Learning Self-Control

S. Nageeb Ali

University of California, San Diego

March 1, 2011

Motivation

- Temptation → Commitment.
- But what does a DM know about his temptations?
- E.g. in quasi-hyperbolic discounting,

$$U(u_t, u_{t+1}, ...) = E_t \left(u_t + \beta \sum_{\tau=1}^{\infty} \delta^t u_{t+\tau} \right).$$

- Usual practice fixes DM's beliefs at $\hat{\beta}$.
    - Sophistication: $\hat{\beta} = \beta$.
    - Naivete: $\hat{\beta} = 1$.
    - Partial naivete: $\hat{\beta} \in (\beta, 1)$.
- Beliefs influence commitment (and dynamic) choice.

Conceptual Issues

- Partial sophistication: beliefs incompatible with experience.
- Difficult to understand when solution concept is appropriate.

"I think that behavioral economics would be well served by concerted attempts to provide learning-theoretic (or any other foundations) for its equilibrium concepts. At the least, this process might provide a better understanding of when the currently used concepts apply...." – Drew Fudenberg

Conceptual Non-issues

Leaves open big questions:

- (When) Is *Sophistication* = *Long-run limit of learning*?
- How does the *technology* of commitment affect learning?
- What is the pathway from *Naivete* → *Limit of learning*?
- Who becomes sophisticated and who remains naive?

Conceptual Non-issues

Leaves open big questions:

- (When) Is *Sophistication* = *Long-run limit of learning*? ✓
- How does the *technology* of commitment affect learning?✓
- What is the pathway from *Naivete* → *Limit of learning*?
- Who becomes sophisticated and who remains naive?

Dual Selves

- Long-run *Planner* chooses a menu in each period.
- Myopic *Doer* picks from menu based on i.i.d. taste-shock and persistent temptation.
- Planner does not know extent of Doer's temptation, but learns over time through Bayesian updating.

Learning

- Commitment vs. flexibility → Experimentation.
    - Flexibility necessary for learning.
    - But is costly if Doer has strong temptations.
- Learning may be incomplete.
- Necessary and sufficient condition on commitment technology for as-if sophistication.

FCD

Full Commitment Distinguishability

$\equiv$ for every (θ_G, θ_B),

there exists a commitment technology such that Planner can fully commit θ_B and not θ_G.

Consumption-savings setting: FCD ✓
Addiction: ~~FCD~~
Costly self-control / willpower: FCD ✓(sometimes)

Main Result: Sufficiency

FCD

$\Downarrow$

Globally adequate learning regardless of δ and prior.

Globally adequate learning
$\equiv$ For every Doer type, Planner eventually attains same payoffs as fully informed Planner.

Main Result: Necessity

$\neg$FCD

$\Downarrow$

For every δ, learning is inadequate for some open set of priors.

Inadequate learning
$\equiv$ Strictly positive measure of types for which Planner fails to attains same payoffs as fully informed Planner with strictly positive probability.

Related Literature

- Dual Selves: Thaler and Shefrin (1981), Bernheim and Rangel (2004), Fudenberg and Levine (2006, 2010a,b).
- Commitment vs. Flexibility: Gul and Pesendorfer (2001, 2005), Amador, Werning, and Angeletos (2006).
- Learning: Easley and Kiefer (1988), Aghion, Bolton, Harris, and Julien (1991), Fudenberg and Levine (1993a,b).
- Partial naivete: Many papers here; you've either read or written them anyway.

Example

For *context*, consider the "Gym Environment":

- In each period, DM chooses to work out $(a_t = 1)$ or not $(a_t = -1)$.
- Firm charges lump-sum L in each period.
- DM rejects contract: payoff of 0 in that period.
- DM accepts contract:
  - Pays lump-sum.
  - Immediate cost c_t uniform from $[0, 1]$.
  - (Delayed) Benefit of $b \in [0, 1]$.

Doer: Exercising Options

Doer chooses whether to exercise if contract is signed:

- No temptation: $c_t \leqslant b$.
- Temptation: $c_t \leqslant \theta b$ for $\theta < 1$.
- In either case, Doer is myopic.

Planner: Choosing Contract / Menu

Planner pays for membership iff:

$$\mu_0 b\left(b-\frac{b}{2}\right)+(1-\mu_0)\,\theta b\left(b-\frac{\theta b}{2}\right)\geqslant L.$$

Standard Sophisticated about temptation
Uncertain about Doer's Type

Learning

If Planner signs a contract, he can learn from Doer's exercise choices.

- Suppose Planner observes a_0 but not c_0. (Will relax later).
- If Planner signed contract at $t = 0$:

$$\frac{\mu_1}{1-\mu_1} = \left(\frac{\mu_0}{1-\mu_0}\right) \times \underbrace{\left(\frac{b}{\theta b}\right)}_{a=1} \underbrace{\left(\frac{1-b}{1-\theta b}\right)}_{a=-1}$$

Dynamic Programming

$$V(\mu) = \max \left\{ \begin{array}{l} \text{Today's Value} + \text{Discounted Expected Value} \\ 0 \end{array} \right\}$$

Solution: Planner enrolls iff $\mu \geqslant \mu^*$.

Beliefs are endogenous but converge a.s.

Eventual Beliefs and Choices

Theorem

1. *If the Doer is tempted* (θ), *Planner eventually stops enrolling a.s.*

$$\Pr\left(\lim_{t\to\infty}\mu_t < \mu^*|\theta\right) = 1.$$

2. *If the Doer is not tempted, with positive probability, the Planner's always enrolls and with positive probability, stops enrolling.*

$$\Pr\left(\lim_{t\to\infty}\mu_t \in \underbrace{[0,\mu^*)}_{Ineffiency} \cup \{1\} | \textit{Not tempted}\right) = 1.$$

Partial Commitments

Suppose that Doer of either type can be *nudged* to exercise through rewards.

- Exercise iff $c_t \leqslant \theta b + z$.

Planner can sign a *commitment contract* in which

- Planner sets $z = b - \theta b$.
- Pays upfront $(1 - \theta)(\theta b)^2$.

Contract: Zero expected transfers, and induces first-best when Planner is confident that Doer's type is θ.

Globally Adequate Learning

Fact

Commitment contracts $\Rightarrow$ Globally adequate learning.

General Framework

Generalizes examples in several ways:

- Continuum of types.
- Partial commitments come in two forms: *Nudges* and *Menus*.
    - Nudges *influence* payoffs of Doer, e.g., Antabuse, commitment contracts, promises and peer-based shame mechanisms.
    - Menus *restrict* choices of Doer, e.g., illiquid assets.
    - Paper studies both; for talk, will focus on menus.
- Planner can observe signals of past taste-shocks.

Setting

- Action $a_t \in A \equiv [\underline{a}, \overline{a}]$ is chosen in period $t = 0, 1, 2, ...$
- In each period, state $s \in \mathcal{S} \equiv [\underline{s}, \overline{s}]$ is drawn, iid with cdf F.

Planner's Payoffs

Planner has payoffs $u(a, s)$ that are

- Strictly quasi-concave in a for each state s,
- Satisfy strict single crossing in (a, s)

$$\Rightarrow a_P(s) \text{ is single-valued and non-decreasing in } s.$$

Assume unique $\hat{a}$ that is *ex ante* optimal.

Commitment

Planner chooses a *menu*, a closed and non-empty subset of actions, M.

- $\mathcal{F}$ is the set of all *logically feasible* menus.
- $\mathcal{M}$ is the set of all *economically feasible* menus.
- $\mathcal{M}$ is closed (in the Hausdorff metric topology).
- $\mathcal{M}$ contains full flexibility ($M = A$) and full commitment ($M = \{\hat{a}\}$).

Doer

Doer of type θ solves

$$Max_{a \in M} W(a, s, \theta)$$

where W is:

- Continuous, strictly quasi-concave
- Satisfies strict single-crossing property in (a, s) and (a, θ).

$$\Rightarrow a_D(s, \theta, M) \text{ is non-decreasing in } s \text{ and } \theta.$$

Temptation

Assumption

The Doer is tempted to undertake lower actions than the Planner:

$$u(a, s) \succeq W(a, s, \theta)$$

by the single-crossing condition for every θ.

Full Information Benchmark

$$\pi(\theta, M) = \int_{\mathcal{S}} u\left(a_D(s, \theta, M), s\right) dF.$$

$$\pi^*(\theta) \equiv \max_{M \in \mathcal{M}} \pi(\theta, M).$$

$$\hat{\pi} = \int_{\mathcal{S}} u(\hat{a}, s)\, dF.$$

Uncertainty, Learning, and Feedback

- Planner begins with prior μ_0.
- After each period, Planner obtains signal about prior state.
- History h^t denotes history of commitments, actions, and signals in periods $0, \ldots, t-1$.
- μ_t is relevant posterior.

Dynamic Programming

$$V(\mu;\delta) = \max_{M\in\mathcal{M}} \left\{ (1-\delta) \int_{\Theta} \pi(\theta, M)\, d\mu + \delta \int_{P(\Theta)} V(\mu';\delta)\, dQ(\mu, M) \right\}$$

Adequacy

Definition

Learning is *adequate* for a type θ if the Planner's payoffs when uncertain eventually converge to the full information benchmark.

$$\Pr\left(\lim_{t\to\infty} V(\mu_t;\delta) = \pi^*(\theta)\,|\theta\right) = 1.$$

Role of Commitments in Learning

- If the Planner retains some flexibility for Doer to choose different actions, empirical frequency of actions identify type.
- Full commitment impedes learning: for some types, the Planner may wish to fully commit.

$$\widehat{M}(\theta) = \{M \in \mathcal{M} : a_D(s, \theta, M) = \hat{a} \text{ for almost all } s\}$$

$$\widehat{\Theta} = \{\theta \in \Theta : \pi^*(\theta) = \hat{\pi}\}$$

Full Commitment Distinguishability

Definition

FCD is satisfied if for almost every $\hat{\theta}$ in $\hat{\Theta}$ and every θ not in $\hat{\Theta}$,

$$\hat{M}(\hat{\theta}) \not\subseteq \hat{M}(\theta)$$

Main Results

Sufficiency

Theorem

If an environment satisfies FCD, then for all priors and discount factors, learning is globally adequate.

Intuition: If $\hat{\theta}$ and θ are in support, use commitment that distinguishes them. Repeat.

Main Results

Necessity

Suppose that $\mathcal{M} = \mathcal{F}$, or is the set of all feasible interval menus.

Theorem

If an environment fails for FCD, then for all discount factors, learning is inadequate for some open set of priors.

Intuition: A failure of FCD $\Rightarrow$ costly experimentation.

Role of Patience

Theorem

Regardless of FCD, for every μ_0,

$$\lim_{\delta \to 1} V(\mu_0; \delta) = \int_{\Theta} \pi^*(\theta)\, d\mu_0$$

Force similar to Aghion, Bolton, Harris, and Jullien (1991), and Fudenberg and Levine (1993b).

- Approximate payoffs with a finite set of commitments.
- Choose each commitment a large number of times.
- Settle on commitment that appears optimal.

Difficult to distinguish patience from naivete through menu choice.

Application to Savings

$$E\left[\sum_{t=0}^{\infty} \delta^t u_t U(c_t)\right]$$

- $u_t \in [\underline{u}, \overline{u}]$ with $0 < \underline{u} < \overline{u}$ and $E[u] = 1$.
- $U(c_t)$ is a CRRA utility function with coefficient $\sigma \geqslant 0$.
- Planner begins with wealth y_0, and future wealth, $y_t = R\left(y_{t-1} - c_{t-1}\right)$.

Tempted to Overconsume

$$E\left[\sum_{t=0}^{\infty} \delta^t u_t U(c_t)\right]$$

- Planner's Solution: $c_P(u) \times y$.
- Doer's (ideal) consumption: $c(u, \theta) \times y$
  - Strictly decreasing in θ, where $\theta \in [\underline{\theta}, 1]$.
  - Highest type has no bias: $c(u, 1) = c_P(u)$.
  - Can capture present-bias where Doer has discount factor $\theta\delta$.

Commitment: Illiquid Assets

- Illiquid assets are a natural commitment technology to consider.
- Planner purchases $s_t \times y_t$ of illiquid wealth at the beginning of time t.
- Constrains Doer to choose from $[0, (1 - s_t)y_t]$ in period t.

If Planner could commit to singleton, set $\hat{s}$ to be the optimal full commitment.

$$\hat{s} = \delta^{\frac{1}{\sigma}} R^{\frac{1-\sigma}{\sigma}} \quad (1)$$

FCD in Savings Environment

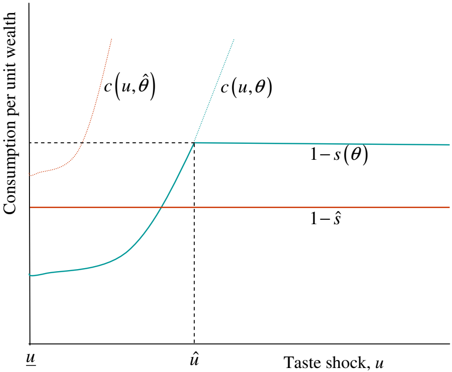

Figure I: Consumption Caps

Dotted curves indicate the Doer's ideal consumption for each taste-shock, and solid lines indicate the Doer's actual consumption when the Planner selects commitment optimally.

Result

Theorem

Learning is globally adequate for all priors and discount factors.

Caveat: Learning is still costly and can make DM poorer.

Conclusion

- Paper offers condition for Bayesian learning to yield sophistication.
- Results highlight dynamic benefits of partial commitments.
- Methodologically, framework shows tractability of dual self models.

Conclusion

- Learning can fail when individuals aren't Bayesian, have bounded memories, and have self-serving beliefs.
- Also, learning about new environment and self-control is hard.
- Commitment may have other costs that are not modeled (and may require self-control).
- Empirical challenges in identifying self-awareness from choices, but much exciting work underway.