5. You Can Fool Some of the People…

> You may fool all the people some of the time, you can even fool some of the people all of the time, but you cannot fool all of the people all the time. *Abraham Lincoln.*

What can economic theory reasonably hope to say? Any model is an idealization in which many things that are thought to be relatively unimportant are ignored: decision costs, social preferences, costs of acquiring information, and so forth and so on. Moreover in applied work it is necessary to adopt specific mathematical functions which are at best approximations to an underlying reality. A caricature of *homo economicus* asserts that in the laboratory everyone is selfish and that all the participants understood the instructions. Or more strongly that all students always get all exam questions correct – the falsity of which even academic economists must surely be aware.

Modern economic theory is not such a caricature. As we have seen Nash equilibrium sometimes predicts well – and sometimes does not. Whether a theory that is sometimes right and sometimes wrong is useful depends on whether we can predict when it will be correct. For example, Newtonian mechanics does poorly at speeds close to that of light, but is very useful at lower speeds. It is true that Nash equilibrium is a core concept in modern economic theory. It is, however, the starting point of economic theory, not the ending point – economists have developed a set of tools that enables us to determine when Nash equilibrium is a reasonable approximation and when not.

I have discussed the theory of social preferences and will subsequently discuss learning theory. Besides these specific models, economists have theories that enable us to understand what happens when everyone is a little "irrational" and a few people are very "irrational."

Approximate Equilibrium

In standard Nash equilibrium it is assumed that every player makes the best choice possible. In 1980 Roy Radner introduced the weaker concept of approximate Nash equilibrium: this supposes only that each player makes a relatively good choice. In a correct model choosing the best option given beliefs is essentially a tautology. Given that models are never correct there is no reason to presume that theoretical players do better than "relatively well."

The idea that players do "well" but not "perfectly" can be found in some of the earliest behavioral criticisms of standard economics. Simon's 1956 notion of *satisficing* behavior – for which he won the Nobel Prize in Economics – supposes that people are satisfied and stop attempting to learn if they achieve a desirable goal that falls short of the very best possible. In Simon's theory this goal is based on historical data about how well the decision-maker has done in the past.

Although it is not widely known modern economics incorporates satisficing concepts in two ways. The first is through the notion of habit formation where preferences change over time as experience is acquired. More on that later. The second is through the notion of approximate optimization.

The idea of approximate optimization is hardly new and scarcely originates either with Simon or Radner. The traditional theory of competitive behavior is a model of approximate optimization. That is, in practice and in any economic model, a trader always has a little bit of market power – even the smallest wheat trader can change prices a tiny bit in her favor by withholding a some wheat from the market. But in practice nobody is going to take the time and effort to figure out how to manipulate a market in order to garner a few cents. The theory of competitive behavior supposes that traders ignore the possibility of such small gains.

The use of approximate optimization is also widespread in the modern economic theory of learning. To take two examples: in Foster and Young’s 2003 paradigm of the hats it is assumed that a player only try new things if there is evidence of a strategy that works at least a bit better than the status quo. In Fudenberg and Levine [1995] players are assumed to randomize between nearly indifferent alternatives even though this results in slightly less than the optimum payoff. This randomization provides strong protection against an opponent who is cleverer than you are.

The notion of approximate equilibrium is also important for measurement. Given the objective play of other players, and what a player actually did, we can ask “how much more money could that player have earned?” In Nash equilibrium the answer is zero – it is not possible to do better. In approximate Nash equilibrium the answer may be positive – and is often referred to by the greek letter ε, which in mathematics is traditionally used to refer to a small number. Notice that modifying Nash equilibrium to allow an ε loss contains two possibilities. One may be that a player consistently earns a bit less than she might. The other is that she occasionally earns a lot less than she might. That is “all of the people some of the time" and "some of the people all of the time." The latter possibility – people occasionally earning a lot less than they might is of particular importance when the population is large, since it implies that a small fraction of the population will be "misbehaving" quite a lot.

Turning back to measurement, ε is our measure of how much the "true" preferences of the player differ from the preferences that we have written down. So we allow the possibility that the true "payoff" from a choice might be somewhat larger than captured by the model, but by no more than ε. In effect ε is a measure of the approximation we think we made when we wrote down a formal mathematical model of player play, or of the uncertainty we have about the accuracy of that model.

The measure of "success" for Nash equilibrium should not be whether play "looks like an equilibrium" but whether ε is small. Take the case of ultimatum bargaining. Fudenberg and Levine [1997] computed the losses to players playing less than a best-response as averaging $0.99 per player per game out of the $10.00 at stake. What is especially striking is that most of the money is not lost by second players to whom we have falsely imputed selfish preferences, but rather by first movers who incorrectly calculate the chances of having their offers rejected. As we have noted, however, a first player who offers a 50-50 split may not realize that she could ask for and get a little bit more without being rejected, nor if she continues to offer a 50-50 split, will she learn of her mistake.

The message here is not that the theory worked well, but rather that the failure of the theory is much less than a superficial inspection suggests. Simply comparing the prediction of subgame perfection to the data indicates that players offered $5.00 when they should have offered $0.05. Yet a more reasonable measure of the success of the theory is that players lose only $0.99 out of the possible $10.00 that they can earn.

Equilibrium: The Weak versus the Strong

The problem with approximate (or ε-) equilibrium is not that it makes inaccurate predictions, but that it makes too many predictions. The ultimatum bargaining game is a perfect example: with $\varepsilon = \$0.99$ half of the offers at $5.00 is an approximate equilibrium – and so is all the offers at $0.05.

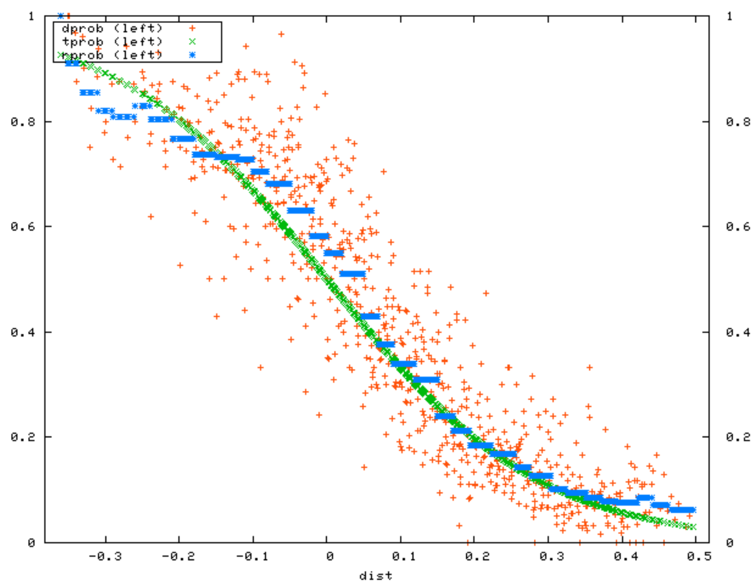
Weak predictions are not a good thing in a theory. Yet a theory that is sometimes weak and sometimes strong can be useful if it lets us know when it is weak and when it is strong. When there is a narrow range of predictions – as in the voting game, or in games such as best shot or competitive bidding – the theory is useful and correct. When there is a broad range of predictions such as in ultimatum bargaining the theory is correct, but not as useful.

The role for behavioral economics – if there is to be one – is not to overturn existing theory, but to strengthen it. The evidence is strong that psychological factors are weak compared to economic factors, but in certain types of games that may make a great deal of difference.

Voting Redux

To get a sense of the limitations of existing theory, it is useful to take a look under the hood of the voting game described earlier. At the aggregate level the model predicts with a high degree of accuracy. However, as anyone who has ever looked at raw experimental data can verify, individual play is very noisy.

The figure below from Palfrey and Levine [2007] summarizes the play of individuals in the voting experiment. Depending on the probability of being pivotal (deciding the election) and on the cost of participation, we can calculate for each player how costly it is to participate. This is shown on the horizontal axis. If – in a given election – the cost is positive the player should not vote; if it is negative then the player should vote. The vertical axis is the actual frequency with which voters participated. The red dots are the results of individual elections. The blue dots are averages of the



red dots for each level of participation cost, and the green curve is a theoretical construct described below.

The theory of Nash equilibrium says that we should observe a "best response" function that is flat with the probability of participating equal to one for all negative losses (gains) and flat with a probability of zero for all positive losses. This is far from the case: some players make positive errors, some make negative errors. However in this voting game the errors tend to offset each other. Over voting by one voter causes other voters to want to under vote, so aggregate behavior is not much affected by the fact that individuals are not behaving exactly as the theory predicts. A similar statement can be made about the competitive auction and other games in which equilibrium is strong and robust. By way of contrast in ultimatum bargaining a few players rejecting bad offers changes the incentives of those making offers: They will wish to make lower offers – moving away from the subgame perfect equilibrium, not towards it.

A key feature of the individual level data is that behavior is sensitive to the cost of "mistakes." That is, voters are more likely to play "sub-optimally" if the cost of doing so is low. The same is true in ultimatum bargaining: bad offers are less costly to reject than good ones, and are of course rejected more frequently.

Quantal Response Equilibrium

One response to the fact that in some games such as ultimatum bargaining equilibrium theory makes weak predictions is to try to explicitly model psychological forces to get a more accurate model that can make more exact predictions. A more naïve approach is to ignore psychological forces entirely and just assume that costly deviations from equilibrium are less likely than inexpensive ones. This captures the important fact that when incentive are weak play is less predictable. It leads to a theory known as *quantal response equilibrium* (or QRE) introduced by McKelvey and Palfrey in 1995. It is built on the standard logistic choice model introduced to economics by McFadden [1980].

QRE supposes that play is somewhat random. It assumes a non-negative numerical parameter usually represented by the greek letter λ. This parameter describes how noisy choices are. If $\lambda = 0$ the player simply chooses a strategy at random. As λ grows large her play approaches the best response of Nash equilibrium. For intermediate values of λ strategies with higher payoffs are more likely to be used than those with lower payoffs, but there is still a chance that lower valued alternatives will be chosen.

In a Nash equilibrium players must play optimally given their beliefs and their beliefs must be correct. In a QRE players must employ probabilities consistent with λ given their beliefs and their beliefs must be correct. Rather than a best response they play a "quantal response."

To give an idea how this theory works in the voting experiment we can estimate a common value of λ for all players. The corresponding equilibrium probabilities of play are given by the green curve in the figure above. This does an excellent job of describing individual play – although it makes roughly the same predictions for aggregate play as Nash equilibrium.

While QRE is useful in explaining a many experimental deviations from Nash equilibrium in games where Nash equilibrium is weak, it captures only the cost side of preferences. That is, it recognizes – correctly – that departures from standard "fully rational" selfish play are more likely if they are less costly in objective terms, but it does not attempt to capture the benefits of playing non-selfishly. It does not well capture, for example, the fact that under some circumstances players are altruistic, and in others spiteful.

Selling a Jar of Pennies

Enough theory – would you like to make some money? Here is a surefire way to do it. Put a bunch of pennies in a jar, and get together a group of friends. Then auction off the jar of pennies. You

will find if you have about thirty friends that you can sell a $3.00 jar of pennies for about $10.00.

This illustrates an important phenomenon known at the *winner's curse*. Your friends all stare at the jar and try to guess how many pennies there are. Some under guess – they may guess that there are only 100 or 200 pennies. They bid low. Others over guess – they may guess that there are 1,000 pennies or more. They bid high. Of course those who overestimate the number of pennies by the most bid the highest – so you make out like a bandit.

According to Nash equilibrium this shouldn't happen. Everyone should rationally realize that they will only win if they guess high, so they should bid less than their estimate of how many pennies there are in the jar. They should bid a lot less – every player can guarantee they lose nothing by bidding nothing. So in equilibrium, they can't on average lose anything, let alone $7.00.

QRE – by recognizing that there is a small probability that people aren't so rational – makes quite a different prediction. People no doubt perceive that there is some most possible profit they could make by getting the most number of pennies at zero cost. Let's call this amount of utility U. They also perceive that there is some least possible profit by getting a jar with no pennies at the highest possible bid. Let's call that utility u. As a formal mathematical theory, QRE says that the ratio of probabilities between two different strategies is a function of λ times the difference in utilities – specifically that the ratio of the probability between two bids that give utility U,u is $\exp[\lambda(U-u)]$ where $\exp$ stands for the exponential function of mathematics. Now whatever is the difference in utility between two strategies it cannot be greater than that between U and u. What this means is that the probability of the highest possible bid is always at least some number $p>0$ that may depend on how many bids are possible, but not on how many bidders there are or what strategies they employ.

What happens as the number of bidders grows? Each bidder according to QRE has at least a p probability of making the highest possible bid. With many bidders it becomes a virtual certainty that one of the bidders will (unluckily for them) make this high bid, so with enough bidders, QRE assures the seller a nice profit.

Break Left? Or Right?

The role of approximate equilibrium, of QRE, and of altruism can be seen in analyzing the game of *Matching Pennies.* Each player has a penny, and secretly places it heads up or heads down. If the two pennies match – both heads or both tails – one player, the matching player, wins both pennies; if the two pennies do not match her opponent wins both pennies.

Matching pennies is an example of a zero sum game: one player's gain is the other's loss. It is not a new game – it is described in Conan Doyle's "The Final Problem" written in 1893. In that story Sherlock Holmes is being pursued by his arch-enemy the brilliant but evil Professor Moriarity. If Holmes can escape to France he wins; if Moriarity can catch Holmes first Moriarity wins. The climactic conclusion of the story finds Holmes on a train bound for Dover and Moriarity pursuing Holmes on another train. The only stop is at Canterbury. If both get off at the same stop Moriarity catches Holmes (the "pennies" match) and Moriarity wins. If they get off at different stops Holmes wins. Despite the supposed brilliance of Holmes and Moriarity, their creator Conan Doyle was not a terribly good game theorist – in the story Holmes reasons that Moriarity thinks he is going to Dover, so he gets off at Canterbury while Moriarity continues to Dover and loses the game. But why does not the supposedly brilliant mathematician Moriarity understand Holmes reasoning so get off at Canterbury himself? And why does not Holmes anticipating this get off at Dover? Despite the fact that we can repeat this logic endlessly there is a Nash equilibrium – it necessarily requires that players choose randomly. If each has a 50% chance of getting off at Canterbury or Dover, then each has a 50% chance of winning the game no matter what the other player does.

Does that sound realistic? Choosing randomly? The problem of evading capture does not occur only in novels. The best selling book ever released by the RAND Corporation is their 1955 table of random numbers. Folklore has it that at least one captain of a nuclear submarine kept it by his bedside to use in plotting evasive maneuvers. More familiar are sporting events. The soccer player kicking a penalty goal must keep the goal keeper in the dark about whether he will kick to the right or the left of the goal; the tennis player must be unpredictable as to which side of the court she will serve to, the football quarterback must not allow the defense to anticipate run or pass, or whether the play will move to the right or the left, and the baseball catcher must keep the batter uncertain as to how his pitcher will deliver the ball. Indeed, at one time in Japan catchers were equipped with small mechanical randomization devices with which to call the pitch – this was later ruled unsporting and banned from play.

In 2001 – in a paper published in the leading journal in economics – Holt and Goeree studied several variations of Matching Pennies in the laboratory. In the first variation the payoffs were 80 for the winner and 40 for the loser. As in other versions of matching pennies the only Nash equilibrium is for players to randomize 50-50 – and indeed, unlike Holmes and Moriarity – they did just that. The table below shows the theoretical Nash equilibrium of 50% and in parentheses the actual fraction of subjects that chose the corresponding row and column. As you can see it is quite close to 50%.

	50% (48%)	***50% (52%)***
50% (48%)	80,40	40,80
50% (52%)	40,80	80,40

This type of randomization is called a *mixed strategy* equilibrium. Fifty-fifty is a particularly easy strategy to implement, and even though Conan Doyle couldn't figure it out the experimental participants did. However, the theory of mixed strategy equilibrium is peculiar in that it predicts that each player must randomize so as to make his opponent indifferent. This implies that in a mixed strategy equilibrium each player's play depends only on his opponents payoffs and not on his own. This can be counterintuitive.

To study randomization Holt and Goeree changed the payoffs by increasing (from 80 to 320) or decreasing (from 80 to 44) the payoff to Player 1 in the upper left corner. In theory this should change Player 2's equilibrium play, but Player 1 should continue to randomize 50-50. The two tables below show the theoretical predictions of Nash equilibrium and in parentheses what actually happened: far from continuing to randomize 50-50 Player 1 played the row containing the highest payoff at least 92% of the time.

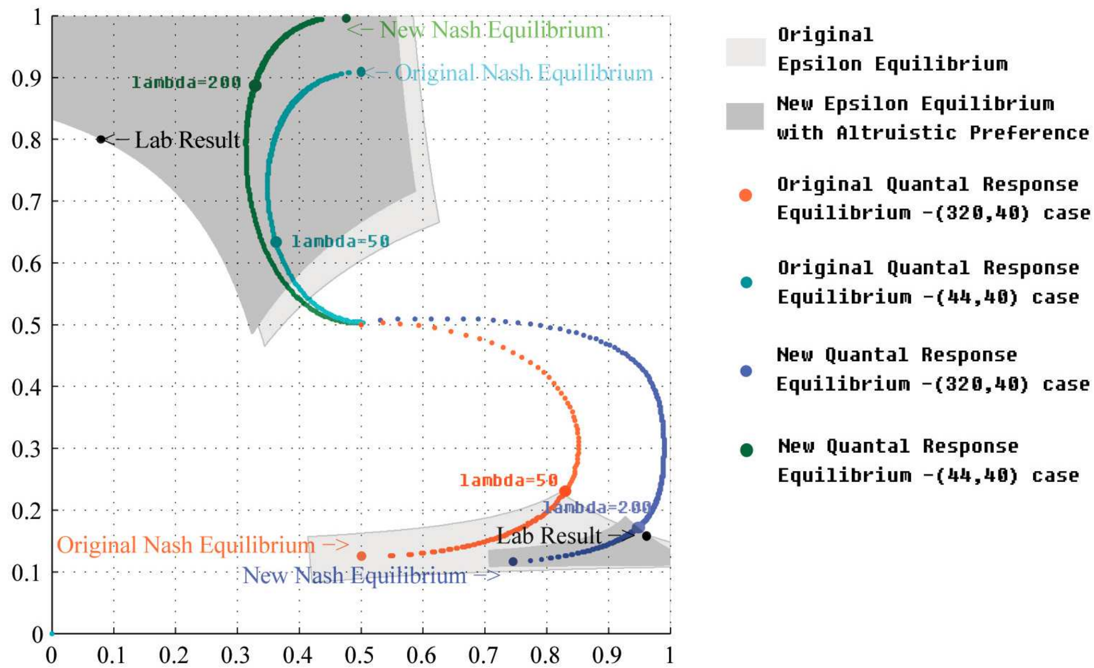
	12.5% (16%)	***87.5% (84%)***
50% (96%)	320,40	40,80
50% (4%)	40,80	80,40

	87.5% (80%)	***12.5% (20%)***
50% (8%)	44,40	40,80
50% (92%)	40,80	80,40

As is the case with some of the earlier experiments, the theory here does about as badly as it can: the theory predicts equal probability between the two rows, but the actuality is that one row is played pretty much all the time. However: unlike the other experiments this one involves players who are inexperienced in the sense that they only got to play the game once. From the perspective of learning theory there is no reason we should expect to see a Nash equilibrium. Never-the-less it is interesting to see how well our theoretical tools work in understanding what happened.

The figure below is taken from Levine and Zheng [2010]] and illustrates our main concepts. The vertical axis is the frequency with which Player 1 chooses the *Top* row; the horizontal axis the frequency with which Player 2 chooses the *Left* column. The laboratory results are shown by the black dots labeled *Lab Result* with the upper left dot corresponding to the second matrix – the 44 game, and the lower right dot corresponding to the first matrix – the 320 game. The theoretical prediction of Nash equilibrium – that Player 1 (and only Player 1) randomizes 50-50 – are labeled as *Original Nash Equilibrium*.

We consider several different ways of weakening the theory of selfish Nash equilibrium. The first is by computing all the approximate equilibrium in which the losses are no greater than



those actually suffered by the participants. This is the light gray shaded region. The second is by computing the QRE corresponding to differing levels of noisy decision making. These are the light blue and red curves that begin at the respective Nash equilibria and – as decision making becomes more noisy – move eventually towards the completely random outcome where both players simply make each choice with equal 50% probability. The dark gray region and the green and dark blue curves also examine approximate and QRE – but do so under the hypothesis that players are altruistic.

To understand what this diagram does and does not show, it is useful to start with QRE. One prediction of quantal response is a tendency toward the middle. For example in the 320 game Player 2 plays *Left* in Nash equilibrium 12.5% of the time. Quantal response says that errors in play will push that towards the middle – toward a 50-50 randomization, and indeed we see that in actuality 16% rather than 12.5% of Player 2's play *Left*. This in turn has a substantial impact on the incentives of Player 1: with "too many" player 2's playing *Left*, the best thing for Player 1 to do is to play *Top* and try to get the 320 – and again this is what we see participants do. We see it also in the diagram. As we vary the parameter of noisy choice away from Nash equilibrium and perfect best response we see that QRE play shifts towards to the right – towards the lab result with more Player 1's playing *Top*. Similarly in the 44 game, "too many" player 2's play *Right* – 20% rather than 12.5% – and this tilts the Player 1's towards playing *Down*. Again, the initial effect of increasing the noise parameter is to move the QRE towards the lab result.

Eventually, when the noise becomes too great, QRE approaches a pure 50-50 randomization. What the diagram also shows is that this happens "too soon" in the sense that play in the QRE "starts back" towards 50-50 before it gets to the laboratory result. That effect is much more pronounced in the 44 game than the 320 game.

Next consider altruism. This is potentially important in the 320 game since Player 2 by giving up 40 can increase the payoff of Player 1 by 280 – you don't have to be that generous to take such an opportunity. This also can explain why "too many" Player 2's play *Left*. If we assume a combination of errors due to quantal response and some altruistic players, it turns out we can explain the 320 game quite well, as the curve combining the two effects passes more or less directly through the laboratory result.

In the 44 game the situation is different. Even combining altruistic players with quantal response errors quantitatively we can explain only about half the laboratory result. Here the approximate equilibrium regions can help us understand what is going on. Notice that in the 320 game the approximate equilibrium region while wide is not very tall. While there are many possible strategies by player 1 that are consistent with a relatively small loss, there are very few strategies by player 2: Player 2 must play *Right* with between about 10% and 20% probability. On the other hand, in the 44 game approximate equilibrium indicates we can say little beyond Player 1 should play *Top* more frequently than *Bottom* and Player 2 should play *Right* more frequently than *Left*. The reason for this is not hard to fathom In the 320 game incentives are relatively strong: by making a wrong choice players can lose between 40 and 280. In the 44 game by making a wrong choices player can lose between 4 and 40. Naturally when incentives are less strong the set of approximate equilibrium is larger and we are less able to make accurate predictions of how players will play.

Finance Theory and Noise Traders

The notion of approximate equilibrium, especially in the form of QRE, is widely used in experimental economics. But has it taken root in mainstream economics? In the analysis of real economic problems? Like most tools in economics it is applied by economists where it is relevant – where there is empirical and conceptual reason to think that it is important. No where is this more true than in the theory of information in financial markets – and here, in the form of noise traders – it a key tool of analysis.

Central to any theory of financial markets is the extent to which they are “informationally efficient,” meaning how well do they incorporate information available to investors about economic circumstances. In a world in which you cannot fool anybody ever the tiniest bit of information would typically be revealed nearly instantaneously – leading to the conundrum that nobody could profit from inside information, and so nobody would bother to acquire any in the first place.

On the other hand – you surely can fool some of the people some of the time – and this idea far from being ignored by economists is the foundation of the modern theory of information in financial markets. It originates in modern form in the dissertation of Anat Admati, published in 1985 in *Econometrica* the leading journal in economic theory. The idea was picked up by Fischer Black. Black’s description of noise traders – the small but important irrational component of the market – was published in 1986, and Google assures us there have been some 1328 follow-on papers. Black is hardly an obscure figure: he avoided joining his co-author Myron Scholes on the stand to receive the Nobel prize in Economics by the time honored tradition of dying too soon. In the event, it would be ridiculous to assert as many commentators do that the central finding in modern finance theory is that markets are informationally efficient.

Conclusion

The chapter started with a quote attributed to Abraham Lincoln: “you can not fool all of the people all of the time.” By way of contrast modern rational expectations theory seems to say "you can not fool anybody ever." Are economists fools for being slavish disciples of so ridiculous a doctrine? We are not. Modern economic theory is much closer to Abraham Lincoln's point of view than it is to the popular caricature of rational expectations. Approximate equilibrium, quantal response equilibrium and the introduction of noise traders are all widely used methods designed to admit into rational expectations theory the idea that small irrationalities abound. It is fair to say that the basis of modern economics is that most people are rational most of the time. This is far from a slavish devotion to a ridiculous doctrine – it well captures the spirit of Abraham Lincoln.