Exp Econ (2007) 10:105–122
DOI 10.1007/s10683-006-9135-z

Mental processes and strategic equilibration: An fMRI study of selling strategies in second price auctions

**David M. Grether · Charles R. Plott ·
Daniel B. Rowe · Martin Sereno · John M. Allman**

Received: 23 June 2004 / Revised: 16 May 2006 /
Accepted: 16 May 2006 / Published online: 13 February 2007

Abstract This study is the first to attempt to isolate a relationship between cognitive activity and equilibration to a Nash Equilibrium. Subjects, while undergoing fMRI scans of brain activity, participated in second price auctions against a single competitor following predetermined strategy that was unknown to the subject. For this auction there is a unique strategy that will maximize the subjects' earnings, which is also a Nash equilibrium of the associated game theoretic model of the auction. As is the case with all games, the bidding strategies of subjects participating in second price auctions most often do not reflect the equilibrium bidding strategy at first but with experience, typically exhibit a process of equilibration, or convergence toward the equilibrium. This research is focused on the process of convergence.

In the data reported here subjects participated in sixteen auctions, after which all subjects were told the strategy that will maximize their revenues, the theoretical equilibrium. Following that announcement, sixteen more auctions were performed. The question posed by the research concerns the mental activity that might accompany equilibration as it is observed in the bidding behavior. Does brain activation differ between being equilibrated and non-equilibrated in the sense of a bidding strategy? If so, are their differences in the location of activation during and after equilibration? We found significant activation in the frontal pole especially in Brodmann's area 10, the anterior cingulate cortex, the amygdala and the basal forebrain. There was significantly more activation in the basal forebrain and the anterior cingulate cortex during the first

---

D. M. Grether · C. R. Plott (✉) · J. M. Allman
California Institute of Technology
e-mail: cplott@hss.caltech.edu

D. B. Rowe
Medical College of Wisconsin

M. Sereno
University of California San Diego

sixteen auctions than in the second sixteen. The activity in the amygdala shifted from the right side to the left after the solution was given.

Keywords Experiments · Auctions

JEL Classification D440, L620, L810, C930, C900

1 Introduction

The development of tools to measure reactions associated with mental activity opens the possibilities of fruitful interactions between economists and biologists, especially neuroscientists. This study is the first to join that interface with a focus on equilibration. Economic theory is typically a theory of equilibrium and equilibration for individuals, markets or, more generally, systems. While both market models and game theory models have both equilibrium and equilibration defined in terms of observable variables, the theory itself, when applied to the individual, is often developed in terms of the unobservable beliefs and objectives that the individual might hold. Thus, the theory is often developed from assumptions about unobservable states of mind along with the use of "as if" assumptions, to which theorists often resort when confronted by unobservables.

The application of the resulting "as if" methodology is widespread, well understood and remarkably successful in developing models for the behavior of both individuals and complex systems of individuals. Included are postulates of optimizing behavior and strategy, as well as postulates about the consequences of strategies that others might employ. Intentions, attributions of intentions and strategic thinking, none of which can be observed, play a role in the theory. At the base of the success of the theory, in spite of the lack of observability of key variables, rests a theory of equilibrium and equilibration. It is that particular feature of this complex that we focus on here. Can we observe mental activities associated with choice behaviors that are interpreted as equilibration in terms of economic models? Can the mental activities be interpreted as part of equilibration in an economic sense? Can the "as if" propositions be pushed to deeper points of analysis and physical phenomena?

Developments in the use of laboratory techniques in experimental economics together with the technological advancements of fMRI make possible collaboration between economics and neuroscience (Glimcher, 2003; McCabe, 2001; Smith et al., 2002) that might narrow the scope for speculation about what is taking place in the mind as economic decisions are made. Some may argue that economists are not interested in which areas of the brain are active during different tasks and while the foundation for such interest is yet to be proved we feel that the elements of such a foundation are obvious. Economic theory contains many instances in which understanding and motivation play a role. Other regarding preferences in public goods environments, sources of valuation in contingent valuation studies, attitudes of fairness and reciprocity in exchange situations, spite and punishment play a role in enforcement. These are all motivations that become intertwined with preferences in models. Concepts of purpose and intent surface in the context of collusion and coordination. Concepts of risk as opposed to ambiguity create divisions in theories of behavior under uncertainty. One does not have to look deeply in the literature to find theories based on preference types, such as liquidity traders as opposed to insiders in theories of finance. These are wide classes of discussions that proceed on a foundation that might be made stronger if the locations of brain activity can lead us to a better identification of motivation and intention.

Distinguishing between equilibrium behavior and out of equilibrium behavior could have important implications for the study of economics. Economic theory contains many models of equilibria and for those models observing behavior of individuals who are not equilibrated does not provide a valid test of the theory. Similarly, the theory contains models of convergence and dynamics but in markets containing many people "non movement of process, quantities and other economic magnitudes" are sometimes hard to detect due to the randomness of uncoordinated behavior of agents that might be in equilibrium individually. Knowing which data represent equilibrium behavior and which do not would be a significant breakthrough. The implications of commonly observed behavior that appears inconsistent with economic theory (Ellsberg, 1961), are quite different if the behavior is equilibrium behavior as opposed to confused responses of subjects searching for a solution (Plott, 1996). The heuristics and biases literature (Kahneman et al., 1982) , for example, contains several well documented and replicated examples of behavior apparently inconsistent with expected utility theory. Are these behaviors the results of "disequilibrium seat of the pants" responses to unfamiliar tasks or do these behavior reflect equilibrium responses of individuals not well described by neoclassical economic theory? Identifying the different types of behavior from mental activity could help to resolve this and possibly other issues. Furthermore, as questions posed for fMRI studies advance the methodology itself will have a need to become informed about whether activations in areas of the brain are stabilized in some sense in response to specific cognitive processes or are transient reflecting aspects of disequilibrium.

In this paper we take an admittedly tentative first step towards understanding equilibrium versus out of equilibrium actions at the individual level of analysis. We report the results of experiments conducted while subjects were inside an MRI scanner. Subjects participated in a series of auctions in which the optimal response is to accurately reveal preferences, which seem obvious once understood. Indeed, with training which sometimes includes explaining the optimal strategy, subjects typically adopt the strategy and in some cases do so only after a few trials. In the experiments conducted, optimal and non-optimal strategies are easily observed and then can be interpreted as distinguishing between equilibrium and disequilibrium behavior. Such decisions are made while the subject is undergoing fMRI so the differences in mental process can be observed during and after equilibration in an economic sense.

2 Experimental procedures

2.1 The task

The subject is given a coupon with a face value stated in the experimental currency called francs. Francs were converted to dollars at the end of the experiment at a fixed known exchange rate (300f = \$1). The subject may keep the coupon and redeem it for cash at the end of the experiment or sell it and earn the sales price instead. The subject's task is to state a reservation price, i.e. the lowest price at which the subject is willing to sell the coupon. That price will be compared with a bid randomly generated independently of the subject's price. If the random bid is above the reservation price, the subject sells the coupon at the higher price. Otherwise the subject keeps the coupon. The experimental task is based upon the method developed by Becker et al. (1964) for determining subjects' valuations of objects. We refer to this procedure as BDM. We used francs rather than direct cash values so that if we alter the design and change the number of tasks in a session the exchange rate can be adjusted to keep subjects earnings at a reasonable level. In practice, subjects earned about $70 for the session, which ran about an hour and a half.

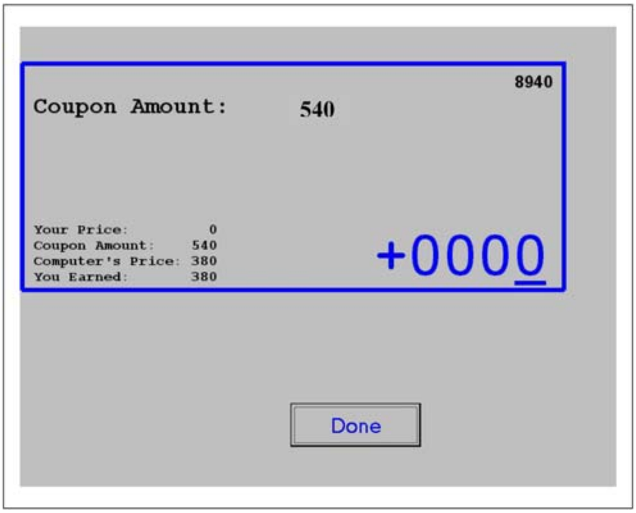
If one wants to maximize rewards, the optimal strategy is to be "truthful" in the sense of bidding exactly the amount of the coupon. Thus, if the subjects name an amount less than the value to them of the coupon, they risk selling it for less than the value. On the other hand, if they name an amount that is more than the coupon is worth to them, there is the possibility of rejecting bids that they would prefer to having the coupon. By naming any amount other than the true value of the coupon they cannot gain anything, but may lose something.

Regardless of the simplicity, major elements of this problem are present in almost every economic decision. First, is a notion of independence between subjects' bids and the bids to which theirs will be compared. If the bids are somehow related, then subjects would have an incentive to place bids with an objective of influencing the future bids used by the experimenter. Secondly, there is a notion of optimizing behavior from the point of view of the subject, as opposed to the set of all bidders. A subject who is concerned about being "fair" or concerned about appearance could behave much differently from theoretical predictions. Third, is the recognition that one strategy is dominant: it is optimal regardless of the outcome of the randomness. Recognition of dominance is an exercise in abstract logic that leads to the conclusion of the type of strategy that is optimal. Such interactions among motive, understanding, logic, beliefs and strategy have always been a challenge to economists, who traditionally have only observed behaviors as data.

Subjects make sixteen bidding decisions and after each are informed of the amount of money they accumulate as a result. After the first sixteen auctions, subjects were told the correct strategy, that is, the strategy that will ensure them the largest amount of money. Instructions are attached in the Appendix. These were read prior to any decision or experience in the machine. Once in the machine, the screens they saw are typical of the example in Fig. 1.The Instructions about the optimal strategy, which were as presented after the first sixteen auctions, are attached at the end of the general instructions. The objective of the research reported here is to study the subjects while mental activity is taking place both before and after the instructions about the optimal strategy are given to them. Presumably, if the decisions of the subject are not equilibrated before the instructions are given to them, they would be so afterward. The research questions we address are twofold. What locations in the brain show activation in the task? Are there systematic differences in brain activation before and after subjects have found the solution to the problem?

2.2 Session procedures

Subjects were recruited primarily from the upper level, undergraduate student population of University of California at San Diego. However, the subjects did



8940

Coupon Amount: 540

Your Price: 0
Coupon Amount: 540
Computer's Price: 380
You Earned: 380

+0000

Fig. 1 Sample display

include graduate students, a postdoctoral fellow and a working adult. Each subject was first given a safety orientation and, if willing, signed the consent form. Next the subject read the Keypad Instructions for operating the specially designed non-metallic fiber optic button response unit (BRU) and then practiced with the BRU while seated in front of a laptop with the keypad to the right of the computer. There are no metal parts that could cause interference. The unit has four buttons that can be used to correspond to any four keys on the keyboard. In these experimental sessions, the buttons on the BRU corresponded to: Left, Right, Up and Down. Numbers were displayed on the screen as four digit numbers with all columns initialized at zero. The Right (Left) button caused the cursor to move one column to the right (left) with a wrap around to the far left (right) if the cursor was already in the rightmost (leftmost) column. The Up (Down) button increased (decreased) the digit in the active column by one. All buttons had analogous wrap around features. When the subject was comfortable with the BRU the subject was given the selling instructions. At this point, any questions about the procedures were answered.

The subject was then positioned in the scanner with the BRU taped to the right side and the right arm taped on the BRU to alleviate the need to hold the arm still. The projected output from the laptop was aligned on a screen so that the subjects can see the information on the experimental tasks projected into the scanner. Once the subject was properly placed in the scanner and able to read information displayed the subject is given a second practice session practice with the BRU under the new condition of lying down in the scanner. As the positioning in the scanner takes several minutes and could be distracting, the instructions for the bidding task are summarized on a screen projected into the scanner (See Appendix instructions). At this point, the experiment, which consists of four scanning session is conducted. The data were collected from a block design experiment lasting 512 seconds in which participants were given eight trials of stimuli A, B, and C. Task A consisted of the participant viewing simple blue text from gray screen, determining a number, and entering the number using a button response unit all in 22 seconds as described below. Task B consisted of the participant receiving simple blue text feedback displayed on a gray screen for 10 seconds. Task C was a control stimulus which consisted of a neutral gray blank screen. Scanning was performed using a 1.5 Tesla Siemens Magneton in which 24 interleaved axial slices of size 64×64 were acquired using a single shot gradient echo pulse sequence with full Cartesian k-space coverage having a $\text{FOV} = 19.2$ cm, $\text{FA} = 90^\circ$, and $\text{TE} = 40$ ms. After Fourier image reconstruction, each voxel has dimensions in mm of 3×3 in plane and 5 mm through plane resulting in a volume of 45 cubic mm. Observations were taken every 4 seconds or $\text{TR} = 4000$ ms so that there are 128 in each voxel. A custom program was written that alternately displayed a Bidding Task (A and B) and a blank screen (C). The program also reads the subject's input from the BRU and records each keystroke, the time elapsed, the bids and earnings. The program includes a display calibration screen to align the projector and a Raw Display Program, which reads and displays the Raw Data.

3 The statistical model

The data are panel data with the unit of observation being x (i, t), the hemodynamic response of voxel i at time t. The basic model is (Rowe, 2001, 2003):

$$x(i,t) = a(i) + b(i)^{*}t + c(i,1)^{*}S(1,t) + \cdots + c(i,k)^{*}S(k,t) + \varepsilon(i,t). \quad (1)$$

In the equation, the $S\,(k, t)$ are the values of K possibly unobserved stimuli often called reference functions that the voxels are responding to. Each voxel has its own intercept and trend as well as coefficients for contributions from the k responses to the stimuli or reference functions. We included the trend term to deal with a possible confound between time in the scanner and activity and other potential sources of drift (Smith et al., 1999). We take the disturbances, $\varepsilon\,(i, t)$ to be matrix normal with contemporaneous covariance matrix Ψ and uncorrelated over time.

If the reference functions are unobserved, the model is similar to a factor analysis model. In factor analysis it is usually assumed that the factors (reference functions in this application) are orthogonal. We do not impose that condition here, so the model is perhaps better described as a latent variable model.

Classical statistical methods for the analysis of fMRI data assume a generic prespecified hemodynamic response termed a reference function (Bandettini et al., 1993; Friston, et al., 1994). But the choice of an *a priori* known and fixed reference function may not be appropriate. Much effort has been devoted to incorporating information from individual subjects in the determination of a reference function that is correlated with the time series in each voxel for activation (Frank et al., 1998; Kershaw and Ardekani, 1999; Dishbrow et al., 2000; Genovese, 2000; Gossl et al., 2001; Friston et al., 2002a,b). Usual practice is to estimate differences in response by introducing the boxcar functions (so-called due to their shape) taking the value 1 during a treatment and -1 or 0 when the treatment is not on, shifting them to allow for the hemo-dynamic lag and smoothing or tapering them to reduce discontinuities. The estimates are usually interpreted as giving the differences in the average response between the treatment being on or off, though the transformations and smoothing make the interpretation more complicated.

We proceed in a Bayesian fashion, and quantify available knowledge in the form of priors for the models parameters. The coefficients of the reference functions are taken to be jointly normal with covariance matrix Ψ, and mean matrix C_0 (the natural conjugate prior, see Rowe (2003, pp. 207–211). An uninformative prior is used for the trend and intercept coefficients. The reference functions are taken to be matrix normal with prior mean $S_{0,}$ the columns of which are the assumed to be boxcar functions representing the experimental treatments, and covariance matrix R.

The Bayesian statistical approach utilized here determines the underlying response or reference function. The reference function is viewed as the underlying response due to the presentation of the experimental tasks. The reference function need not be assumed to be known or fit into the standard on/off or rise/fall format and it may change (possibly nonlinearly) over the course of the experiment. Instead of subjectively choosing a known reference function as in standard GLM procedures, prior information as to its value is quantified. This prior information is combined with the data and a reference function is determined statistically using the information contributed from every voxel. In the Bayesian approach, all the voxels contribute to "telling us" the underlying response due to the presented experimental stimulus. In practice, we do not know the true underlying time response function. The coefficients of the reference functions have the same interpretations as coefficients in familiar regression models. That is, a significant coefficient means that the expected response varies significantly with change in the reference function. For a fuller discussion, see Rowe (2000, 2001). For an elementary but elegant discussion of Bayesian inference, see Edwards et al. (1963). The Bayesian approach to the analysis of fMRI data is becoming accepted as providing useful methods for analyzing such data. This is evidenced by the number of recently published manuscripts including those by the Karl Friston (Friston et al., 2002a,b). Further, the widely distributed software SPM has incorporated Bayesian analysis (see, Friston and Penny (2003)).

The variance covariance matrices, Ψ and R, are taken to have inverted Wishart distributions with parameters $(\nu,\ Q)$ and $(\eta,\ V)$ respectively. The number of components was initially taken to be one, and later we allowed for two reference functions. We use an empirical Bayes approach and take the prior means for the coefficients of the reference functions to be the estimated coefficients of the voxels' activations on boxcar functions. We obtain the conditional densities of the parameters given the values of the other parameters and the hyper parameters. We estimate the mode of the posterior distribution by the method of iterated conditional modes (see Lindley and Smith (1972) and Rowe (2001) for details), and we obtain our estimates of the reference function (s) from the posterior mode. We proceed conditional upon the reference functions that we estimated from the mode of the posterior distribution. That is, we look at the relationship between the hemodynamic responses in voxels and the statistically determined Bayesian reference functions, using the Bayesian reference functions in place of the boxcar functions. The reference functions are in all cases significantly correlated with the boxcar functions, generally tracking them, but with extra detail and some shifts. Rather than shifting the boxcar functions for the hemodynamic lag, this approach allows the data to dictate the shifts.

The procedures for calculating the statistical significance of the activation were based on the AFNI program and related software. For each scan, we computed the histogram of the t statistics for the full brain and picked the value that corresponded to a fixed percentile of the distribution (typically 95%). We defined regions of interest (ROIs) prior to the analysis, and only used regions we could identify without reference to activation maps. For the frontal pole, for example, we took to all of the brain forward of minus 55 (Talairach) on the y-axis. We identified all clusters of voxels in ROIs such that at each voxel in the cluster the t statistic was at least as large as the cut off level. Voxels were defined as being in the same cluster if their centers were no more than 5.7 mm apart. We calculated the bandwidths of the Gaussian filter corresponding to the spatial correlations in the data using the program 3dFWHM. We simulated the clusters in the ROIs using areas of roughly the same configurations as the ROIs. In the simulations samples of independent standard normal variables are generated and smoothed with the Gaussian filter estimated from the data. Each simulation was preformed 10,000 times and the number of clusters tabulated by size of cluster. Rather than perform separate simulations for each scan, we performed the simulations using the maximal bandwidths from the four scans for each subject. As larger bandwidths lead to more clusters this procedure is conservative.

The procedures for calculating the differences in activation were similar. The time series for scans one and two were added and the series for scans three and four subtracted from them. Once these calculations were performed, the methods were the same as those described.

4 Results

The results are partitioned into two subsections. The first is focused on the bidding behavior as viewed from the point of view of economics models. The second subsection is related to mental activity as rendered from the fMRI.

4.1 Bidding results

Previous experiments suggest that the task we chose is one in which subjects tend to spend some time developing a strategy and the strategy once developed tends to be game theoretically correct. After developing such a strategy the typical subject tends to use it with little or no change. While we could never be sure of what might be taking place in the mind we can observe choices and we can use conventions for determining equilibrium behavior in terms of choice. In that sense it is possible to observe brain activity during a non-equilibrated period and a subsequent equilibrated period where the latter are understood in the sense of economic theory. The first result of this section demonstrates that equilibration processes, as seen by economic theory, were observed on average. Table 1 shows the time structure of absolute difference between bid and the value of the ticket for each subject. Compared are the first sixteen auctions with the last sixteen. Equilibration occurs when the difference is zero. For no subject is the median lower in the first sixteen and for most it is higher. The data can be viewed a different way in Fig. 2. Shown there are the eight bids for each subject during each

Table 1 Absolute value of the bid minus the value of the ticket

	Auctions 1 to 16				Auctions 17 to 32			
Subject	Mean	Std. Dev.	Median	Number ≤ 10	Mean	Std. Dev.	Median	Number ≤ 10
1	419.1	111.5	399	0	388	94.2	0	13
2	.1	.3	0	16	.2	.5	0	16
3	83.9	92.0	65	5	5.9	4.4	5	14
4	23.3	84.5	1	15	1	0	1	16
5	73.8	76.3	60	1	59.4	24.9	60	0
6	0.0	0.0	0	16	0	0	0	16
7	379.2	145.5	400	1	0	0	0	16
8	187.3	97.9	210	0	15.7	28.0	5	11
9	14.1	30.1	0	12	0	0	0	16
10	78.8	35.8	80	1	580	2245.6	0	12[a]
11	6.3	3.6	0	16	0	0	0	16
12	31.9	29.9	25	5	25.0	30.6	10	8
13	263.0	161.7	278	1	129.7	125.9	95	2
14	29.4	55.4	7.5	9	0	0	0	16
15	51.6	22.5	50	0	50	26.7	50	0
16	91.9	107.0	61.5	3	0	0	0	16
17	167.1	224.4	17	7	53.8	215.0	0	15

[a] Deleting one observation at 9000 reduces the mean and standard deviation to 18.7 and 38.9, respectively.

of the four sessions plotted against the value of the ticket for which the subject was bidding. Equilibrium theory predicts that all bids should lie on the 45 degree line. Subjects were told the equilibrium strategy at the start of the third series of eight. The equilibration tendencies are obvious in the data.

Result 1. Overall, subjects exhibited convergence toward the equilibrium strategy with more rapid convergence exhibited after the domination features of the equilibrium strategy were explained.

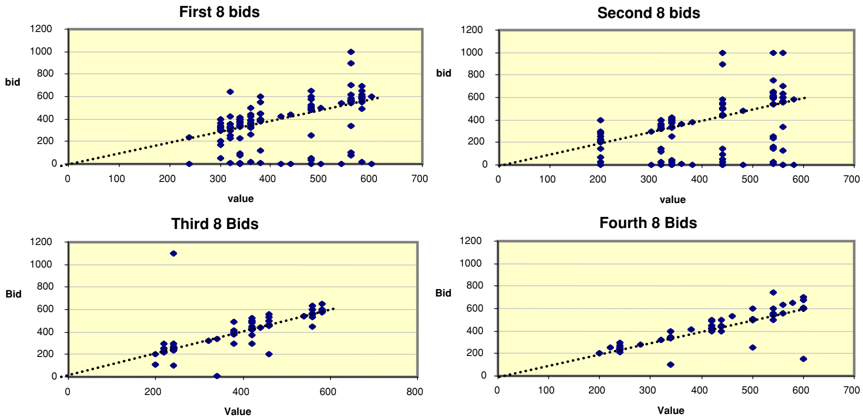


Fig. 2 Pooled bids of all individuals

Support. Each of the 17 subjects participated in 32 auctions, producing a total of 544 bids. Of these, 241 were exactly equal to the value of the coupon. There were 75 bids equal to coupon value in the first 16 trials and 166 during the second 16 trials. All coupon values and the predetermined bids were in multiples of ten (subjects were not told this). Thus, any bid within 10 of the coupon value earned the same amount as bidding the value exactly. More bids, mainly from the second 16 trials (203 compared to 108) satisfied this criterion. The magnitudes of the differences were not large in general. Averaging over all subjects, the average differences were 42.3 points for the first eight trials and 47.9 for the second eight trials. The differences for the second half of the session were much smaller: 0.5 points for the last eight trials and 9.1 for the third eight (ignoring one outlier at 9000).

The figure illustrates the existence of variability across subjects. One concern was that the task would be so obvious to the subjects that they would see the solution immediately and indeed for some subjects that was the case as is summarized by the next result and support.

Result 2. Convergence properties differ across the behavior of individuals.

Support. Of the seventeen subjects who participated in scans, four nearly always used the dominant strategy from the very beginning (subjects 2, 4, 6 and 11). An additional three subjects (subjects 3, 9 and 14) converged to the dominant strategy without being told for a total of seven subjects that consistently used the dominant strategy without being given the solution. In addition, seven subjects followed the correct strategy after being given the solution. Even after being given the solution, subjects frequently experimented a bit before adopting the dominant strategy. All but three subjects adopted the dominant strategy by the end of the experiment (two consistently priced over the value and one priced below). Initial behavior, however, was heterogeneous. Six subjects began by pricing over the value, three consistently set prices below the value, four gave prices both over and under the value and four used the dominant strategy from the beginning.

The source of the observed variability differs from individual to individual and during the first phases appears to be related to strategic behavior based on misunderstandings or confusion about the full implication of various strategies. The misguided strategy seems to come in two forms. Some individuals place a higher bid believing they will receive a higher price not realizing that in a second price auction they are exposed to an opportunity for loss. Other subjects offer at low prices explaining the strategy by saying that they wanted to make the sale. Both properties can be seen in the first panels of Fig. 2. Notice two separate groupings of data, one with a tendency to be above the 45° line of the optimal strategy and the other much below the line. On average, there was a greater tendency to over bid with 205 prices set above coupon value and 95 set below.[1] Notice that in the lower panels the two groups have merged into one as subjects tended to adopt the optimal strategy.

---

[1] Explanation of the tendency to over bid involves some controversy. Kahneman et al. (1991) interpret this as an inducement effort", on the other hand Plott and Zeiler (2003) suggest that it results from subject's misunderstandings of the task.

Those subjects who did not adopt the dominant strategy of giving the coupon value as their reservation price seemed to adopt some simple alternative e.g. rounding up the coupon value to the next multiple of 100, or alternating above and below. One subject alternated by being over the coupon value by 30 points and by under 70 points.

A scanner is hardly an ideal decision making environment. The subjects are motionless on their backs and in a tube looking at a screen through mirrored glasses. The button response unit (BRU) is taped to them, and the scanner, when active, makes a lot of noise. Given this, the overall performance is reasonable. Pilot experiments with subjects who were not in scanners exhibited similar convergence.

4.2 Results from fMRI: Mental activity measurements

The brain areas activated in these experiments include the frontal polar cortex, which is known from previous studies to be involved in calculation, and a group of brain structures that receive input from the dopaminergic neurons in the midbrain that signal expectation of reward. These structures include the basal forebrain, the amygdala and the anterior cingulate cortex.

Two additional results emerge from the brain scans. The first of these (Result 3) identify the location in which significant activities were observed. The second (Result 4) summarizes the changes in mental activity across the treatments.

Each subject participated in four sets of bids each eight minutes and thirty-two seconds long. The sequence was twenty-two seconds during which the value of the coupon was displayed (see supporting materials for the display) and the subject entered their reservation price, ten seconds for feedback, followed by thirty-two seconds of blank screen. For technical reasons we were unable to analyze the MRI data from the first 6 subjects; the analysis is based on subject 7 through 17.

Result 3. Significant activation occurs in four areas: the frontal pole (especially Brodmann's Area 10), the amygdala, the anterior cingulate cortex and the basal forebrain.

Support. The activations are summarized in Table 2. The table gives the estimated significance levels for the largest cluster of voxels in the indicated regions all of which exceeded a predetermined cutoff level. Frequently, there were multiple clusters found which were significant at conventional levels of significance. The significance levels shown are, therefore, conservative. As stated above, the anterior cingulate activity was greater in the first sixteen auctions, which is shown in Fig. 3, a sagittal view (at $x = 0$) depicting the difference between the intensity over the first 16 rounds minus rounds 17–32. The right amygdala was more active in the first 16 rounds than the left (see Fig. 4).

Result 4. Activation levels changed over the course of the treatments. The anterior cingulate activity was greater during the first (during theoretical equilibration) trials as was the right amygdala. During the second sixteen trials (after theoretical equilibration), amygdalar activity shifted to the left side and the anterior cingulate activity reduced.

Support. The right amygdala was more active in the first 16 rounds than the left (see Fig. 2). In the final rounds, the amygdalar activity shifted to the left side. Table 3

Table 2 Significance levels of responses for the four scans

Region subject	Frontal pole	Anterior cingulate cortex	Basal forebrain	Left amygdala	Right amygdala
7	a a c a		a a a a	- - - a	c d
	a c d c		d a - b		- - - d
8	a b a c	- - d	a a a a	b b b c	
	a a a b	b	b b a	c - b	- a
9	c d b d	- c	a a a a	- c	b - a c
	- a a a	- a a	a a - a	- a c	c - c a
10	a a a a	c b a d	c a a a	- - d b	a b
	a c a a	- c - b	- a a a	- a b a	- a
11	c a a a		b a d c	a a b c	
	a a a a		d - d b	a c c a	
12	a a a a	a - - b	a - - c	b - b	b - c
	a a a a	a - b	a b a d	- c	c - c
13	a a a a	- a d	- b		
	a c a a	- - a	a d c	c	- c - c
14	a a a a	- c - a	- - a a		
	a a a a	- - b a	- d a a		- - - b
15	c a a	a c		- c c	
	a a a a	- b	c a c a	c - a c	
16	a a a a	- - a	- a - b	c - a	- - a
	a a a a		c a d		
17	a a a	a	a b - c		d - c
	b a a a	- a a a	a d a a		- - d

The letters in each cell give the significance levels for the four scans in order.
The first row gives significance levels for the first component, and the second row gives significance levels for the second component.
Legend: a $p < .0005$; b $p < .005$; c $p < .05$; d $p < 0.1$; blank or - $p > .10$

gives the significance levels for the differences in the responses between the first two scans (before subjects were given the solution to the problems) and the last two scans (after the solution was given). Table 4 contains the same information for the scans immediately before and after subjects were given the solution (scans 2 and 3). In the tables we report significance levels for the largest clusters of each sign (a minus sign indicates the cluster is significant with a negative sign indicating greater activation in the later scans).

5 Discussion

The results reported here suggest the existence of a correspondence between the phenomenon of equilibration in decision rules as it is interpreted by economic theory and the physiological traces of mental processing. Such a correspondence holds a potential for joining two sciences in useful ways. On one hand, as already noted, a concept of equilibrium is central to much of economic theory. The theory interprets behavior "as if" certain mental processes were taking place and introspection about those processes often serves to motivate new theory. In the absences of direct observations of mental activity, competing interpretations of observed behavior are difficult to resolve.

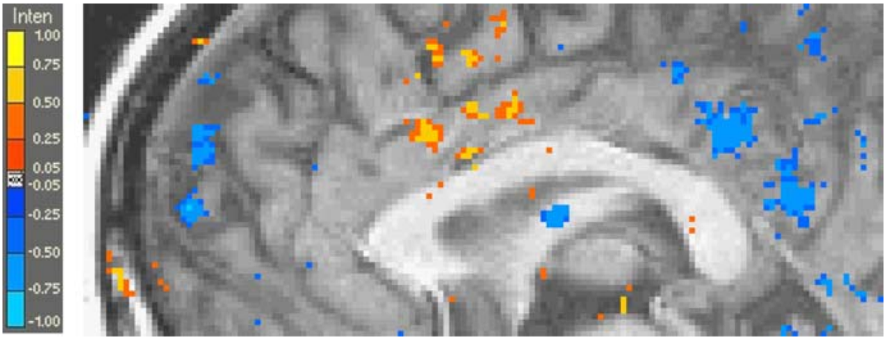


Fig. 3 Average response for the first 16 rounds minus the average for rounds 17–32. Note the activity in the anterior cingulate cortex is greater in the first 16 rounds. Sagittal view at $x = 0$. Features on the right side of the brain are shown in the right side of the figure

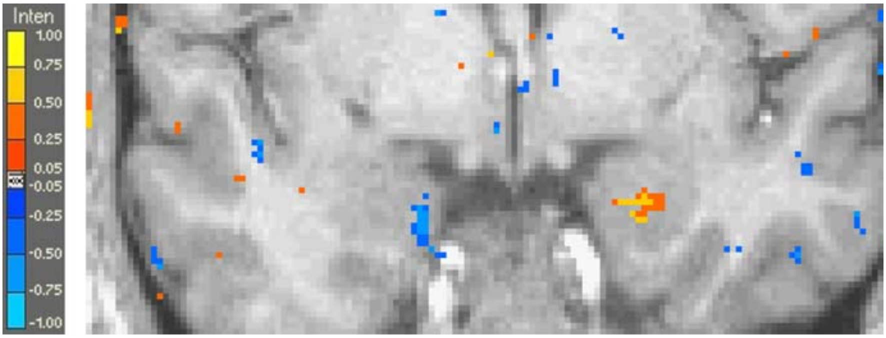


Fig. 4 Average response for the first 16 rounds minus the average for rounds 17–32. Note the activity is greater on the right amygdala in for the first 16 rounds, and shifts to the left side for rounds 17–32. Coronal view at $y = 1$. Features on the right side of the brain are shown in the right side of the figure

For example, an individual who is confused and exhibits unchanging behavior while resolving a decision problem could be interpreted as having attained equilibrium. Furthermore, measurements of behavior tend to be limited to those that are suggested by the theory, which is itself conditioned on an assumption that mental processes cannot be observed. Thus, for economics the discovery suggests the possibility of enriching the set of observable variables that might be interpreted by theory. Similarly, on the other hand, the discovery connects neuroscience with a new set of behaviors and suggests new interpretations of the areas of the brain in which mental activities are taking place.

The task was chosen so that we expected to image brain activity during a non-equilibrated or searching stage and after equilibration in the sense of economic theory. To this end, the task was not perfect. If a subject happens to adopt the equilibrium strategy, by accident or without realizing that it is indeed a correct strategy, there was no way no way to tell. In our experiment, for example, it is possible that subjects could respond with the face value of the coupon because it is a salient number and not because they had determined that this was in fact the best possible response. Though a few subjects adopted the optimal strategy without being given the solution, all subjects were given the optimal strategy after the first two scans. We do not know if those

Table 3 Response Scans 1&2 minus 3& 4. Differences computed for each voxel

Subject	Frontal pole	Anterior cingulate cortex	Basal forebrain	Amygdala
		First component		
7	a, a-		b-	c
8	a		a	a
9	a,a-	d	a,a-	
10	a,a-		a-	a
11	a-		a-	
12	a,a-	a	c	
13	a,a-	b	a	
14				
15	a			
16	a,a-			
17	a,a-	b		
		Second component		
7	a			
8	c			
9		c	a	
10	a,a-		a-	
11	a,a-		a	
12	a,a-	a,a-	b,a-	a-
13	a,a-		a	
14	a,a-	a-		
15	a-			
16				
17	c, a-			

Legend: a $p < .0005$; b $p < .005$; c $p < .05$; d $p < .10$.
Entries with a minus sign following are for regions with negative values: i.e. the average response in scans 3 and 4 was greater than the average response in scans 1 and 2.

subjects using the optimal strategy before getting the solution actually understood the task rather than responding with the coupon value as a focal point.

With the qualifications as above, we conservatively interpreted observations from the second two scans as being equilibrated and those from the first two as being non-equilibrated. In this regard, we see more activation during the first two scans in the anterior cingulate and the right amygdala than during the later scans. Thus, we interpret the phenomena of disequilibration with the anterior cingulate, which dampens with the process of equilibration. The findings and interpretation are consistent with other findings reported in the literature.

The strongest and most consistent activation was located near the frontal pole, primarily in Brodmann's area 10 (Table 5). This area has been previously been found to be active in other tasks involving monetary reward and risk assessment (Rogers et al., 1999; Ernst et al., 2001). Area 10 is also activated in mental calculation (Rickard, 2000) which the subjects would have been performing in this task. Area 10 is activated when an expected financial reward is received (Knutson et al. 2003). Area 10 is both absolutely and relatively much larger in humans than in other primates which suggests that some of the circuitry related to economic decision-making is a phylogenetic specialization in the human brain (Allman et al., 2002; Semendeferi et al., 2001).

Table 4 Response Scan 2 minus Scan 3. Differences computed for each voxel

Subject	Frontal pole	Anterior cingulate cortex	Basal forebrain	Amygdala
		First component		
7	c		a	
8	a			
9	b,c-	d	d,a-	b
10	a,a-			
11	a		a	
12	a,a-		a-	
13	a,a-		a	
14				
15	a,a-			
16				
17	a,a-	a	a	
		Second component		
7	a,		a	a
8	a		a	b,b-
9		a		
10	a,a-		a,a-	
11	a,a-		a-	
12	a-		a-	
13	c-			d
14	a,a-		a,a-	a
15	a,a-			
16	a,a-			
17	d, a-	a-	c-	

Legend: a $p < .0005$; b $p < .005$; c $p < .05$; d $p < .10$.
Entries with a minus sign following are for regions with negative values: i.e. the average response in scan 3 was greater than the average response in scan 2.

The anterior cingulate cortex, basal forebrain, and amygdala all receive strong dopaminergic input from the midbrain, which signals the expectation of reward (Schultz, 2002). The anterior cingulate was active during the first 16 auctions before the subjects were told the optimal bidding strategy and less active during the second 16 auctions. The anterior cingulate cortex is active in difficult tasks involving considerable uncertainty (Critchley et al., 2001). The anterior cingulate cortex is also active in tasks that require novel as opposed to routine solutions (Raichle et al., 1994). The anterior cingulate cortex is the source of an EEG signal associated with financial loss in gambling (Gehring and Willoughby, 2002), which would presumably not occur with the application of the optimal bidding strategy. More broadly speaking, the anterior cingulate cortex is involved in maintaining physiological homeostasis through the regulation of a wide variety of mechanisms necessary for life including hunger and thirst. Liotti et al. (2001). This broader view of a homeostatic role for anterior cingulate is consistent with the lack of activation in these structures during economic decision making during equilibrium conditions.

The basal forebrain contains many reward-related neurons (Schultz, 2000). Activity in the ventral striatal component of the basal forebrain is specifically related to cumulative financial reward during a series of gambles (Elliott et al., 2003). The amygdala

Table 5 Area 10 significant activation

		Component 1				Component 2			
Scan		1	2	3	4	1	2	3	4
Subject	7				•				
					a				
	8						•		
							a, a		
	9			•			•		•
							a		b
	10	•		•	•	•	•	•	•
		a		a	a	a	a	a	a
	11	•	•	•	•	•	•	•	•
		c	a, b	a	a	a	a	a	a
	12	•	•	•	•	•	•	•	•
		a	a	a	a	a, a	a	a	a
	13	•	•	•	•	•		•	•
		a	a	a	a	a		a, a	a
	14	•	•		•	•	•		•
		a	a		a	a	a		a
	15	•	•	•		•	•	•	•
		b	a	a		a	a	a	a
	16	•	•	•	•	•	•	•	•
		a	a	a	a	a	a	a, a	a
	17	•	•	•		•	•	•	•
		b	a	a		a	a	a	a

A dot indicates that part of Area 10 is contained in a cluster that is significant at the level indicated based on simulations of clusters in the frontal pole. Legend: $p < .0005$; b $p < .005$; c $p < .05$; d $p < .10$. Multiple letters refer to separate clusters of voxels.

is strongly associated with fear (Ledoux, 1995) and it is possible that the initial activation on the right side during the disequilibrium phase may have arisen from the subjects' fear of financial loss. However, the activation of the left amygdala during the second sixteen auctions after the subjects had been informed of the optimal strategy may be associated with their increased certitude of receiving financial reward. There is increasing evidence that amygdala is involved in positive reward assessment as well as its more familiar role in fear (Baxter and Murray, 2002; Schultz, 2002). Recently, Fried et al. (2001) have shown through micro dialysis that the dopaminergic input to the amygdala in human subjects increases during cognitive tasks not involving fear.

The fact that the location of neural activity changes as equilibration takes place means that in imaging studies, equilibration and time must be considered in addition to features of the task. Clearly there is much to be done. In pilot studies, we included control tasks, simple calculations requiring response on the BRU, and found that subjects were distracted by the second task, forgetting their experience in the earlier auctions. Thus, we did not include control tasks because of the expected interference in subjects' finding the optimal strategy in the selling task. A natural extension would be to bring subjects back after a few days to redo the task. Assuming that they then understand the task, we could use these extra trials to compare their now (presumed to be) equilibrated behavior with the scans from the earlier sessions. At this time it should be able to introduce other control tasks. Other measurements could be made simultaneously. For example, we could measure respiration, heart rate, skin conductivity and eye movement could all be monitored to look for correlations with possible equilibration. We view this work as only the first step in what could be a long term project to understand the relations between physiological responses and equilibration.

The task studied involves numerical responses. While this is true of many economic decisions, it is by no means true of all. Choices from non-numerically valued alternatives are basic to economics, and could also be studied. The task studied has a single individual acting alone with alternative bids predetermined by a computer. Bargaining, or any type of other regarding behavior is not included in the situation studied. A natural extension would be to have subjects buying and selling in markets with other participants.

Acknowledgment The research support of the Caltech Laboratory for Experimental Economics and the National Science Foundation are gratefully acknowledged. The paper was presented at the Conference at the University of Minnesota Conference on Experimental Research in 2000 and circulated as a working paper in April 2001.

References

Allman, J. M., Watson, K., & Hakeem, A. (2002). Two phylogenetic specializations in the human brain. *Neuroscientist*, *8*, 335–347.

Bandettini, P. M., Jesmanowicz, A., Wong, E. C., & Hyde, J. S. (1993). Processing strategies for time-course data sets in functional MRI of the human brain. *Magnetic Resonance in Medicine*, *30*, 161–173.

Becker, G. M., DcGroot, M. H., & Marshak, J. (1964). Measuring utility by a single response sequential method. *Behavioral Science*, *9*, 226–232.

Frank, L. R., Buxton, R. B., & Wong, E. C. (1998). Probabilistic analysis of functional magnetic resonance imaging data. *Magnetic Resonance in Medicine*, *39*, 132–148.

Baxter, M. G., & Murray, E. A. (2002). The amygdala and reward. *Nature Reviews, Neuroscience*, *3*, 563–573.

Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron*, *29*(2), 537–545.

Dishbrow, E. A. et al. (2000). Functional MRI at 1.5 tesla: A comparison of the blood oxygenation level-dependent signal and electrophysiology. *Proceedings of the National Academy of Sciences*, *97*, 9718–9723.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review*, *70*, 193–242.

Elliott, R., Friston, K., & Dolan, R. (2000). Dissociable neural responses in human reward systems. *Journal of Neuroscience*, *20*, 6159–6165.

Elliott, R., Newman, J., Longe, O., & Deakin, J. (2003). Instrumental responding for rewards is associated with enhanced neuronal response in subcortical reward systems. *Neuroimage*, *21*(3), 984–990.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axions. *Quarterly Journal of Economics*, *75*, 643–669.

Ernst, M. et al. (2001). Decision-making in a risk-taking task: a PET study. *Neuropsychopharmacology*, *26*, 682–691.

Fried, I. et al. (2001). Increased dopamine release in the human amygdala during performance of cognitive tasks. *Nature Neuroscience*, 201–206.

Friston, K. J., Jezzard, P., & Turner, R. (1994). Analysis of functional MRI time-series. *Human Brain Mapping*, *18*, 153–171.

Friston, K. J., & Penny, W. (2003). Posterior probability maps and SPMs. *NeuroImage*, *19*, 1240–1249.

Friston, K. J., Penny, W. D., Phillips, C., Kiebel, S. J., Hinton, G., & Ashburner, J. (2002a). Classical and Bayesian inference in neuroimaging: theory. *NeuroImage*, *16*, 465–483.

Friston, K. J., Glaser, D. E., Henson, R. N. A., Kiebel, S. J., Phillips, C., & Ashburner, J. (2002b) Classical and Bayesian inference in neuroimaging: applications. *NeuroImage*, *16*, 484–512.

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*, 2279–2282.

Genovese, C. R. (2000). A Bayesian time course model for functional magnetic resonance imaging data. *Journal of the American Statistical Association*, *95*, 691–703.

Glimcher, P. W. (2003). *Decisions, uncertainty and the brain: the science of neuroeconomics*. MIT Press.

Gossl, C., Fahrmeir, L., & Auer, D. P. (2001). Bayesian modeling of the hemodynamic response function in BOLD fMRI. *NeuroImage*, *14*, 140–148.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the coarse theorem. *The Journal of Political Economy*, *98*, 1325–1348.

Kahneman, D., Slovic, P., & Tversky, A. (eds.) (1982). *Judgment under uncertainty: heuristics and biases*. Cambridge University Press.

Kershaw J., & Ardekani, B. A. (1999). Application of Bayesian inference to fMRI data analysis. *IEEE Transactions on Medical Imaging*, *18*, 1138–1153.

Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M., & Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *Neuroimage*, *18*(2), 263–272.

LeDoux, J. (1995). *The emotional brain*. New York: Simon & Schuster.

Lindley, D. V., & Smith, A. F. M. (1972). Bayes estimates for the linear model. *Journal of the Royal Statistical Society*, *34*, 1–19.

Liotti, M., Brannan, S., Egan, G., Shade, R., Madden, L., Abplanalp, B., Robillard, R., Lancaster, J., Zamarripa, F., Fox, O., & Denton, D. (2001). Brain responses associated with consciousness of breathlessness (air hunger). *Proceeding of the National Academy of Sciences*, *98*, 2035–2040.

McCabe, K., Houser, D., Ryan, L., Smith, V., & Tromard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences*, *98*(20), 11832–11835.

Plott, C. R. (1996). Rational individual behavior in markets and social choice processes: the discovered preference hypothesis. In K. J. Arrow, E. Colombatto, M. Perlman and C. Schmidt (eds.), *The rational foundations of economic behavior.*

Plott, C. R., & Zeiler, K. (2003). The willingness to pay/willingness to accept gap, the ‘endowment effect,’ subject misconceptions and experimental procedures for eliciting valuations. *American Economic Review* (submitted).

Raichle, M., Feiz, J. A., Videen, T. O., MacLeod, A. M., Pardo, J. V., Fox, P. T., & Petersen, S. E. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex*, *4*, 8–26.

Rickard, T., Romero, S., Basso, G., Wharton, C., Flitman, S., & Grafman, J. (2000). The calculating brain: an fMRI study. *Neuropsychologia*, *38*, 325–335

Rowe, D. B. (2001). Bayesian source separation for reference function determination in fMRI. *Magnetic Resonance in Medicine*, *46*, 374–378.

Rowe, D. B (2003). *Multivariate Bayesian statistics: models for source separation and signal unmixing*. Chapman and Hall: CRC Press.

Russo, E. (2000). Debating the meaning of fMRI. *The Scientist*, *14*, 18–20.

Schultz, et al. (2000). Reward processing in primate orbital frontal cortex and basal ganglia. *Cerebral Cortex*, *10*, 272–284.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *36*, 241–263.

Semendeferi, K., Armstrong, E., Schleicher, A., Zilles, K., & Van Hoesen, G. (2001). Prefrontal cortex in humans and apes: a comparative study of area 10. *American Journal of Physical Anthropology*, *114*, 224–241.

Smith, A. et al. (1999). Investigation of low frequency drift in fMRI signal. *Neuroimage*, *9*, 526–533.

Smith, K. et al. (2002). Neuronal substrates for choice under ambiguity, risk, certainty, gains and losses. *Management Science*, *48*, 711–718.