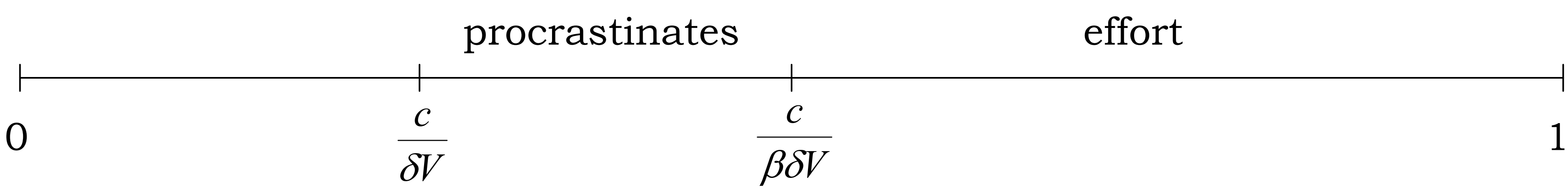
Benefit V if success (0 otherwise).

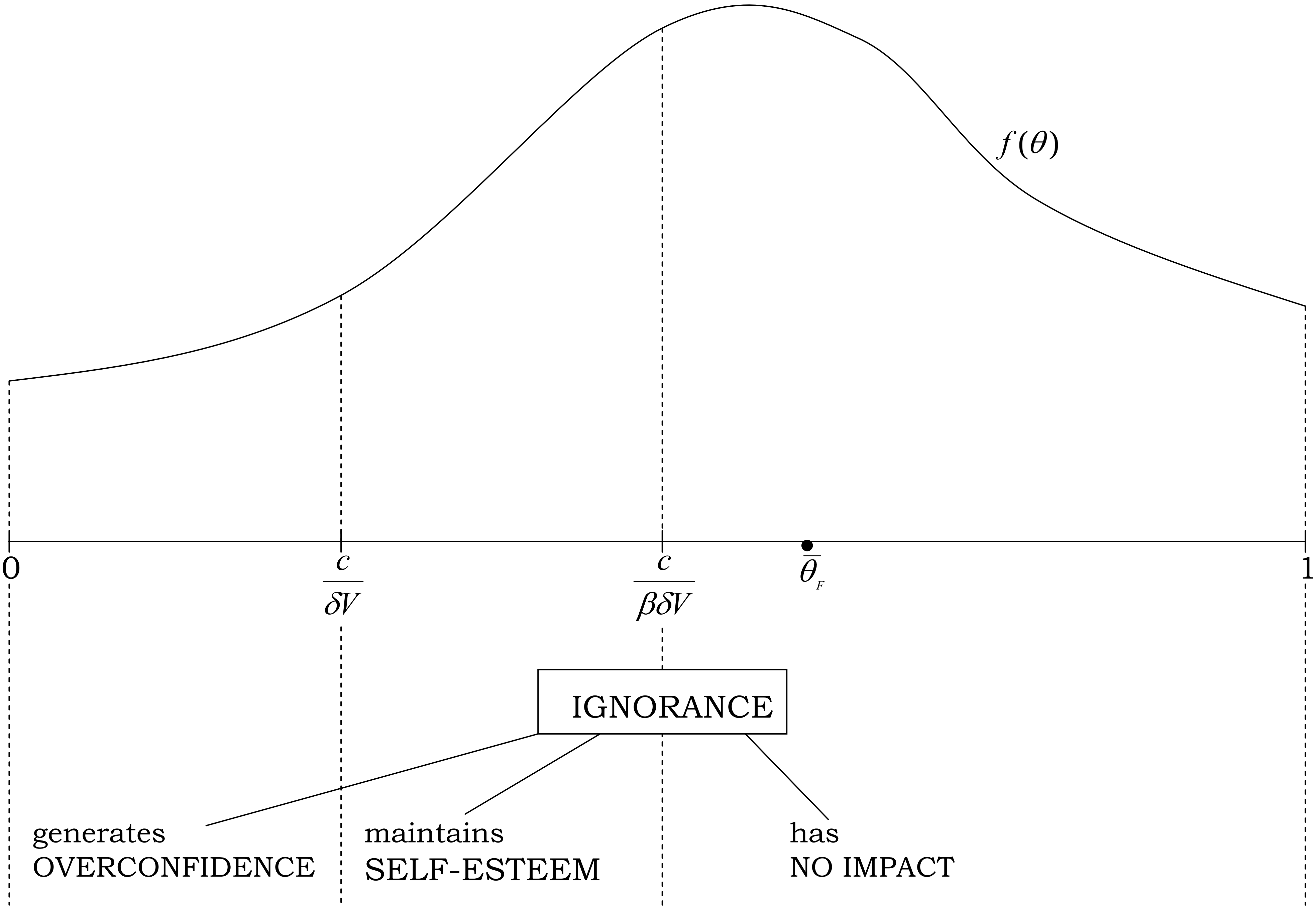
Self 1's payoff :

$$u_1 + \beta\delta u_2 = -c + \beta\delta\theta V \quad \text{if effort (0 otherwise).}$$

Self 0's payoff :

$$u_0 + \beta(\delta u_1 + \delta^2 u_2) = u_0 + \beta\delta(-c + \delta\theta V) \quad \text{if effort } (u_0 \text{ otherwise}).$$





$f(\theta)$
0
$\frac{c}{\delta V}$
$\frac{c}{\beta\delta V}$
$\bar{\theta}_F$
1
IGNORANCE
generates
OVERCONFIDENCE
maintains
SELF-ESTEEM
has
NO IMPACT

VALUE OF INFORMATION

(ex-ante belief manipulation)

• Self 0 has prior distribution *F(θ).* Chooses between remaining uninformed, or learning the actual θ (for free).

Case (a): Mean $\bar{\theta}_F$ *of F(θ) exceeds* $\dfrac{c}{\beta\delta V}$ → Self 1 works when uninformed.

Value of information = βδ ×

$$I_F = \int_0^{\frac{c}{\delta V}} (c - \delta\theta V)\; dF(\theta) - \int_{\frac{c}{\delta V}}^{\frac{c}{\beta\delta V}} (\delta\theta V - c)\; dF(\theta).$$

Avoid overconfidence; independent of β

Lose confidence-building, ↗ in β

• The lower β, the lower I_F

• The value of information can be *negative*:

agent *prefers not to know* his own ability, so as to preserve motivation.

Case (b): Mean $\bar{\theta}_F$ *lies below* $\dfrac{c}{\beta\delta V}$ → Self 1 shirks when uninformed.

$$I_F = \int_0^{\frac{c}{\beta\delta V}} (\delta\theta V - c)\; dF(\theta).$$

•Information is always valuable. The *lower* β*,* the more so.

• A more time-inconsistent agent is more “desperate” for information that might restore his self-confidence and motivation

⇒ Will *gamble for resurrection* of his self-esteem: undertake very hard/risky tasks with high informational content.

WHAT TYPES OF INDIVIDUAL ARE MOST EAGER TO MAINTAIN THEIR SELF-CONFIDENCE?

Definition (Comparison of self-confidence levels): *An individual with distribution F over ability θ has higher self-confidence than an otherwise identical individual with distribution G if the likelihood ratio is monotonic :*

$$\frac{f(\theta)}{g(\theta)} \quad \text{is increasing in } \theta.$$

Proposition 1: *If an individual prefers not to receive information in order to preserve his self-confidence, so will any individual with higher initial self-confidence.*

People with higher initial perceptions of their abilities (e.g., due to past successs) are the most insecure / defensive about ego-threatening information.

$$I_F = \int_0^{\frac{c}{\beta\delta V}} F(\theta)\, d\theta - \left(\frac{1-\beta}{\beta}\right) c\, F\left(\frac{c}{\beta\delta V}\right) \le 0$$

iff

$$\int_0^{\frac{c}{\beta\delta V}} \frac{F(\theta)}{F(c/\beta\delta V)}\, d\theta \le \left(\frac{1-\beta}{\beta}\right) \left(\frac{c}{\delta V}\right)$$

$$\text{Now}: F \succ_{MLRP} G \Rightarrow \frac{F(\theta)}{F(c/\beta\delta V)} < \frac{G(\theta)}{G(c/\beta\delta V)}$$

for all $\theta \le c/\beta\delta V$.

So $I_G < 0 \Rightarrow I_F < 0$

SELF-HANDICAPPING

- Commonly observed: Inadequate preparation, withholding of effort, drinking before task, choking under pressure…

- Self-setting of overambitious goals or tasks.

- Explanation: agent incurs costs to make performance less uninformative about his enduring ability.

- WHO IS MORE LIKELY TO SELF HANDICAP?

Let F be more self-confident than G.

Date-0 costs of *not* learning ability: $h_0(F)$, $h_0(G)$.

More complex: previous effect (more to lose from bad news), plus 2 new ones:

(a) Often $h_0(F) > h_0(G)$: more able have higher cost of botching date-0 task. But not always.

(b) $F\left(\dfrac{c}{\beta\delta V}\right) < G\left(\dfrac{c}{\beta\delta V}\right)$: self-handicapping cost unconditional. But impact on date-1 decision- making is less frequent if more self-confident (less likely to receive bad news): reduces return on h_0 "investment" in non-information.

An individual with priors *F(θ)* self-handicaps iff:

$$\left(\frac{1-\beta}{\beta}\right)c-\delta V\int_0^{\frac{c}{\beta\delta V}}\frac{F(\theta)}{F(c/\beta\delta V)}d\theta \geq \frac{h_0(F)}{\beta\delta F(c/\beta\delta V)}$$

Now, if $F \succ_{MLRP} G$ there are three effects:

- The more self-confident individual has *more to lose* from bad news: LHS is larger under *F* than under *G;*

- The more self-confident individual is *less likely* to receive bad news, so investing in ignorance pays off with lower probability: $F(c/\beta\delta V) \leq G(c/\beta\delta V)$, which tends to make RHS larger *F* than under *G;*

- *Direct cost* of self-handicapping may be higher under *F* (e.g., foregone output from failing current task); but need not be (e.g., less likely to fail exam):

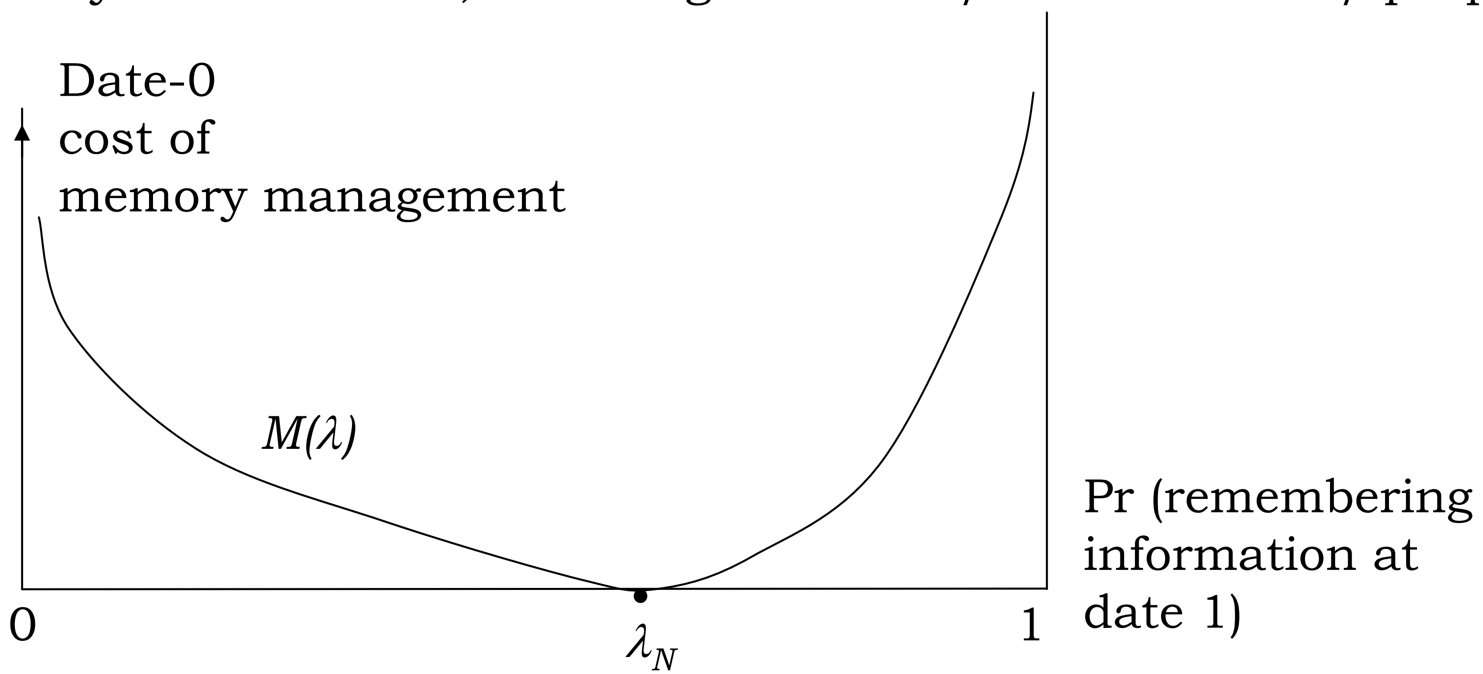
$$h_0(F) \gtrless h_0(G)$$

3. THE PSYCHOLOGICAL "IMMUNE SYSTEM" (ex-post belief manipulation)

- How do people deal with information that is damaging to self-esteem?
- Try to discount /challenge / explain away / forget / not think about it
- *Self-deception: c*entral issue is consistency between
  - hot attributions / motivated beliefs (self-esteem maintenance)
  - cold attributions (cognition, rationality).
- *Fact:* memory is imperfect, attention is limited, awareness is selective.

SELECTIVE MEMORY/ATTENTION/AWARENESS

- *Self-enhancing information:*
  - rehearsal (linger over positive feedback), manage cues,
  - keep hard evidence, select environment / interactions appropriately.
- *Self-threatening information:*
  - seek contradictory information, excuses; rehearse it.
  - create distraction (e.g., fight); repress memory; drink to forget.
  - destroy hard evidence, avoid negative cues / environments / people.

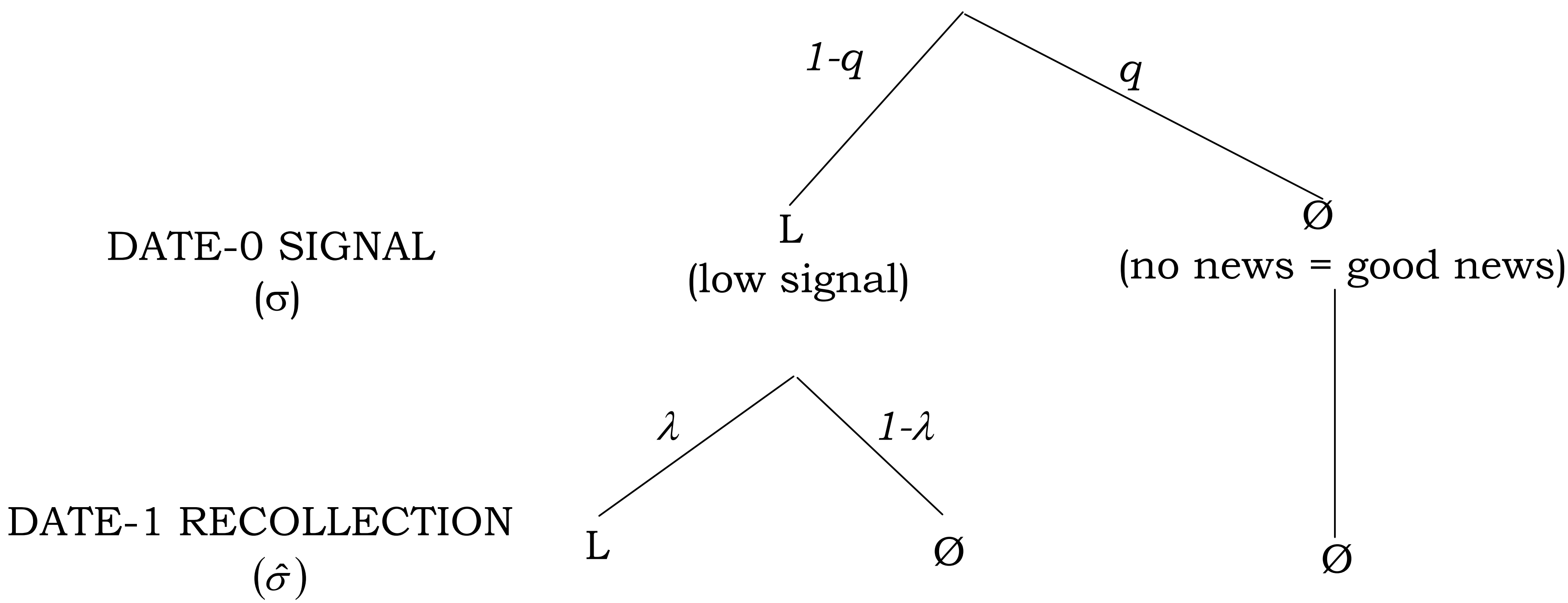


- **Hypothesis 4 (memory/awareness management):** The individual can, at a cost, increase or decrease the probability of remembering an event and/or its interpretation.

But we maintain Bayesian rationality:

- **Hypothesis 2' (awareness of selectivity of memory):** While the individual can manipulate his/her conscious self-knowledge, (s)he is aware that incentives exist that result in selective memory.

ENDOGENOUSLY SELECTIVE MEMORY AND SELF-DECEPTION



- Self 0's expected ability given date-0 signal σ:

$$\equiv \begin{cases} \theta_L \text{ if } \sigma = L \\ \theta_H \text{ if } \sigma = \phi \end{cases} \qquad\qquad (\theta_H > \theta_L)$$

- Self 1's inference problem: assess the reliability of "no recollection"

$$r^* = \frac{q}{q + (1-q)(1-\lambda^*)}$$

- Self 1's expected ability : θ_L if $\hat{\sigma} = L$, but if $\hat{\sigma} = \phi$, then :

$$\theta(r^*) = r^* \theta_H + (1 - r^*)\,\theta_L$$

• Self 0 faces some uncertainty about date-1 cost c : distribution $\Phi(c)$

Uncertainty resolved at date 1.

• Self 0's payoff (gross of memory management cost), when Self 1 has recollection $\hat{\sigma}$:

$$\beta\delta \int_0^{\beta\delta \mathrm{E}(\theta|\hat{\sigma})} \left(\delta \mathrm{E}\left(\theta \mid \sigma\right) - c\right) d\Phi(c)$$

• When Self 0 receives self-threatening news $(\sigma = L)$:

$$U_T(\theta_L) = \beta\delta \int_0^{\beta\delta\theta_L} \left(\delta\theta_L - c\right) d\Phi(c)$$

$$U_C(\theta_L | r^*) = \beta\delta \int_0^{\beta\delta\theta(r^*)} \left(\delta\theta_L - c\right) d\Phi(c)$$

Confidence building v.s. overconfidence.

Confidence building v.s. overconfidence:

$$U_C(\theta_L | \; r^*) - U_T(\theta_L) =$$

$$= \left(\int_{\beta\delta\theta(r^*)}^{\delta\theta_L} (\delta\theta_L - c) d\,\Phi(c) \;-\; \int_{\delta\theta_L}^{\beta\delta\theta(r^*)} (c - \delta\theta_L) d\,\Phi(c) \right).$$

Confidence-building, gain ↘ in β

Overconfidence, loss ↗ in β

PERFECT BAYESIAN EQUILIBRIUM

(1) *Optimal memory / awareness management*

$$\lambda^* \in \underset{\lambda}{\operatorname{argmax}} \left\{ \lambda \mathrm{U_T}(\theta_L) + (1-\lambda)\, U_C(\theta_L \mid \; r^*) - M(\lambda) \right\}$$

(2) *Rationality*

$$r^* = \frac{q}{q + (1-q)(1-\lambda^*)}$$

Solving the game:

•Start with case of costless memory/awareness management: $M \equiv 0$. *Proposition 2, Figure 1.*

•Incorporate costs afterwards. *Proposition 3, Figures 2.1 to 2.3.*

COSTLESS MEMORY/AWARENESS MANAGEMENT

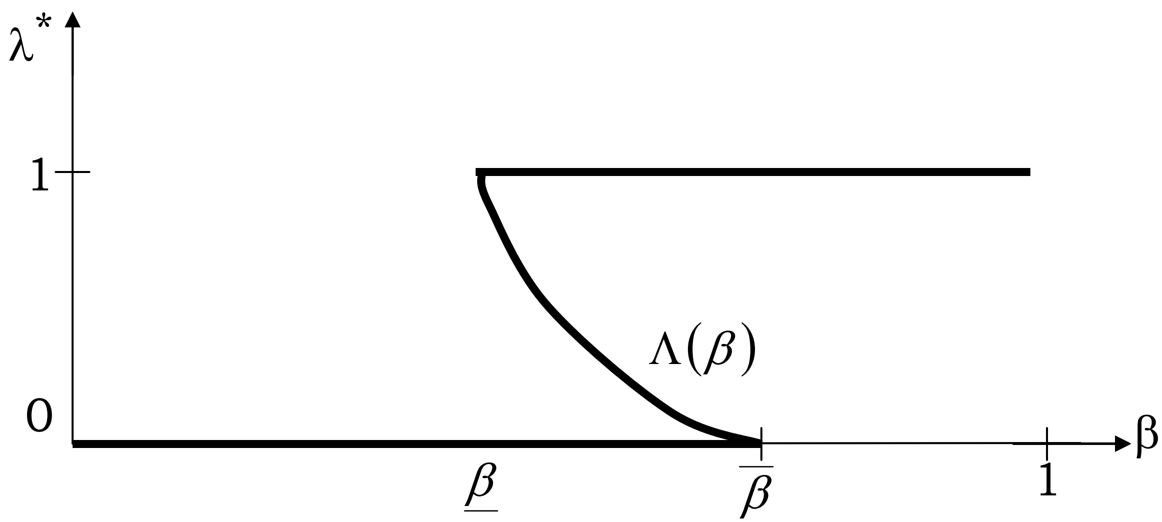


Figure 1: The equilibrium with costless recall/repression

(a) Impact of β:

- confidence-building effect increases with time-inconsistency (1- β)
- overconfidence effect decreases with it.

(b) Self-fulfilling expectations.

The higher the degree of repression $1\text{-}\lambda$ by Self 0,

$\rightarrow$ the more Self 1 discounts "no bad news"

$\rightarrow$ the lower the risk of overconfidence

$\rightarrow$ the greater Self 0's incentives to censor bad news.

- Same person can be in different “modes” of self-esteem management:
  - Active self-esteem maintenance / “looking on the bright side”, denial
  - Realistic / “accept who you are”

- Very different from *ex-ante* decisions not to become informed, self-handicap, etc. Such “strategic ignorance” (Carillo-Mariotti 97) always yields unique outcome, which is optimal ex-ante.

- *Costly memory management:* similar conclusions.
  Comparative statics with respect to repression cost.

Costly Memory/Awareness management:

- Let the memory cost function be:

$$M(\lambda) = a(1-\ln\lambda) + b\big(1-\ln(1-\lambda)\big).$$

Thus $\lambda_{\mathrm{N}} = a/(a+b)$. Let the distribution of effort costs be uniform.

- For any candidate equilibrium belief *r* of Self 1, can then compute:

\- The memory strategy λ*(r)* of of Self 0 consistent with Bayes' rule.

\- Self 0's marginal incentive to forget above/below the rate λ*(r)* :

$$\psi(r,\beta) = r(\Delta\theta)\left(\frac{\beta^2\delta^3}{\bar{c}}\right)\left((1-\beta)\theta_L - \frac{\beta r}{2}(\Delta\theta)\right) - r\left(\frac{1-q}{q}\right)\left(\frac{(a+b)q-(aq+b)r}{(1-r)(1-q)}\right),$$

where $\Delta\theta \equiv \theta_H - \theta_L$. Sign of $\psi(\mathrm{r},\beta)$ = that of a 3$^{\mathrm{d}}$ deg ree polynomial in r.
Either one or three solutions $r \in [q,1]$.

- **Proposition 3:** case $b = 0$ (only repression/forgetting is costly)

$\Rightarrow$ complete solution.

WELFARE ANALYSIS

•Two views in psychology literature (academic & popular) :

• *SELF-HELP*: “think positive,” “avoid negative thoughts,” “enhance your self-esteem,”…

• *KNOW YOURSELF*: “accept who you are,” “be honest with yourself”, …
Either as alternative equilibrium strategy, or possible commitment.

Which is best (ex-ante)?

• Externality among self-esteem states:
- selective memory/ awareness helps when *low* self-esteem
- but creates *self-distrust* which is hurtful when *high* self-esteem

• *Proposition 4:* says when either effect dominates $\Rightarrow$ "self-traps".

May want to : - be free to actively manage self-esteem/memory,
- commit to always face the truth.

Is self-deception ultimately beneficial or harmful?

• Gain: raises self-confidence/motivation when it would otherwise be low, i.e. from θ_L to $\theta(r^*)$, following $\sigma\sigma = L$.

Benefit from self-enhancement ≥ overconfidence cost + management cost ⇒ expected gain ≥ 0 when σ = *L:*

$$\lambda \mathrm{U}_{\mathrm{T}}(\theta_L) + (1-\lambda)\, U_C(\theta_L \mid r^*) - M(\lambda) \geq \mathrm{U}_{\mathrm{T}}(\theta_L) - M(1)$$

•But: lower self-confidence/motivation when it would otherwise be high, i.e. only $\theta(r^*)$ instead of θ_H, followin $g\,\sigma = \varnothing$.

This is the *self-doubt* effect: Self 1 always discounts good news (absence of bad news), so when they are in fact true, ie. when σ = ∅*,* there results a loss > 0:

$$U_T(\theta_H \mid r^*) = \beta\delta \int_0^{\beta\delta\theta\,(r^*)} (\delta\theta_H - c)\, d\,\Phi(c) < U_T(\theta_H \mid 1)$$

•Let $W(\lambda^*, r^*)$ denote ex-ante welfare in when Self 0 plays λ^* and Self 1's beliefs are r^* .

•Compare equilibrium $W(\lambda^*, r^*)$ with truthful recall equilibrium (or commitment strategy) $W(1, 1)$.

•Difference is:

$$\Delta W(\lambda^*, | r^*) = (1-q)\left((1-\lambda^*) \int_{\beta\delta\theta_L}^{\beta\delta\theta(r^*)} (\delta\theta_L - c) d\Phi(c) - M(\lambda^*) + M(1) \right)$$
$$- q\left(\int_{\beta\delta\theta(r^*)}^{\beta\delta\theta_H} (\delta\theta_H - c) d\Phi(c) \right).$$

•First term is ≥ 0, second term is ≤ 0, and generally < 0.

• When $M \equiv 0$, the first the first term = 0 in *any* mixed strategy eqbm.

• When $M \not\equiv 0$, not true: "surplus" $M(1) - M(\lambda^*) - (1-\lambda^*) M'(\lambda^*) > 0$

•Provide examples where $\lambda^* = 0$ eqbm. is superior, or where $\lambda^* = 1$ eqbm. (or commitment) is better.

Proposition 4: *If the function*

$$\Gamma(Z,\beta) \equiv \int_0^Z (Z - \beta c)\, d\Phi(c)$$

is concave in z, then ex-ante welfare is higher if all bad news are censored from memory than if they are always recalled. The function Γ*(Z,β) is concave in Z iff the cost density decreases fast enough*

$$-\frac{\partial \ln \varphi(c)}{\partial \ln c} > \frac{2-\beta}{1-\beta} \quad \text{for all} \quad c \in [\underline{c}, \bar{c}].$$

It is convex if the inequality is reversed.

- The value of information (always knowing the true θ) rather than having uninformed posterior θ*(q)* is equal to δ ×

$$\Gamma(\beta\delta(q\theta_H + (1-q)\theta_L), \beta) - q\Gamma(\beta\delta\theta_H, \beta) - (1-q)\Gamma(\beta\delta\theta_L, \beta).$$

BELIEFS AND MAKE-BELIEFS

(1) Distribution of actual (expected) abilities:

– a fraction $1 - q$ of individuals are of low ability, $\theta = \theta_L$ (more generally, low expected ability), having received a negative signal, $\sigma = L$.

– a fraction q, having received $\sigma = \emptyset$, have high (expected) ability, $\theta = \theta_H$

Average ability is $q\,\theta_H + (1\text{-}q)\,\theta_L = \theta^*(q)$.

Assume $q < 1/2$, so that median ability is θ_L.

(2) Distribution of self-evaluations:

– a fraction $(1\text{-}q)(1\text{-}\lambda)$ *overestimate* their ability by $\theta(r^*) - \theta_L = r^*(\theta_H - \theta_L)$

– a fraction q *underestimate* their ability by $\theta_H - \theta(r^*) = (1 - r^*)(\theta_H - \theta_L)$

Therefore:

• If the costs of repression/forgetting a are low enough, can have $(1\text{-}q)(1\text{-}\lambda) > 1/2$, perhaps even $(1\text{-}q)(1\text{-}\lambda) \lesssim 1$. Thus, most people believe themselves to be *more able than they actually are, more able than average,* and *more able than the majority* of individuals.

• Adding those who had truly received the signal $\sigma = \emptyset$, the fraction of the population who think they are better than average is even larger: $1\text{-}\lambda\ (1\text{-}q)$.

• The remaining minority think, correctly, that they are worse than average; ⇒ low motivation, undertake little, achieve even less. Fit experimental findings of depressed people as *"sadder but wiser'"* realists, compared to their non-depressed counterparts who are much more likely to exhibit self-serving delusions.

•Bayes' law does not constrain the *skewness* in the distribution of biases (Carrillo and Mariotti (1997)). Can be very skewed.

• Only requires that the *average* bias across the $(1-q)(1-\lambda)$ optimists and the q pessimists be zero: $(1-q)(1-\lambda)\, r^* - q(1- r^*) = 0.$

Intuitively: Tradeoff between relative proportions of overconfident versus underconfident agents in the population, and their respective degrees of over- or under-confidence.

• Zero average bias does not preclude self-esteem maintenance from having *aggregate* economic effects. They do affect total effort, output and welfare.

•What explanatory power would be gained/lost by departing from Bayesian rationality?

More naïve posteriors:

Let us allow the agent at date 1 to remain less than fully aware of the fact that, at date 0, any negative signal σ = *L* would have been forgotten with probability λ. Now fails to remember not just *what* he forgets, but also *that* he forgets:

$$r_{\pi}(\lambda) = \frac{q}{q + \pi(1-q)(1-\lambda)}$$

$\pi \in [0,1]$ parametrizes the extent of Bayesian rationality.

See Figure 3 for analysis of equilibria. Shows that:

(a) Models' results are robust, as long as people are reasonably (even if not fully) objective in their inference.

(b) While explanatory power gained by departing from Bayesian rationality is somewhat limited (perception biases need no longer sum to zero, aggregate real effects may be magnified).
much can be lost if the departure is too drastic: without sufficient introspection, one can not account for "self-traps"".

DEFENSIVE PESSIMISM

- People most often concerned with boosting their self-esteem, but there are also instances where seek to *minimize* their achievements, or convince themselves that the task at hand will be *difficult or unlikely to succeed,* rather than easy:

- Student studying for exams may discount his success on previous ones as attributable to luck or lack of difficulty.
- A young researcher or lawyer may understate the value of his prior achievements, as compared to what will be required to obtain tenure or partnership.
- A dieting person who lost a moderate amount of weight may decide that he "looks fatter than ever", no matter what others or the scale may say.

Such "defensive pessimism" can be captured with a very simple variant of our model. These are situations where:

(a) the *underlying motive* for information-manipulation is unchanged: alleviate the shirking incentives of future selves.

(b) ability is a *substitute* rather than a complement to effort in generating future payoffs.

This clearly gives the agent an incentive to discount, ignore /repress signals of *high* ability, such as previous achievements, as these would increase the temptation to "coast" or "slack off"".

Substitutability may arise directly in the performance "production function". More interestingly, it will typically occur when

the *reward for performance* is of a *"pass-fail"* nature.

E.g., obtaining a diploma, making a sale, being hired or fired (tenure, partnership), marriage and divorce.