When Does "Economic Man" Dominate Social Behavior?

Colin F. Camerer[1]* and Ernst Fehr[2,3]†

The canonical model in economics considers people to be rational and self-regarding. However, much evidence challenges this view, raising the question of when "Economic Man" dominates the outcome of social interactions, and when bounded rationality or other-regarding preferences dominate. Here we show that strategic incentives are the key to answering this question. A minority of self-regarding individuals can trigger a "noncooperative" aggregate outcome if their behavior generates incentives for the majority of other-regarding individuals to mimic the minority's behavior. Likewise, a minority of other-regarding individuals can generate a "cooperative" aggregate outcome if their behavior generates incentives for a majority of self-regarding people to behave cooperatively. Similarly, in strategic games, aggregate outcomes can be either far from or close to Nash equilibrium if players with high degrees of strategic thinking mimic or erase the effects of others who do very little strategic thinking. Recently developed theories of other-regarding preferences and bounded rationality explain these findings and provide better predictions of actual aggregate behavior than does traditional economic theory.

Most economic analyses are built on two major simplifying assumptions about human nature: Individuals are assumed to be rational decision makers and to have purely self-regarding preferences. The modeling of complex social phenomena often involves simplifying assumptions like these; otherwise, models may quickly become mathematically intractable. The rationality assumption consists of two components: first, individuals are assumed to form, on average, correct beliefs about events in their environment and about other people's behavior; second, given their beliefs, individuals choose those actions that best satisfy their preferences. If individuals exhibit, however, systematically biased beliefs about external events or other people's behavior or if they systematically deviate from the action that best satisfies their preferences, we speak of bounded rationality. Preferences are considered to be self-regarding if an individual does not care per se for the outcomes and behaviors of other individuals. Self-regarding preferences may, therefore, be considered to be amoral preferences because a self-regarding person neither likes nor dislikes others' outcomes or behaviors as long as they do not affect his or her economic well-being. In contrast, people with other-regarding preferences value per se the outcomes or behaviors of other persons either positively or negatively. A large body of evidence accumulated over the last three decades shows that many people violate the rationality and preference assumptions (*1*, *2*) that are routinely made in economics (*3*). Among other things, people frequently do not form rational beliefs, objectively irrelevant contextual details affect their behavior in systematic ways, they prefer to be treated fairly and resist unfair outcomes, and they do not always choose what seems to be in their best interest.

It seems obvious that these violations of the rationality and preference assumptions will appear in the behavior of aggregate entities like markets and organizations or in political processes. This view is premature, however, because many experiments also indicate that a share of the subjects do not violate the above assumptions and, as we will show, the existence of these subjects may cause aggregate outcomes to be close to the predictions of a model that assumes that everyone is rational and self-regarding. The question is therefore how the interactions among heterogeneous subjects shape the aggregate outcome. The intuition into the processes at work can be sharpened by considering how self-regarding individuals and strong reciprocators (*4*) interact in both sequentially and simultaneously played prisoners' dilemma (PD) games. Recent research has documented the existence of a substantial share of strong reciprocators who exhibit a particular form of other-regarding behavior (*5*). Strong reciprocators show a combination of altruistic rewarding, which is a predisposition to reward others for cooperative, norm-abiding behaviors, and altruistic punishment, which is a propensity to impose sanctions on others for norm violations. Strong reciprocators bear the cost of rewarding or punishing even if they gain no individual economic benefit from their acts (*4*).

Cooperation in the Presence of Strong Reciprocators

A PD can be illustrated by a situation in which two geographically separated individuals, A and B, have the chance to engage in a mutually beneficial economic exchange. A and B each possess a good that they value, say, at 10, but each player values the other player's good higher, say at 20. Therefore, if the players send their goods to the exchange partner, they both end up with a more highly valued good than if they retain their goods. There would be no problem if the players could sign a contract that an impartial court could enforce. In the absence of such contract-enforcement institutions, however, the situation represents a PD: A is better off keeping his good, irrespective of whether B sends his good to A. Because the situation is symmetric, B faces the same economic incentives and both players will, therefore, forego the opportunity for a mutually beneficial exchange if they are self-regarding. Strong reciprocators, however, are willing to send their good if they know or believe that the exchange partner will also do so. Thus, the exchange may take place in the presence of strong reciprocators. But what happens if a strong reciprocator (say, player B) faces a self-regarding player A and both players know each other's preferences? If the PD is played simultaneously, i.e., if the goods have to be sent off at the same time, no exchange will take place because B anticipates A's decision to retain the good and does likewise. Thus, the existence of the self-regarding player A induces the strong reciprocator B to behave noncooperatively as well. If the exchange is structured sequentially, however, with A sending off his good first, exchange will take place because A knows that B will only send his good if he first receives A's good. Player A knows, therefore, that if he does not send his good first, no exchange will take place; if, instead he sends his good first, B will reciprocate and both players will be better off. It is, therefore, in A's self-interest to send the good in the sequential exchange; the existence of the strong reciprocator induces the self-regarding player to behave cooperatively in this situation.

The existence of strong reciprocators may generate cooperative outcomes most of the time, even if both players are completely self-regarding but have reason to believe that they face a strong reciprocator with positive probability. Suppose, for example, that there are r = 51% reciprocators in the population and 49% self-regarding players. Suppose further that A and B play the sequential PD, say, 10 times

[1]California Institute of Technology, Pasadena, CA 91125, USA. [2]Institute for Empirical Research in Economics, University of Zurich, Blümlisalpstrasse 10, 8006 Zurich, Switzerland. [3]Collegium Helveticum, Schmelzbergstrasse 25, 8092 Zürich, Switzerland.

*To whom correspondence should be addressed. Division HSS 228-77, Caltech, Pasadena, CA 91125, USA. E-mail: camerer@hss.caltech.edu

†To whom correspondence should be addressed. Institute for Empirical Research in Economics, University of Zurich, Blümlisalpstrasse 10, 8006 Zurich, Switzerland. E-mail: efehr@iew.unizh.ch

and player A does not know whether B is a strong reciprocator. In this situation it is still rational for a self-regarding player A to cooperate (i.e., send the good) even in the final period as long as he believes that he faces a strong reciprocator with more than a 50% chance. Why? If A defects, he knows for sure that the opponent will also do so (i.e., will not send the good) so that A's payoff is 10. But if A cooperates, his expected payoff is $r \times 20 + (1 - r) \times 0$ if $r > 50\%$. A self-regarding player B therefore has a strong incentive to reciprocate A's cooperation in all but the final period because otherwise B would reveal that he is not a reciprocator and this would induce player A to stop his cooperation immediately. Thus, the mere belief that there are reciprocators generates strong cooperation incentives even among purely self-regarding players to gain a reputation by mimicking the behavior of strong reciprocators. In fact, it has been shown that reputation incentives emerging from the belief that the opponent might be a strong reciprocator (tit-for-tat player) may drive cooperation among purely self-regarding players even in simultaneous cooperation games (*6*) and even if r is very small (*7*, *8*).

Theory also shows that a relatively small minority of strong reciprocators can generate cooperative outcomes in one-shot n-person PD games ($n > 2$) if the players are given an explicit punishment opportunity (*9*). Suppose that after the players in the PD have made their choices, they can punish the other players at a cost to themselves. Self-regarding players will never punish in this situation because the game is one-shot and thus there are no future benefits from current investments into punishment. Strong reciprocators will, however, punish defectors even in one-shot situations, providing strong incentives for the self-regarding players to cooperate. Experimental evidence has shown that this threat of punishment may generate very high cooperation rates in stable groups in situations where self-regarding players alone would reach zero cooperation (*10*, *11*) (Fig. 1). However, in the absence of an explicit punishment opportunity, cooperation converges to very low levels (Fig. 1). In fact, theory shows that even a small minority of self-regarding players suffices to induce a large majority of reciprocators to defect in the simultaneous n-person PD (*9*).

All these examples illustrate an important lesson: Individuals who violate the assumptions of economics may create powerful economic incentives for Economic Man to change his behavior, but depending on the economic structure, the existence of Economic Man may also create strong incentives for those with bounded rationality or other-regarding preferences to behave like Economic Man. This principle not only applies to questions of cooperation but is likely to play a role in many other domains, including behavior in bargaining encounters, in competitive markets, as well as in coordination behavior in organizations or in society at large.

The Effects of Competition in the Presence of Strong Reciprocators

To show how the interactions between strong reciprocators and self-regarding individuals shape bargaining behavior, we consider the ultimatum game (*12*), in which a buyer offers a price p to a seller, who can sell an indivisible good. For simplicity, assume that the buyer values the good at 100 and the seller values it at 0. The buyer can make exactly one offer to the seller, which the latter can accept or reject. Trade takes place only if the seller accepts the offer. If the seller is self-regarding, she accepts even a price of 1 because 1 is better than nothing. Thus, a self-regarding buyer will offer $p = 1$ so that the seller earns almost nothing from the trade. Strong reciprocators reject such unfair offers, however, preferring no trade to trading at an unfair price. In fact, a large share of experimental subjects reject low offers in this game, across a wide variety of different cultures (*13*, *14*), even when facing high monetary stakes (*15*, *16*). This fact induces many self-regarding buyers to make relatively fair offers that strong reciprocators will accept. Often the average offers are around $p = 40$, and between 50% and 70% of the buyers propose offers between $p = 40$ and $p = 50$ (Fig. 2A). The behavior of both buyers and sellers changes dramatically, however, if we introduce just a little bit of competition on the seller's side (*13*, *17*, *18*). Assume, for example, that instead of one there are two sellers who both want to sell their good. Again the buyer can make only one offer which, if accepted by one of the sellers, leads to trade. If both sellers reject, no trade takes place; if both sellers accept, one seller is randomly chosen to sell the good at the offered price. Almost all buyers make much lower offers (Fig. 2A) in this situation, and almost all sellers accept much lower offers (Fig. 2B). In fact, if one introduces five competing sellers into this game, prices and rejection rates converge to very low levels such that the trading seller earns only slightly more than 10% of the available gains from trade (Fig. 2, A and B).

Early research on the ultimatum game interpreted the egalitarian outcomes in this game as a sign that people enforce a norm of fairness (*19*). However, if all people obey norms of fairness, why does the price sink to such low levels in the presence of only a little bit of competition among sellers? Heterogeneity in other-regarding preferences is again the key to answering this question, but even if we assume that only a share of the people are strong reciprocators, we still face a puzzle. After all, as Fig. 2, A and B, shows, almost all the experimental subjects make low offers and accept low offers under competitive conditions. Why do fair-minded strong reciprocators, when in the role of a buyer, make such low offers and why do they, when in the role of a seller, accept these low offers? A simple answer to this question would be that competition changes people's preferences; it makes them more selfish. But this would be an uninformative answer; if one can arbitrarily choose the kind of preference that explains an observed behavior, one can explain every behavior and, hence, in fact, nothing. Therefore, the challenge is to explain these facts on the basis of a given distribution of strong reciprocators and selfish subjects (*9*, *17*).

The low rejection rate of reciprocal sellers under competition can be explained if one recognizes their motives. Much research has shown that strongly reciprocal subjects have the goal of punishing unfair behavior or of establishing a fair distribution of outcomes (*2*, *20*). Competition undermines or removes the possibility of meeting these goals in a heterogeneous population of self-regarding and reciprocal sellers. A rational reciprocal seller knows that there is a positive probability that the competing seller(s) will act self-regarding, i.e., will accept any positive offer.

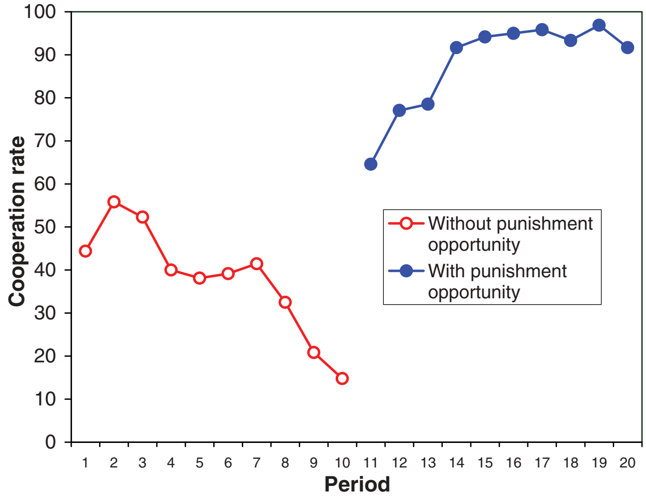

Fig. 1. Cooperation rate in an n-person prisoners' dilemma game with stable groups (*11*). During the first 10 periods, subjects had no opportunity to punish defectors. From period 11 onward, each subject could punish at a cost every other group member after observing their cooperation and defection choices.

Moreover, the more competing sellers there are, the higher the probability that there will be at least one self-regarding seller. If a competing seller accepts a low offer, the reciprocal seller can no longer punish the buyer by rejecting his offer because the buyer can enforce the low price regardless of the reciprocal seller's behavior. Thus, rejections are futile and, therefore, reciprocal sellers will also accept low offers. Here again we encounter a situation where self-regarding agents induce reciprocal agents to behave like self-regarding agents. (*21*)

Recent models of other-regarding preferences explain the phenomena discussed above very well. These models are based on a taste for reciprocation (*22*–*25*), the desire for equitable outcomes (*9*), a distaste for unequal income shares (*26*), or a concern for helping the least well off and the total payoff of the group (*27*). For example, the theory of inequity aversion (*9*), which assumes a share of people with a desire for equitable outcomes, explains why cooperation fails in the absence of a direct punishment opportunity and why it flourishes when the same players have the opportunity to punish group members. The same approach also accounts for the rather egalitarian outcomes in the ultimatum game, while also explaining the low prices in market games with competition (Fig. 2, A and B). In addition, the model predicts when competition does not remove fair behavior. Assume, for example, that the value of the good is not fixed at 100 but is given by 10q, where q measures the quality of the good and is determined by the sellers' effort. If the quality of the good is difficult to enforce through legally binding contracts, the seller has some leeway in determining q, which implies that he can reestablish equity by selling a low-quality good to an unfair buyer. Note that this opportunity to reestablish equity is also available in the presence of competing sellers. Thus, if q is difficult to enforce through contracts, a seller with a preference for equity provides an economic incentive even for self-regarding buyers to treat the seller fairly because otherwise the seller provides low quality. Experimental evidence in fact shows that competition has little impact on prices under these circumstances (*28*) because the buyers' price offers are mainly driven by the concern to ensure high quality. This result could not occur if there were only self-regarding sellers.

Bounded Rationality and Strategic Complementarity

A useful pair of concepts for understanding when aggregate behavior is, or is not, consistent with full economic rationality are "strategic substitutability" and "strategic complementarity" (*29*, *30*). In consumer theory, goods are substitutes if they satisfy similar needs such as, for example, chicken and beef. Therefore, higher chicken consumption will, assuming all else is held constant, be associated with lower beef consumption. Goods are complements if having more of one good enhances demand for another (e.g., peanut better and jelly). The intuition behind substitutes and complements can be extended to strategic contexts. Strategies are complements if agents have an incentive to match the strategies of other players. Strategies are substitutes if agents have an incentive to do the opposite of what the other players are doing. For example, if a firm can earn more profit by matching the prices chosen by other firms, then prices are strategic complements. If firms can earn more profit by choosing a low price when other firms choose high prices (and vice versa), then prices are strategic substitutes.

The idea of strategic substitution and complementarity was first developed in studies of firm interactions (*29*, *30*) but extends naturally to the interaction of economic agents with limited and unlimited rationality (*31*, *32*). When economic choices are substitutes, then rational agents have an incentive to behave in the opposite way to that of less-rational agents. Therefore, the rational agents' behavior will counteract the impact of less-rational agents on aggregate behavior. However, when choices are complements, then it pays for rational agents to mimic the behavior of the less-rational agents. Therefore, the rational agents' behavior amplifies the impact of less-rational agents on aggregate behavior. There is, in fact, evidence indicating that under strategic substitutability, a minority of rational individuals may suffice to generate aggregate outcomes that are predicted by a fully rational model (*33*) whereas under strategic complementarity, a small minority of irrational individuals may cause outcomes that are completely at odds with the rational model (*34*).

To see the amplifying influence of bounded rationality, consider a simple game in which many players choose numbers from 0 to 100 at the same time. The average number is calculated and multiplied by 2/3. The player whose

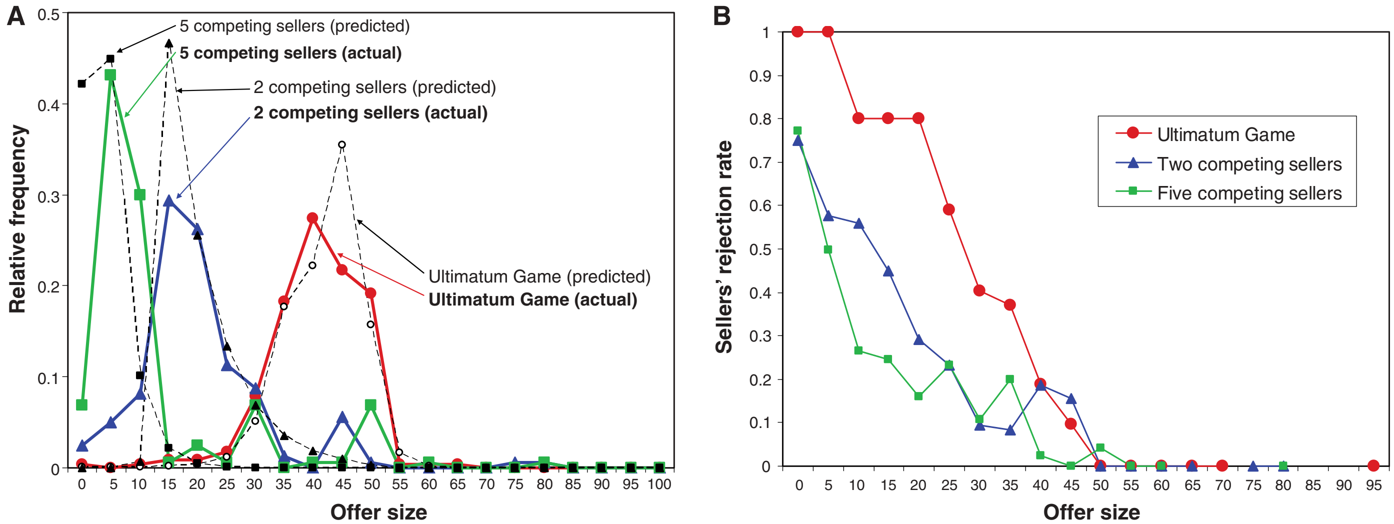

Fig. 2. Behavior of buyers and sellers in the ultimatum game and in market games with competing sellers (*17*). In all games the buyer can make a price offer between 0 and 100 for an indivisible good with value 100. (**A**) The distribution of accepted price offers across conditions. In the ultimatum game most prices are between 40 and 50. If there are two competing sellers, most prices are between 10 and 25; in the case of five competing sellers, the large majority of prices is between 5 and 10. The dotted lines show the predictions of a fairness model for each of the three conditions (*17*). The model has one free parameter to fit the data and combines the theory of inequity aversion (*9*), which assumes heterogeneous preferences for equitable outcomes, with stochastic best reply behavior (*50*). (**B**) Sellers' rejection rate across conditions. More competition leads sellers to reject low offers less frequently.

number is closest to 2/3 of the average (in absolute value) wins a fixed prize. Suppose the numbers are interpreted as the time at which economic actions are taken. Then the game is a simple model of economic situations like introducing a new product in a growing market, or selling stocks in a rising bull market, because players want to move earlier than other players (i.e., choose lower numbers), but the optimal time to move depends on when the average player moves. This game is often called a ''beauty contest,'' after a passage in John Maynard Keynes's influential economics book (*35*). Keynes describes the stock market as a beauty contest in which investors try to figure out what stocks other investors find attractive. Spotting the stocks, which other investors will soon find attractive, earlier enables savvy investors to buy low and sell at a higher price, when the attractiveness of the stocks becomes obvious to all investors and prices rise.

The beauty contest game has been played with dozens of groups of subjects, including small groups of students, highly trained subjects (professional game theorists), and large newspaper contests in which thousands of readers mailed in number entries (*36–38*). The basic patterns of numbers are similar across many groups. Figure 3 shows the data from one study (using a multiplier of 0.7 rather than 2/3) with Singapore engineering students (*37*). Number choices are widely distributed. The most common choice was 35, a good choice if you believe choices of others are random (so the expected average will be 50). Some other subjects choose numbers from 20 to 30, as if they anticipate how other subjects are responding to perceived choices that are random.

The equilibrium concept that is most widely used in game theory offers a clear analysis about this game. In a game-theoretic (Nash) equilibrium, every player guesses accurately what others will do and chooses a ''best response'' strategy, which will give the highest expected payoff, given the guesses. The unique equilibrium in the 2/3 game is to choose zero. Intuitively, if players think the average will be a number X, they should choose $(2/3)X$. But if they believe others guess accurately, then other players will choose a best response to $(2/3)X$, which is $(4/9)X$. If all players are reasoning accurately about the reasoning of other players, they should then choose $(2/3)(4/9)X$, and so on. Imposing the restriction that all players guess correctly what other players will do leads to an equilibrium in which choices must equal beliefs, or $X = (2/3)X$, which implies that $X = 0$. Intuitively, if some players choose numbers N above 0, they should anticipate that other players will choose $(2/3)N$ and should lower their choices; the only combination of optimal response and accurate belief is when all players choose 0. Although this reasoning is logically persuasive, it leads to a bad prediction about what will happen, and also gives bad advice. The Nash equilibrium is an inaccurate prediction because strategies are complements: If a player thinks others will pick high numbers, that player should choose a high number too, which means that if limitedly rational players choose numbers that are above the equilibrium of 0, then even rational players should deviate from the equilibrium by choosing high numbers as well.

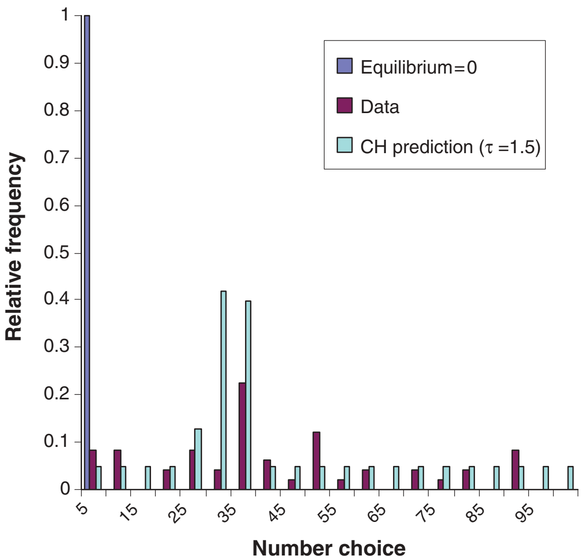

Fig. 3. Number choices and theoretical predictions in beauty contest games. In the beauty contest game players choose numbers from 0 to 100 (x axis, bins of five numbers except 0 to 5). The closest number to 0.7 times the average wins a fixed prize. The relative frequency of number choices is shown on the y axis (*37*). There are large numbers of choices at 50 and 35. The equilibrium prediction is 0. The CH model (*38*) with $\tau = 1.5$ predicts a spread of choices across the 0 to 100 range, and frequent choices of 35, 29, 26, and 25 (resulting from one to four steps of thinking). The actual mean is 39.9. The CH predicted mean is 34.9.

A business entry game illustrates the opposite pattern, in which limits on rationality have diminished impact when strategies are substitutes. Consider a business entry game involving 12 firms. Firms can stay out of a new market and earn a payoff of 0.5, or can enter a competitive market with a capacity c, where c is the number of firms that can coexist profitably, and c is an even number (2, 4, …10). If c of the firms enter, or fewer, then all firms who enter earn a payoff of 1. If more than c firms enter, then all the entering firms earn 0. A smart firm that is neutral toward risk will enter if it believes the chance that there will be $c - 1$ entrants or fewer is less than 50%. In equilibrium, exactly c firms enter when the capacity is c. If more than c entered, then some firms made a forecasting mistake and should have stayed out; if fewer than c entered, some firms that did not enter should have entered. Notice that near the equilibrium entry choices are strategic substitutes—if firms think too many firms will enter, they prefer to stay out; and if they think too many firms will stay out, they should enter.

In experiments, approximate equilibration occurs across different values of c, even in single-shot game experiments when subjects must choose at the same time without communicating beforehand (*2*). This empirical result is surprising because all firms would prefer to enter and earn 1 than to stay out, so the firms must somehow collectively resolve the problem of coordinating which of the firms will enter (earning the highest payoff) and which will stay out, for different values of c. But approximate equilibration occurs instantly, without negotiation, communication, or learning over time. As Kahneman wrote, ''To a psychologist, it looks like magic'' (*39*).

In the beauty contest game, behavior is reliably far from the equilibrium of 0. In the business entry game, behavior is surprisingly close to the equilibrium of c entries, even without learning or communication. The fact that numbers are strategic complements in the beauty contest game, and entry choices are strategic substitutes in the entry game, provides an important clue to explaining why the beauty contest results are far from equilibrium and the entry game results are close to equilibrium. Can a unified theory of bounded rationality explain the opposite results in the two games, reflecting the different impacts of strategic substitutability and complementarity? The answer is yes.

One class of theories that can explain the conflicting results in both games is a ''cognitive hierarchy'' (CH) approach (*38*, *40*, *41*). In theories of this type, there is a distribution of the number of steps of iterated strategic reasoning that players can do. The fraction of players who do k steps of thinking is $f(k)$. Zero-step players just randomize across their strategies. Higher-step players think they are playing against players who do fewer steps of thinking than they do. The model can be closed mathematically by assuming a distribution $f(k)$, with a precise specification of the beliefs of k-step players about the distribution of players who do less reasoning than they do. A reasonable specification of beliefs is that k-step players believe, overconfidently, that they are responding to players who do 0 to $k - 1$ steps of thinking. A simple specification of $f(k)$ that fits data from many different games is a Poisson distribution, which is fully characterized by a single parameter τ, the average number of steps of thinking (*38*). When τ is 1.5 (an estimate that fits many games well), the Poisson $f(k)$ drops off very rapidly; only 8% of players do more than three steps of thinking. This low percentage reflects the intuition that because doing many steps of thinking is mentally difficult, and is constrained by working memory, three or more thinking steps are rare without special training or practice. Not surprisingly, experiments across different subject pools also show differences in the average number of thinking steps τ across groups (*38*).

In the beauty contest game with a multiplier of 0.7, the CH model (with τ = 1.5) generates a distribution of numbers across the entire range, a spike of one-step choices at 35, and two-step choices at 29 (Fig. 3). This simple model fits the basic features of the data more accurately than the equilibrium of zero (which was chosen by only 2% of the subjects). The model can sometimes be improved further by including other thinking types, such as players who choose very low numbers because they think, usually mistakenly, that many others will do so as well.

The same CH model that can explain limited progress toward the equilibrium of 0 in the beauty contest game can also predict why players converge close to the equilibrium instantly in the entry game. In the entry game with 12 players, zero-step players ignore the value of *c* and enter with probability 0.5 for every value of *c*. One-step players stay out when *c* is 2, 4, or 6 (because they think there will be too many zero-step entrants) and enter when *c* is 8 or 10. Two-step players have a more nuanced strategy, responding to their beliefs about the combination of entry by 0-step and 1-step players. When τ = 1.5, two-step players stay out when *c* = 2, because too many 0-step players enter, but they enter when *c* = 4 or 6, because the one-step players stay out for those values of *c* and make it optimal to enter. They stay out when *c* = 8, because entry by one-step players crowds the market, but they enter when *c* = 10. Including each higher *k*-step type smoothes out the deviation between the perceived rate of entry of the average of lower-step thinkers, and the equilibrium rate of entry even further, because of strategic substitution (players stay out when they think too many players will enter, and enter when they think too few will enter). The result is an aggregate entry function, averaging across players using different numbers of steps of thinking, which predicts entry that rises monotonically in the capacity *c*, but also predicts too much entry at low *c* and too little entry at high *c* (Fig. 4). This simple model is one explanation for the "magic" of approximate equilibrium entry rates without learning or communication in experiments.

The beauty contest and entry games show how the same unified model of bounded rationality, made precise in the cognitive hierarchy approach, can explain when behavior is far from equilibrium, in the beauty contest game, and when behavior is surprisingly close to equilibrium, in the entry game. In the beauty contest game, strategies are complements, so players who do limited thinking cause even rational players to choose high numbers. In the entry game, choices are substitutes so the influence of boundedly rational players on aggregate behavior is largely erased.

Strategic substitutability and complementarity can also play an important role in financial markets, where these forces can amplify or diminish the impact of limits to rationality on aggregate outcomes. This can be illustrated by prediction markets for bets on events, and stock markets. In prediction markets, an upcoming event is defined precisely so that bets can be settled, such as the future price at which commodities will sell in a few months, the outcome of a political election, or a newsworthy happening like the capture of Osama Bin Laden. In prediction markets for events, assets are created that pay a fixed sum if the event occurs. Traders buy and sell the asset. The price that is established, normalized by the payout, gives a market-wide probabilistic estimate that the event will occur. For example, the Tradesports Web site (www.tradesports.com) prices on 13 October 2005 implied a 19% probability that Osama Bin Laden would be captured by June 2006.

Many studies have found that prices in prediction markets are remarkably accurate forecasts of events (*42*). Orange juice prices are very sensitive to cold weather in Florida, which causes freezes and drives up prices by reducing supply. Futures prices for juice are therefore influenced by forecasts of freezes. However, one study showed that futures prices for juice are more accurate forecasts of the chance of a freeze than U.S. National Weather Service meteorological forecasts (*43*). The Iowa Political Stock Markets (www.biz.uiowa.edu/iem) also forecast actual election results more accurately than expensive opinion polls in more than 75% of hundreds of different elections at many levels and in different countries (*44*). Sixty days before presidential elections, the Iowa market absolute forecast error of vote share is only 2% (*28*, *45*).

Prediction markets forecast accurately because poorly informed traders provide a clear opportunity for better-informed traders to make money. Better-informed traders who express their confidence by making large trades can be sure to collect when an event either does or does not occur, at a known time in the near future. For example, if a better-informed trader knows that the asset is undervalued (i.e., the event is more likely to occur relative to the prevailing market opinion), he will buy the asset from the poorly informed traders. Thus, substitutability again diminishes the impact of less rational actors.

A contrasting case is stock markets. Stock prices respond to new information rapidly, and using public information (like past price trends) to beat the market is difficult. At the same time, because stocks are claims on profits of an ongoing enterprise, there is never a fixed future time at which the true value of a firm is established and bets are settled once and for all. As a result, well-informed traders cannot always guarantee a profit at the expense of traders with limited rationality. In fact, institutional constraints such as performance pressure, and impediments to selling shares short (betting that stock prices will fall), mean that if stock prices are bad estimates of the value of a firm, large well-capitalized investors cannot always guarantee a profit by betting against the market (*46*). The fact that "noise traders" add volatility to stock prices creates a special kind of risk for smart investors, which makes them reluctant to bet against noise traders and bring stock prices closer to the fundamental value of firms (*47*). As a result, when there are institutional constraints trading strategies can be strategic complements, so that well-informed investors can be forced to follow a poorly informed crowd, rather than betting against it. A striking example is the mispricing of "twin shares," such as Royal Dutch/Shell. Royal Dutch/Shell is a single company whose economic value, until very recently, was divided into two separate shares, Royal Dutch

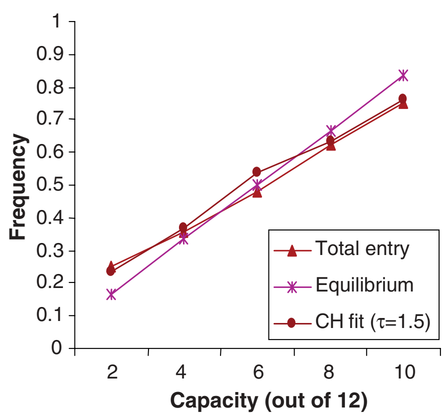

Fig. 4. Rates of entry in business entry games and theoretical predictions (in percentage terms). Actual and predicted rates of entry in games with N = 12 firms and capacity c equal to 2, 4, 6, 8, and 10. Identity line (*) shows Nash equilibrium entry rates (entry equals capacity). Actual entry in experiments [▲, (*38*)]. Entry predicted by CH model (•). Actual entry is monotonic in capacity c, but there is too much entry for low c and too little entry for high c. Actual entry rates (▲) are close to the Nash equilibrium (*). The entry rates predicted by CH (•) are consistent with the deviations between actual and equilibrium entry rates.

shares (traded in Amsterdam) and Shell shares (traded in New York). Based on a merger agreement, the Royal Dutch shareholders are legally entitled to 60% of the combined entity's cash flows and the Shell company is entitled to 40%. If the prices of the shares reflect their economic value, then the ratio of Royal Dutch and Shell share prices should always be 60/40, or 1.5. In fact, the ratio of the two stock prices has wandered away from 1.5 for many years, from 30% too low in 1981 to 15% too high in 1996 (*48*). If the ratio is too high, investors can potentially profit by selling Royal Dutch shares short and buying Shell shares and waiting for the ratio to fall toward 1.5. But betting that the price ratio will revert to its economic value of 1.5 is inherently risky because markets are volatile. When the ratio is well above 1.5, it often rises even further away. In fact, hedge funds that made highly leveraged bets that the ratio would return to 1.5 are exposed to risk. If investors with short horizons are nervous about betting heavily on reversion to the 1.5 ratio, their nervousness keeps them from making large bets on rapid reversion, which in turn keeps that ratio from rapidly reverting to 1.5, which validates their nervousness.

The contrast between prediction and stock markets reiterates the basic theme of this review. In prediction markets there is a known future time at which bets will be settled based on event. As a result, trades are strategic substitutes because a well-informed trader can profitably bet against a poorly informed one with little risk. But stock values are never decided at a clear point in the future. So prices can drift far from economic fundamentals for many years, as the Royal Dutch/Shell case shows. Rational traders who recognize the mispricing and bet against it might have to wait years to earn their due. As Keynes wrote, "markets can stay irrational longer than you can stay liquid." So, trading strategies are complementary when rational traders have an economic incentive to go along with the crowd for extended periods of time.

Alternative Models and Future Directions

The examples discussed in this review show that heterogeneity in other-regarding preferences and bounded rationality, along with the structure of social interactions, determine when collective outcomes are close to predictions based on rationality and self-regarding preferences, or are far from those predictions. Under certain conditions, models based on self-regarding preferences and homogeneous rationality predict aggregate behavior rather well, even though many people exhibit rationality limits and other-regarding preferences (*49*). However, under strategic complementarity, even a small proportion of other-regarding or boundedly rational players may suffice to generate collective outcomes that deviate sharply from models of Economic Man. The new models of heterogeneous social preferences and bounded rationality explain these puzzling results in a unifying way because they explicitly take heterogeneity and incentive interactions between different types of individuals into account. Therefore, they can explain when Economic Man dominates aggregate outcomes and when he fails to do so.

There are many other social domains in which the mixture of heterogeneous social preferences and rationality limits are likely to create profound effects on aggregate behavior. In companies, matching different workers to appropriate jobs, based on their preferences and rationality, implies interesting variation in the nature of employment contracts and firm-level outcomes. Designing well-functioning economic institutions, to help poor countries grow richer, depends on a good model of human behavior. Governments, philosophers, and lawyers are concerned about crafting policies that protect consumers with rationality limits that are swamped by information and choices, while protecting the freedom of choice of expert consumers. Understanding the biological basis for differences in preferences and rationality bounds, and locating their neural circuitry, will also help social sciences, and will inform neuroscience about important kinds of higher-order cognition. A better understanding of when the useful caricature of Economic Man dominates markets, or is dominated by social preferences and rationality limits, will inform all these enterprises and could lead to a more unified, and powerful, approach to both biological and social sciences of human behavior.

References and Notes

1. D. Kahneman, P. Slovic, A. Tversky, *Judgment Under Uncertainty—Heuristics and Biases* (Cambridge Univ. Press, New York, 1982).
2. C. F. Camerer, *Behavioral Game Theory—Experiments in Strategic Interaction* (Princeton Univ. Press, Princeton, NJ, 2003).
3. In principle, economic methods such as optimization under constraints are strictly neutral with regard to the nature of people's preferences. These methods certainly allow for the formalization of other-regarding motives, but in almost all applications, economists assume that preferences are purely self-regarding.
4. H. Gintis, *J. Theor. Biol.* **206**, 169 (2000).
5. E. Fehr, U. Fischbacher, *Nature* **425**, 785 (2003).
6. D. Fudenberg, E. Maskin, *Econometrica* **54**, 533 (1986).
7. D. Kreps, R. Wilson, P. Milgrom, J. Roberts, *J. Econ. Theory* **27**, 245 (1982).
8. C. Camerer, K. Weigelt, *Econometrica* **56**, 1 (1988).
9. E. Fehr, K. M. Schmidt, *Q. J. Econ.* **114**, 817 (1999).
10. E. Ostrom, J. Walker, R. Gardner, *Am. Polit. Sci. Rev.* **86**, 404 (1992).
11. E. Fehr, S. Gächter, *Am. Econ. Rev.* **90**, 980 (2000).
12. W. Güth, R. Schmittberger, B. Schwarze, *J. Econ. Behav. Organ.* **3**, 367 (1982).
13. A. Roth, V. Prasnikar, M. Okuno-Fujiwara, S. Zamir, *Am. Econ. Rev.* **81**, 1068 (1991).
14. J. Henrich *et al.*, *Am. Econ. Rev.* **91**, 73 (2001).
15. R. Slonim, A. E. Roth, *Econometrica* **66**, 569 (1998).
16. L. A. Cameron, *Econ. Inq.* **37**, 47 (1999).
17. U. Fischbacher, C. Fong, E. Fehr, *Working Paper No. 133* (Institute for Empirical Research in Economics, University of Zurich, 2002).
18. B. Grosskopf, *Exp. Econ.* **6**, 141 (2003).
19. W. Güth, R. Schmittberger, R. Tietz, *J. Econ. Psychol.* **11**, 417 (1990).
20. E. Fehr, K. Schmidt, in *Advances in Economics and Econometrics*, M. Dewatripont, L. P. Hansen, S. J. Turnovky, Eds. (Cambridge Univ. Press, Cambridge, UK, 2003), vol. 1, p. 208.
21. Explaining why even reciprocal buyers make unfair offers is less accessible to simple intuitions. See (*9*) and (*17*) for mathematical explanations.
22. M. Rabin, *Am. Econ. Rev.* **83**, 1281 (1993).
23. D. K. Levine, *Rev. Econ. Dyn.* **1**, 593 (1998).
24. M. Dufwenberg, G. Kirchsteiger, *Games Econ. Behav.* **47**, 268 (2004).
25. A. Falk, U. Fischbacher, *Games Econ. Behav.,* in press.
26. G. E. Bolton, A. Ockenfels, *Am. Econ. Rev.* **90**, 166 (2000).
27. G. Charness, M. Rabin, *Q. J. Econ.* **117**, 817 (2002).
28. E. Fehr, A. Falk, *J. Polit. Econ.* **107**, 106 (1999).
29. D. Fudenberg, J. Tirole, *Am. Econ. Rev.* **74 2**, 361 (1984).
30. J. I. Bulow, J. D. Geanakoplos, P. D. Klemperer, *J. Polit. Econ.* **93 3**, 488 (1985).
31. G. A. Akerlof, J. L. Yellen, *Am. Econ. Rev.* **75 4**, 708 (1985).
32. J. C. Haltiwanger, M. Waldman, *Am. Econ. Rev.* **75**, 326 (1985).
33. B. D. Kluger, S. B. Wyatt, *J. Finance* **59**, 969 (2003).
34. E. Fehr, J. R. Tyran, *Am. Econ. Rev.* **91**, 1239 (2001).
35. J. M. Keynes, *The General Theory of Employment, Interest and Money* (Macmillan, London, 1936).
36. R. Nagel, *Am. Econ. Rev.* **85**, 1313 (1995).
37. T. Ho, C. F. Camerer, K. Weigelt, *Am. Econ. Rev.* **88**, 947 (1998).
38. C. F. Camerer, T.-H. Ho, J.-K. Chong, *Q. J. Econ.* **119**, 861 (2004).
39. D. Kahneman, in *Bounded Rational Behavior in Experimental Games and Markets*, R. Tietz, A. Wulf, R. Selten, Eds. (Springer, Berlin Germany, 1988).
40. D. O. Stahl, P. Wilson, *Games Econ. Behav.* **10**, 213 (1995).
41. M. Costa-Gomes, V. Crawford, B. Broseta, *Econometrica* **69**, 1193 (2001).
42. J. Wolfers, E. Zitzewitz, *J. Econ. Perspect.* **18**, 107 (2004).
43. R. Roll, *Am. Econ. Rev.* **74**, 861 (1984).
44. J. Berg, R. Forsythe, F. Nelson, T. Rietz, in *Handbook of Experimental Economic Results*, C. R. Plott, V. L. Smith, Eds. (Elsevier, Amsterdam, in press).
45. Although remarkably accurate in many cases, prediction markets are also sensitive to short-term mistakes: For example, based on early "exit polls," in the 2004 U.S. presidential election the Tradesports and Iowa Political Markets both forecasted a 70% chance of a victory by John Kerry (even though both markets had forecasted a narrow Bush victory the day before the election), because the exit polls were inaccurate.
46. A. Shleifer, *Inefficient Markets—An Introduction to Behavioral Finance* (Oxford Univ. Press, Oxford, 2000).
47. J. B. DeLong, A. Shleifer, L. H. Summers, R. J. Waldman, *J. Polit. Econ.* **98**, 703 (1990).
48. O. A. Lamont, R. H. Thaler, *J. Econ. Perspect.* **17**, 191 (2003).
49. C. R. Plott, *J. Business* **59**, S301 (1986).
50. R. D. McKelvey, T. R. Palfrey, *Games Econ. Behav.* **10**, 6 (1995).

51. C.F.C. gratefully acknowledges support from the Human Frontier Science Program and NSF (grant SES 000376) grants to Caltech. E.F. gratefully acknowledges support from the Research Priority Program of the University of Zurich on the "Foundations of Human Social Behavior" and the Collegium Helveticum.

10.1126/science.1110600