

Models of Thinking, Learning, and Teaching in Games

By COLIN CAMERER, TECK HO, AND KUAN CHONG*

Noncooperative game theory combines strategic thinking, best-response, and mutual consistency of beliefs and choices (equilibrium). Hundreds of experiments show that in actual behavior these three forces are limited, even when subjects are highly motivated and analytically skilled (Camerer, 2003). The challenge is to create models that are as general, precise, and parsimonious as equilibrium, but which also use cognitive details to explain experimental evidence more accurately and to predict new regularities. This paper describes three exemplar models of behavior in one-shot games (thinking), learning over time, and how repeated “partner” matching affects behavior (teaching) (see Camerer et al., 2002b).

I. Thinking

The “cognitive hierarchy” (CH) model starts with 0-step players who randomize equally across strategies; k -step players ($k \geq 1$) believe all other players use only 0 to $k - 1$ steps. We assume the k -type’s beliefs $g_k(h)$ about the proportions of lower-step h types are the normalized true distribution,

$$g_k(h) = f(h) \left/ \sum_{l=0}^{k-1} f(l) \right.$$

$$g_k(h) = 0 \quad \text{for } h \geq k$$

[†] *Discussants:* C. Mónica Capra, Washington and Lee University; Colin Camerer, Caltech; David Cooper, Case Western Reserve University.

* Camerer: Caltech, Pasadena, CA 91125 (e-mail: camerer@hss.caltech.edu); Ho: Haas School, University of California, Berkeley, CA 94720 (e-mail: hoteck@haas.berkeley.edu); Chong: NUS Business School, National University of Singapore (e-mail: bizcjk@nus.edu.sg). This research was supported by NSF grant SES-0078911.

so they have “partially rational” expectations (and $g_k(h)$ approaches $f(h)$ as k grows large). These beliefs are used by k -step thinkers to compute expected payoffs and choose best responses.

Working memory constraints (and doubts about rationality of others) suggest that more and more thinking steps are increasingly rare, expressed by $f(k)/f(k - 1)$ proportional to $1/k$. This assumption implies that $f(k) = e^{-\tau} \tau^k / k!$, the Poisson distribution, where τ is the mean and variance of the number of thinking steps. Axioms and data from more than 80 games suggest that τ is between 1 and 2. Assuming $\tau = 1.5$ makes the model more precise than Nash equilibrium (because it predicts a specific distribution of strategies when there are multiple equilibria). We guess that $\tau = 1.5$ will *never* predict one-shot experimental data worse than Nash equilibrium (under typical lab conditions), and will almost always predict more accurately.

Earlier models of limited thinking include Ken Binmore (1988), Dale Stahl (1993), and Rosemarie Nagel (1995). Among one-parameter models, quantal response equilibrium (QRE) weakens best-response but retains mutual consistency (e.g., Jacob Goeree and Charles Holt, 2002); Mónica Capra’s (1999) theory weakens best-response and consistency simultaneously.

The goal of all these models is to explain why and when equilibration is limited, and when instant equilibration occurs. Two classes of games illustrate how the CH model does this.

In a “ p -beauty contest” (PBC) game, players choose numbers in an interval, say $[0, 100]$. For $p = 2/3$, the player whose number is closest to two-thirds of the average wins a fixed prize. The PBC game is a tool for inferring steps of thinking from choices. One-step players expect an average of 50 (from random choice by 0-steppers) and choose 33. Two-step players choose a best-response to a mixture of 33 and 50, around 26 if $\tau = 1.5$. Equilibrium requires mutual consis-

TABLE 1—DATA AND ESTIMATES OF τ IN PBC GAMES
(EQUILIBRIUM = 0)

Subjects/game	Data		Steps of thinking
	Mean	SD	
Computer scientists	18	17.4	3.8
Game theorists	19	21.8	3.7
Caltech students	23	11.1	3.0
Newspaper	23	20.2	3.0
Portfolio managers	24	16.1	2.8
Economics Ph.D. class	27	18.7	2.3
Caltech $g = 3$	22	25.7	1.8
High school	33	18.6	1.6
1/2 mean	27	19.9	1.5
70-year-olds	37	17.5	1.1
German students	37	20.0	1.1
CEO's	38	18.8	1.0
Game $p = 0.7$	39	24.7	1.0
Caltech $g = 2$	22	29.9	0.8
PCC $g = 3$	48	29.0	0.1
Game $p = 0.9$	49	24.3	0.1
PCC $g = 2$	54	29.2	0.0
Mean:			1.56
Median:			1.30

Notes: "Newspaper" refers to data from three newspaper contests; $g = 2$ and $g = 3$ denote two- and three-player games; "1/2 mean" is a game in which $p = 1/2$; "PCC" is Pasadena City College; "high school" refers to data from high-school students. See Camerer et al. (2000a) for details.

tency, which implies a choice $x^* = (2/3)x^*$, or zero (corresponding to infinite τ in CH). As a practical matter, choosing zero is being too smart for one's own good. The game requires both logical skill and "social intelligence," or knowing how much thinking others will do (and expect).

Table 1 shows the mean and standard deviations of the number choices in beauty-contest games with different subject pools (generally playing for \$20) and group sizes g . The mean number is usually 20–40, although analytical-skilled groups choose lower numbers. Estimates of τ which produce the closest fit to the mean are mostly 1–3 with mean and median around 1.5. Except for some interesting outliers, estimates of τ are remarkably close across a surprising range of subjects (e.g., CEO's, 70-year-olds, and high-school students). Fixing $\tau = 1.5$ predicts a mean of 27, which is too high or low in some groups but is always much more accurate than the Nash prediction of zero.

While limited-thinking models do a good job of explaining limited equilibration in PBC

games, the main challenge is to see whether the same model can explain quite different games. An instructive example is asymmetric matching-pennies games. Two players (1 and 2) choose H or T; if they mismatch the payoffs are (0, 1) (player 2 wins), and if they match the payoffs are (1, 0) from (T, T) and (x , 0) from (H, H) (assume $x > 1$). These games have a bizarre property (common to all mixed-equilibrium games): A player's strategy probabilities depend only on the other player's payoffs. As a result, player 1 is predicted to play H and T equally often, ignoring her possible payoffs of x and 1, and player 2 chooses T $x/x + 1$ of the time.

The CH model overturns the weird player-1 prediction and matches the sensible player-2 prediction. One- and two-step player 1's choose H; if τ is 1.5, this means most row players will pick H (cf. the Nash prediction of 50 percent). Since two-step player 2's anticipate that one-step player 1's choose H, they will choose T. When $x = 9$, for example, the CH model predicts 0.89 and 0.28 play of H by players 1 and 2 (using τ estimated from a sample of 22 games), and Nash predicts 0.50 and 0.10. The experimental frequencies are 0.75 and 0.33. Thus, the CH model corrects the strange player-1 prediction and roughly matches the Nash prediction for player 2's (which is empirically accurate). This pattern generally holds across 22 mixed games.

Limited strategic reasoning results from constraints on the human brain; the CH model takes that constraint seriously and embodies it precisely, to out-predict equilibrium. The model also gives good advice (it has "economic value"): Subjects would have earned more if they used the model to forecast.

II. Learning

Recently, game theorists began to actively research the idea that equilibration arises from learning, evolution, or imitation. Two prominent simple learning models are reinforcement and belief learning (e.g., fictitious play). In reinforcement, strategies have numerical attraction levels which are reinforced (increased) when a strategy is chosen and the result is good. Reinforcement is a good model of animal learning but does not gracefully accommodate the

fact that people often choose strategies that have not been directly reinforced. In fictitious play, players form beliefs based on a weighted average of what others have done in the past, and best-respond given their beliefs. Remarkably, weighted fictitious play is equivalent to a generalized reinforcement model in which unchosen strategies are reinforced by the forgone payoffs they *would have* yielded. This algebraic fact means reinforcement and belief learning are boundary cases of a more general model, such as experience-weighted attraction (EWA) (Camerer and Ho, 1999). In EWA, forgone payoffs are weighted by a parameter δ ; $\delta = 0$ is simple reinforcement, and $\delta = 1$ is fictitious play.

In experimental games with limited power to distinguish models (see Timothy Salmon, 2001) these models are all equally accurate and improve modestly over equilibrium. However, EWA is more robust than reinforcement and belief learning because it is easy to find games where the latter two models predict poorly and the δ parameter repairs their weaknesses. For example, when only one player earns a positive payoff, as in a single-unit auction or *p*-beauty contest, simple choice reinforcement predicts almost no learning (since most players are not reinforced). EWA fixes this mistake, because if players learn which of their unchosen strategies would have won and weight its forgone payoff by δ , they will learn much faster than reinforcement predicts (which matches the lab data better). Belief learning often predicts learning that is too rapid or not sharp enough (in price-matching games) (Ho et al., 2002), which is consistent with $\delta < 1$.

Estimates across many games show that the best-fitting values of δ and ϕ (a decay weight) are usually different than the values implicit in fictitious play (which is empirically the weakest of the simple models) and cluster near the reinforcement value ($\delta = 0$) in mixed games. Parameter values also vary systematically across games. One way to model this variation is to replace fixed parameters with functions of experience, allowing both individual differences in "learning styles" and endogenous cross-game differences. The "fEWA" method (Ho et al., 2002) does this by making ϕ into a "change-detector" function which goes toward 1 if other players are equilibrating and toward 0 when

other players' choices are surprising ("starting afresh" after a surprise). Functional values produced this way closely match the best-fitting values of δ and ϕ from a fixed-parameter specification. For example, a fixed ϕ is estimated to be 0.36 for PBC games (because players realize that what happened more than a couple of periods ago is irrelevant), and the functional values average 0.58. Fixed ϕ is close to 1 in mixed games, because a long history is needed to figure out what randomizing players might do, and the functional values average 0.89. Replacing parameters with functions means that the fEWA model has only one free parameter to be estimated (a response sensitivity), which makes it more parametrically parsimonious than many other theories, and often just as accurate or better (even in a new sample of games collected after the initial model was developed).

III. Teaching

In the adaptive-learning models described above, the learning rules do not use information about payoffs of others and do not account for the future effects of current actions. "Sophistication" can be added by assuming that some players realize others are learning (cf. Stahl, 2000). When the same players are paired repeatedly ("partner-matching" in experimental jargon), a sophisticated player can also "teach" by choosing current actions that influence how their learning partner will behave in the future, to the teacher's benefit. (This model borrows strategic foresight from repeated-game theory but assumes that some players learn from the past and lack strategic foresight.) Models of this sort can explain experimental data on finitely repeated trust games (when teaching benefits both players) and entry-deterrence (when teaching benefits only the teacher) (Camerer et al., 2002b) more accurately than stochastic equilibrium theories.

IV. Conclusion

Behavioral game theories limit the amount of rationality and mutual consistency of players, to explain experimental regularities while retaining precision and cross-game generality. This paper described three examples of new theories: a one-parameter model of steps of thinking; a

hybrid model of learning from experience (and a one-parameter version); and a model of sophistication and strategic teaching in repeated games. If the goal is to predict data accurately, models like these should be explored as potential improvements over equilibrium predictions.

REFERENCES

- Binmore, Kenneth.** "Modeling rational players: Part II." *Economics and Philosophy*, April 1988, 4(1), pp. 9–55.
- Camerer, Colin F.** *Behavioral game theory: Experiments on strategic interaction*. Princeton, NJ: Princeton University Press, 2003.
- Camerer, Colin F. and Ho, Teck-Hua.** "Experience-Weighted Attraction Learning in Normal-Form Games." *Econometrica*, July 1999, 67(4), pp. 827–74.
- Camerer, Colin F.; Ho, Teck-Hua and Chong, Juin-Kuan.** "A Cognitive Hierarchy Model of One-shot Games." Unpublished manuscript, Caltech, 2002a.
- . "Sophisticated EWA Learning and Strategic Teaching in Repeated Games." *Journal of Economic Theory*, May 2002b, 104(1), pp. 137–88.
- Capra, C. Mónica.** "Noisy Expectation Formation in One-Shot Games." Ph.D. Dissertation, University of Virginia, 1999.
- Goeree, Jacob K. and Holt, Charles A.** "Ten Little Treasures of Game Theory, and Ten Intuitive Contradictions." *American Economic Review*, December 2001, 91(5), pp. 1402–22.
- Ho, Teck-Hua; Camerer, Colin F. and Chong, Juin-Kuan.** "Functional EWA: A One-Parameter Theory of Learning in Games." Unpublished manuscript, University of California–Berkeley, 2002.
- Nagel, Rosemarie.** "Experimental Results on Interactive Competitive Guessing." *American Economic Review*, December 1995, 85(5), pp. 1313–26.
- Salmon, Timothy C.** "An Evaluation of Econometric Models of Adaptive Learning." *Econometrica*, November 2001, 69(6), pp. 1597–1628.
- Stahl, Dale O.** "Evolution of $smart_n$ Players." *Games and Economic Behavior*, October 1993, 5(4), pp. 604–17.
- . "Rule Learning in Symmetric Normal-Form Games: Theory and Evidence." *Games and Economic Behavior*, July 2000, 32(1), pp. 105–38.